Multi–colored rooted tree analysis for Runge–Kutta methods for the weak approximation of stochastic differential equations

Andreas Rößler

Darmstadt University of Technology, Fachbereich Mathematik, Schlossgartenstr. 7, D-64289 Darmstadt, Germany

Abstract

A general class of stochastic Runge–Kutta methods for Itô and Stratonovich stochastic differential equation systems with a multi–dimensional Wiener process is considered. The multi–colored rooted tree analysis is applied to calculate order conditions for the coefficients of explicit and implicit stochastic Runge–Kutta methods assuring convergence in the weak sense with a prescribed order. Especially, order conditions and some coefficients for stochastic Runge–Kutta schemes of weak order two are calculated explicitly.

Key words: stochastic Runge–Kutta method, stochastic differential equation, multi–colored rooted tree analysis, weak approximation, numerical method
MSC: 65C30, 65L06, 60H35, 60H10

1 Introduction

In recent years, derivative free Runge–Kutta type schemes have been proposed for the strong approximation of stochastic differential equations (SDEs), see e.g. [5,8,9,14]. Burrage and Burrage [1,2] introduced colored trees for the calculation of order conditions for stochastic Runge–Kutta methods for strong approximation. However, for the weak approximation of SDEs particular schemes have to be developed, see e.g. [5,7,8,10–12,15,16]. Komori, Mitsui and Sugiura [6] applied colored trees in order to calculate coefficients for ROW–type schemes for the weak approximation of Stratonovich SDEs with

Email address: roessler@mathematik.tu-darmstadt.de (Andreas Rößler).

an one-dimensional Wiener process.

The aim of the present paper is to calculate order conditions for a general class of stochastic Runge–Kutta (SRK) methods for the weak approximation of Itô and Stratonovich SDEs with a multi–dimensional Wiener process. Therefore, the multi-colored rooted tree analysis for weak approximation introduced in [10,12,13] is applied. It provides an unified and very efficient way to determine order conditions for SRK methods for both, Itô and Stratonovich SDEs. As an example, order conditions up to order two are calculated for the introduced class of explicit or implicit SRK methods by multi-colored trees and some coefficients for such schemes are presented.

Let $(\Omega, \mathcal{F}, P)$ be a 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 consider the solution $(X_t)_{t \in \mathcal{I}}$ of either a d-dimensional Itô or Stratonovich SDE system

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

for $d, m \geq 1$ and $t \in \mathcal{I}$, where the jth column of the $d \times m$-matrix function $b = (b^{i,j})$ is denoted by b^j for $j = 1, \ldots, m$. The second integral w.r.t. the Wiener process has to be interpreted either as an Itô integral with $*\mathrm{d}W_s^j = \mathrm{d}W_s^j$ or as a Stratonovich integral with $*\mathrm{d}W_s^j = \circ \mathrm{d}W_s^j$. Let $X_{t_0} = x_0 \in \mathbb{R}^d$ be the $\mathcal{F}_{t_0}$-measurable initial value such that for some $l \in \mathbb{N}$ holds $E(\|X_{t_0}\|^{2l}) < \infty$ where $\|\cdot\|$ denotes the Euclidean norm if not stated otherwise. Here, $W = ((W_t^1, \ldots, W_t^m))_{t \geq 0}$ is an m-dimensional Wiener process w.r.t. $(\mathcal{F}_t)_{t \geq 0}$. We suppose that the conditions of the Existence and Uniqueness Theorem are fulfilled for SDE (1) (see, e.g., [4]).

The solution $(X_t)_{t \in \mathcal{I}}$ of a Stratonovich SDE with drift a and diffusion b is also solution of an Itô SDE with the same diffusion b, however with the modified drift $\tilde{a}^i(t,x) = a^i(t,x) + \frac{1}{2} \sum_{j=1}^d \sum_{k=1}^m b^{j,k}(t,x) \frac{\partial b^{i,k}}{\partial x^j}(t,x)$ for $i = 1, \ldots, d$, provided b is sufficiently differentiable. In the following, let $C_P^l(\mathbb{R}^d, \mathbb{R})$ denote the space of all $g \in C^l(\mathbb{R}^d, \mathbb{R})$ with polynomial growth, i.e. there exists a constant $C > 0$ and $r \in \mathbb{N}$, such that $|\partial_x^i g(x)| \leq C(1 + \|x\|^{2r})$ for all $x \in \mathbb{R}^d$ and any partial derivative of order $i \leq l$ [5].

Let a discretization $\mathcal{I}_h = \{t_0, t_1, \ldots, t_N\}$ with $t_0 < t_1 < \ldots < t_N = T$ of the time interval $\mathcal{I} = [t_0, T]$ with step sizes $h_n = t_{n+1} - t_n$ for $n = 0, 1, \ldots, N-1$ be given. Further, let $h = \max_{0 \leq n < N} h_n$ denote the maximum step size.

Definition 1.1 *A time discrete approximation process Y converges weakly with order p to the solution process X of SDE (1) as $h \to 0$ if for each $f \in C_P^{2(p+1)}(\mathbb{R}^d, \mathbb{R})$ there exists a constant C_f, which does not depend on h,* *and a finite* $\delta_0 > 0$ *such that*

$$|E(f(X_t)) - E(f(Y(t)))| \leq C_f \, h^p \tag{2}$$

holds for each $h \in]0, \delta_0[$ *and* $t \in \mathcal{I}_h$.

2 Stochastic Runge–Kutta Methods

We consider a very general class of stochastic Runge–Kutta methods which has been introduced in [10]: Let $\mathcal{M}$ be an arbitrary finite set of multi-indices with $\kappa = |\mathcal{M}|$ elements and let $\theta_\nu(h)$, $\nu \in \mathcal{M}$, be some suitable random variables. For the weak approximation of the solution $(X_t)_{t \in \mathcal{I}}$ of SDE (1), considered either with respect to Itô or Stratonovich calculus, a general class of s-stage stochastic Runge–Kutta methods is then given by

$$\begin{aligned} Y_0 &= x_0 \\ Y_{n+1} &= Y_n + \sum_{i=1}^{s} z_i^{(0,0)} \; a\left(t_n + c_i^{(0,0)} h_n, H_i^{(0,0)}\right) \\ &\quad + \sum_{i=1}^{s} \sum_{k=1}^{m} \sum_{\nu \in \mathcal{M}} z_i^{(k,\nu)} \; b^k \left(t_n + c_i^{(k,\nu)} h_n, H_i^{(k,\nu)}\right) \end{aligned} \tag{3}$$

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

$$\begin{aligned} H_i^{(k,\nu)} &= Y_n + \sum_{j=1}^{s} Z_{ij}^{(k,\nu),(0,0)} \; a\left(t_n + c_j^{(0,0)} h_n, H_j^{(0,0)}\right) \\ &\quad + \sum_{j=1}^{s} \sum_{r=1}^{m} \sum_{\mu \in \mathcal{M}} Z_{ij}^{(k,\nu),(r,\mu)} \; b^r \left(t_n + c_j^{(r,\mu)} h_n, H_j^{(r,\mu)}\right) \end{aligned}$$

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

$$\begin{aligned} z_i^{(0,0)} &= \alpha_i \, h_n & \qquad z_i^{(k,\nu)} &= \sum_{\iota \in \mathcal{M}} \gamma_i^{(\iota)\,(k,\nu)} \, \theta_\iota(h_n) \\ Z_{ij}^{(k,\nu),(0,0)} &= A_{ij}^{(k,\nu),(0,0)} \, h_n & \qquad Z_{ij}^{(k,\nu),(r,\mu)} &= \sum_{\iota \in \mathcal{M}} B_{ij}^{(\iota)\,(k,\nu),(r,\mu)} \, \theta_\iota(h_n) \end{aligned}$$

for $i, j = 1, \ldots, s$ and let $\alpha_i, \gamma_i^{(\iota)\,(k,\nu)}, A_{ij}^{(k,\nu),(0,0)}, B_{ij}^{(\iota)\,(k,\nu),(r,\mu)} \in \mathbb{R}$ be the coefficients of the SRK method. As usual, the weights can be defined by

$$c^{(0,0)} = A^{(0,0),(0,0)} e, \qquad c^{(k,\nu)} = A^{(k,\nu),(0,0)} e, \tag{4}$$

with $e = (1, \ldots, 1)^T$. If $A_{ij}^{(k,\nu),(0,0)} = B_{ij}^{(\iota)\,(k,\nu),(r,\mu)} = 0$ for $j \geq i$ then (3) is called an explicit SRK method, otherwise it is called implicit. We assume that the random variables $\theta_\nu(h_n)$ satisfy the moment condition

$$E\left(\theta_{\nu_1}^{p_1}(h_n)\cdot\ldots\cdot\theta_{\nu_\kappa}^{p_\kappa}(h_n)\right)=O\left(h_n^{(p_1+\ldots+p_\kappa)/2}\right) \tag{5}$$

for all $p_i \in \mathbb{N}_0$ and $\nu_i \in \mathcal{M}$, $1 \leq i \leq \kappa$. The moment condition ensures a contribution of each random variable having an order of magnitude $O(\sqrt{h})$. This condition is in accordance with the order of magnitude of the increments of the Wiener process.

Remark that for a deterministic ordinary differential equation, i.e. SDE (1) with $b \equiv 0$, the SRK method reduces to the well known deterministic Runge–Kutta method, so the introduced class of SRK methods turns out to be a generalization of deterministic Runge–Kutta methods.

3 Colored Rooted Tree Analysis

Following the approach in [10,13] (see also [12]), we give a definition of colored trees which will be suitable for SDEs w.r.t. a multi–dimensional Wiener process. Since each SDE system can be represented by an autonomous SDE system

$$X_t = X_{t_0} + \int_{t_0}^{t} a(X_s)\,\mathrm{d}s + \sum_{j=1}^{m}\int_{t_0}^{t} b^j(X_s)\, * \mathrm{d}W_s^j \tag{6}$$

with one additional equation representing time, we restrict our considerations to an autonomous SDE system in this section.

Definition 3.1 *(1) A monotonically labelled* S-tree (stochastic tree) $\boldsymbol{t}$ *with* $l = l(\boldsymbol{t}) \in \mathbb{N}$ *nodes is a pair of maps* $\boldsymbol{t} = (\boldsymbol{t}', \boldsymbol{t}'')$ *with*

$$\boldsymbol{t}' : \{2,\ldots,l\} \to \{1,\ldots,l-1\}$$
$$\boldsymbol{t}'' : \{1,\ldots,l\} \to \mathcal{A}$$

so that $\boldsymbol{t}'(i) < i$ *for* $i = 2,\ldots,l$. *Unless otherwise noted, we choose the set* $\mathcal{A} = \{\gamma, \tau, \sigma_{j_k}, k \in \mathbb{N}\}$ *where* j_k *is a variable index with* $j_k \in \{1,\ldots,m\}$.

(2) LTS denotes the set of all monotonically labelled S-trees w.r.t. $\mathcal{A}$. *Here two trees* $\boldsymbol{t} = (\boldsymbol{t}', \boldsymbol{t}'')$ *and* $\boldsymbol{u} = (\boldsymbol{u}', \boldsymbol{u}'')$ *just differing by their colors* $\boldsymbol{t}''$ *and* $\boldsymbol{u}''$ *are considered to be identical if there exists a bijective map* $\pi : \mathcal{A} \to \mathcal{A}$ *with* $\pi(\gamma) = \gamma$ *and* $\pi(\tau) = \tau$ *so that* $\boldsymbol{t}''(i) = \pi(\boldsymbol{u}''(i))$ *holds for* $i = 1,\ldots,l$.

So $\mathbf{t}'$ defines a father son relation between the nodes, i.e. $\mathbf{t}'(i)$ is the father of the son i. Furthermore the color $\mathbf{t}''(i)$, which consists of one element of the set $\mathcal{A}$, is added to the node i for $i = 1,\ldots,l(\mathbf{t})$. Here, $\tau = \bullet$ is a deterministic node, $\sigma_{j_k} = \bigcirc_{j_k}$ is a stochastic node with a variable index $j_k \in \{1,\ldots,m\}$ and $\gamma = \otimes$ is the root of the tree. The variable index j_k is associated with the j_kth component of the corresponding m-dimensional Wiener process of the considered SDE. As an example Figure 1 presents two elements of LTS.

In the following, we denote by $d(\mathbf{t}) = \sharp\{i : \mathbf{t}''(i) = \tau\}$ the number of de-

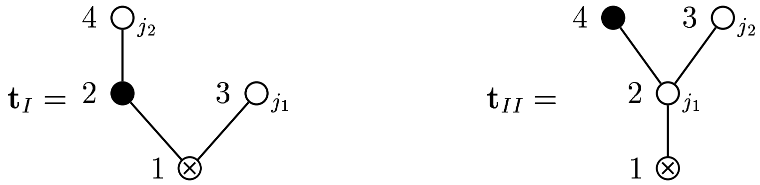


Fig. 1. Two elements of LTS with $j_1, j_2 \in \{1, \ldots, m\}$.

terministic nodes and by $s(\mathbf{t}) = \sharp\{i : \mathbf{t}''(i) = \sigma_{j_k}, k \in \mathbb{N}\}$ the number of stochastic nodes. The order $\rho(\mathbf{t})$ of the tree $\mathbf{t}$ is defined as $\rho(\mathbf{t}) = d(\mathbf{t}) + \frac{1}{2}s(\mathbf{t})$ with $\rho(\gamma) = 0$. The order of the trees $\mathbf{t}_I$ and $\mathbf{t}_{II}$ presented in Figure 1 can be calculated as $\rho(\mathbf{t}_I) = \rho(\mathbf{t}_{II}) = 2$.

Every labelled tree can be written by a combination of three different brackets: If $\mathbf{t}_1, \ldots, \mathbf{t}_k$ are colored trees then we denote by $(\mathbf{t}_1, \ldots, \mathbf{t}_k)$, $[\mathbf{t}_1, \ldots, \mathbf{t}_k]$ and $\{\mathbf{t}_1, \ldots, \mathbf{t}_k\}_j$ the tree in which $\mathbf{t}_1, \ldots, \mathbf{t}_k$ are each joined by a single branch to $\otimes$, $\bullet$ and $\bigcirc_j$, respectively. Therefore proceeding recursively, for the two examples $\mathbf{t}_I$ and $\mathbf{t}_{II}$ in Figure 1 we obtain $\mathbf{t}_I = ([\bigcirc^4_{j_2}]^2, \bigcirc^3_{j_1})^1 = ([\sigma^4_{j_2}]^2, \sigma^3_{j_1})^1$ and $\mathbf{t}_{II} = (\{\bullet^4, \bigcirc^3_{j_2}\}^2_{j_1})^1 = (\{\tau^4, \sigma^3_{j_2}\}^2_{j_1})^1$.

Now, two labelled trees $\mathbf{t}, \mathbf{u} \in LTS$ are called equivalent, i.e. $\mathbf{t} \sim \mathbf{u}$, if they are identical except for their monotonically labels. The set of all equivalence classes under the relation $\sim$ is denoted by $TS = LTS/\sim$. We denote by $\alpha(\mathbf{t})$ the cardinality of $\mathbf{t}$, i.e. the number of possibilities of monotonically labelling the nodes of $\mathbf{t}$ with numbers $1, \ldots, l(\mathbf{t})$.

For example, the labelled trees $([\sigma^3_{j_1}]^2, \sigma^4_{j_2})^1$, $([\sigma^4_{j_2}]^2, \sigma^3_{j_1})^1$ and $(\sigma^2_{j_3}, [\sigma^4_{j_8}]^3)^1$ belong to the same equivalence class as $\mathbf{t}_I$ in the example above, since the indices j_1 and j_2 are just renamed either by j_2 and j_1 or j_8 and j_3, respectively. Finally the graphs differ only in the labelling of their number indices.

For every rooted tree $\mathbf{t} \in TS$, there exists a corresponding *elementary differential*. The elementary differential is defined recursively by $F(\gamma)(x) = f(x)$, $F(\tau)(x) = a(x)$ and $F(\sigma_j)(x) = b^j(x)$ for single nodes and by

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

for a tree $\mathbf{t}$ with more than one node. Here $f^{(k)}$, $a^{(k)}$ and $b^{j\,(k)}$ define a symmetric k-linear differential operator, and one can choose the sequence of labelled

S-trees $\mathbf{t}_1, \ldots, \mathbf{t}_k$ in an arbitrary order. For example, the Ith component of $a^{(k)} \cdot (F(\mathbf{t}_1), \ldots, F(\mathbf{t}_k))$ can be written as

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

where the components of vectors are denoted by superscript indices, which are chosen as capitals.

Definition 3.2 *For $* \in \{I, S\}$ let $LTS(*)$ denote the set of trees $\boldsymbol{t} \in LTS$ having a root $\gamma = \otimes$ and which are build by finite many steps of the form*

a) adding a deterministic node $\tau = \bullet$, or
b) adding two stochastic nodes $\sigma_{j_k} = \bigcirc_{j_k}$, where both nodes get the same new variable index j_k for some $k \in \mathbb{N}$. Additionally, in the case of $* = I$ neither of the two nodes is allowed to be the father of the other.

The nodes are labelled in the order of adding. Further $TS(*) = LTS(*)/\sim$ denotes the equivalence class under the relation $\sim$ restricted to $LTS(*)$ and $\alpha_*(\boldsymbol{t})$ denotes the cardinality of $\boldsymbol{t}$ in $LTS(*)$ for $* \in \{I, S\}$, respectively.

It holds $LTS(I) \subset LTS(S)$. For example, the tree $(\{\sigma_{j_1}^3\}_{j_1}^2, \{\sigma_{j_2}^5\}_{j_2}^4)^1$ belongs to $LTS(S)$ but not to $LTS(I)$. However, the tree $(\{\sigma_{j_2}^4\}_{j_1}^2, \{\sigma_{j_2}^5\}_{j_1}^3)^1$ belongs to $LTS(I)$. The only difference is the sequence of the construction, i.e. the correct father-son relationship for the stochastic nodes (see [10,13] for details). Now, the following Theorem holds (see Thm 3.2, Thm 4.2 and Prop 5.1 in [13]):

Theorem 3.3 *For the solution $(X_t)_{t \in \mathcal{I}}$ of SDE (6), $p \in \mathbb{N}_0$, $f : \mathbb{R}^d \to \mathbb{R}$ with $f, a^i, \tilde{a}^i, b^{i,j} \in C_P^{2(p+1)}(\mathbb{R}^d, \mathbb{R})$ for $i = 1, \ldots, d$, $j = 1, \ldots, m$ and $t = t_0 + h$ holds*

$$E^{t_0,x_0}(f(X_t)) = \sum_{\substack{\boldsymbol{t} \in TS(*) \\ \rho(\boldsymbol{t}) \leq p}} \sum_{j_1, \ldots, j_{s(\boldsymbol{t})/2}=1}^{m} \frac{\alpha_*(\boldsymbol{t})\, F(\boldsymbol{t})(x_0)}{2^{s(\boldsymbol{t})/2}\, \rho(\boldsymbol{t})!} h^{\rho(\boldsymbol{t})} + O(h^{p+1})$$

with $* = I$ in the case of Itô and $* = S$ in the case of Stratonovich calculus.

Next, we give an expansion for the approximation process $(Y(t))_{t \in \mathcal{I}_h}$ defined by the SRK method (3). For $\mathbf{t} \in TS$ let the density $\gamma(\mathbf{t})$ be defined recursively by $\gamma(\mathbf{t}) = 1$ if $l(\mathbf{t}) = 1$ and

$$\gamma(\mathbf{t}) = \begin{cases} \prod_{i=1}^{\lambda} \gamma(\mathbf{t}_i) & \text{if } \mathbf{t} = (\mathbf{t}_1, \ldots, \mathbf{t}_\lambda), \\ l(\mathbf{t}) \prod_{i=1}^{\lambda} \gamma(\mathbf{t}_i) & \text{if } \mathbf{t} = [\mathbf{t}_1, \ldots, \mathbf{t}_\lambda] \text{ or } \mathbf{t} = \{\mathbf{t}_1, \ldots, \mathbf{t}_\lambda\}_j. \end{cases}$$

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*. So for every $\mathbf{t} \in TS$ the function Φ_S is defined recursively by

$$\Phi_S(\mathbf{t}) = \begin{cases} \prod_{i=1}^{\lambda} \Phi_S(\mathbf{t}_i) & \text{if } \mathbf{t} = (\mathbf{t}_1, \ldots, \mathbf{t}_\lambda) \\ {z^{(0,0)}}^T \prod_{i=1}^{\lambda} \Psi^{(0,0)}(\mathbf{t}_i) & \text{if } \mathbf{t} = [\mathbf{t}_1, \ldots, \mathbf{t}_\lambda] \\ \sum_{\nu \in \mathcal{M}} {z^{(k,\nu)}}^T \prod_{i=1}^{\lambda} \Psi^{(k,\nu)}(\mathbf{t}_i) & \text{if } \mathbf{t} = \{\mathbf{t}_1, \ldots, \mathbf{t}_\lambda\}_k \end{cases} \tag{8}$$

where $\Psi^{(0,0)}(\emptyset) = \Psi^{(k,\nu)}(\emptyset) = e$ with $\gamma = (\emptyset)$, $\tau = [\emptyset]$, $\sigma_k = \{\emptyset\}_k$ and

$$\Psi^{(k,\nu)}(\mathbf{t}) = \begin{cases} Z^{(k,\nu),(0,0)} \prod_{i=1}^{\lambda} \Psi^{(0,0)}(\mathbf{t}_i) & \text{if } \mathbf{t} = [\mathbf{t}_1, \ldots, \mathbf{t}_\lambda] \\ \sum_{\mu \in \mathcal{M}} Z^{(k,\nu),(r,\mu)} \prod_{i=1}^{\lambda} \Psi^{(r,\mu)}(\mathbf{t}_i) & \text{if } \mathbf{t} = \{\mathbf{t}_1, \ldots, \mathbf{t}_\lambda\}_r \end{cases} . \tag{9}$$

Here $e = (1, \ldots, 1)^T$ and the product of vectors in the definition of $\Psi^{(k,\nu)}$ is defined by component-wise multiplication, i.e. with $(a_1, \ldots, a_n) \cdot (b_1, \ldots, b_n) = (a_1 b_1, \ldots, a_n b_n)$.

Definition 3.4 *Let $LTS(\Delta)$ denote the set of trees $\boldsymbol{t} = (\boldsymbol{t}', \boldsymbol{t}'') \in LTS$ w.r.t. $\mathcal{A} = \{\gamma, \tau, \sigma_{j_k} : k \in \mathbb{N}\}$ such that*

a) the root is of type $\boldsymbol{t}''(1) = \gamma$ and all other nodes are either deterministic or stochastic nodes, i.e. $\boldsymbol{t}''(i) \in \{\tau, \sigma_{j_k} : k \in \mathbb{N}\}$ for $2 \leq i \leq l(\boldsymbol{t})$,
b) all stochastic nodes own a different variable index j_k, $1 \leq k \leq s(\boldsymbol{t})$, i.e. for two different stochastic nodes $i \neq l$ holds $\boldsymbol{t}''(i) \neq \boldsymbol{t}''(l)$.

Further $TS(\Delta) = LTS(\Delta)/\sim$ denotes the equivalence class under the relation $\sim$ restricted to $LTS(\Delta)$ and $\alpha_\Delta(\boldsymbol{t})$ denotes the cardinality of $\boldsymbol{t}$ in LTS.

It holds $LTS(I) \subset LTS(S) \subset LTS(\Delta)$. Further, each tree $\mathbf{t} \in LTS(\Delta)$ has $s(\mathbf{t})$ different variable indices $j_1, \ldots, j_{s(t)}$ while a tree $\mathbf{u} \in LTS(*)$, $* \in \{I, S\}$, has only $s(\mathbf{u})/2$ different variable indices. Then it holds (see Prop 6.1 in [10]):

Proposition 3.5 *Let $(Y(t))_{t \in \mathcal{I}_h}$ be defined by the SRK method (3). Assume that for the random variables holds $\theta_\iota(h) = \sqrt{h} \cdot \vartheta_\iota$ for $\iota \in \mathcal{M}$ with a bounded random variable ϑ_ι. Then for $f : \mathbb{R}^d \to \mathbb{R}$, $p \in \mathbb{N}_0$ and $t = t_0 + h$ holds*

$$E^{t_0,x_0}\left(f\left(Y(t)\right)\right) = \sum_{\substack{\boldsymbol{t} \in TS(\Delta) \\ \rho(\boldsymbol{t}) \leq p + \frac{1}{2}}} \sum_{j_1, \ldots, j_{s(\boldsymbol{t})}=1}^{m} \frac{\alpha_\Delta(\boldsymbol{t})\, \gamma(\boldsymbol{t})\, F(\boldsymbol{t})(x_0)\, E\left(\Phi_S(\boldsymbol{t})\right)}{(l(\boldsymbol{t}) - 1)!} + O(h^{p+1})$$

provided $f, a^i, b^{i,j} \in C_P^{2(p+1)}(\mathbb{R}^d, \mathbb{R})$ for all $i = 1, \ldots, d$ and $j = 1, \ldots, m$.

4 Order Conditions for Stochastic Runge–Kutta Methods

In order to calculate order conditions for the random variables and the coefficients of the SRK method (3), the approximation Y has to be uniformly bounded. Therefore, we assume that each random variable can be expressed as $\theta_\iota(h) = \sqrt{h} \cdot \vartheta_\iota$ for $\iota \in \mathcal{M}$ with a bounded random variable ϑ_ι. We remark, that this condition is not necessary in the case of explicit SRK methods (see [10,12] for details). Further, we assume that

$$E\Big(z^{(k,\nu)^T} e\Big) = 0 \tag{10}$$

holds for $1 \leq k \leq m$ and $\nu \in \mathcal{M}$. Then the approximation Y by the SRK method (3) has uniformly bounded moments (see Prop 6.2 in [10]).

Definition 4.1 *Let $|\boldsymbol{t}|$ denote the tree which is obtained if the nodes σ_{j_i} of $\boldsymbol{t}$ are replaced by σ, i.e. by omitting all variable indices. Let a tree $\boldsymbol{t} \in TS(*)$ for $* \in \{I,S\}$ with variable indices $j_1, \ldots, j_{s(\boldsymbol{t})/2}$ be given and let $\boldsymbol{u} \in TS(\Delta)$ with variable indices $\hat{j}_1, \ldots, \hat{j}_{s(\boldsymbol{u})}$ denote the tree which is equivalent to $\boldsymbol{t}$ except for the variable indices, i.e. $|\boldsymbol{t}| \sim |\boldsymbol{u}|$ with $s(\boldsymbol{t}) = s(\boldsymbol{u})$. For a fixed choice of correlations of type $j_k = j_l$ or $j_k \neq j_l$, $1 \leq k < l \leq s(\boldsymbol{t})/2$, between the indices $j_1, \ldots, j_{s(\boldsymbol{t})/2}$, let $\beta(\boldsymbol{t})$ denote the number of all possible correlations between the indices $\hat{j}_1, \ldots, \hat{j}_{s(\boldsymbol{u})}$ of tree $\boldsymbol{u}$ such that $\boldsymbol{t} \sim \boldsymbol{u}$ holds. In the case of $s(\boldsymbol{t}) = 0$ or $\boldsymbol{t} \in TS(\Delta) \setminus TS(*)$, $* \in \{I,S\}$, define $\beta(\boldsymbol{t}) = 1$.*

For example, for $\mathbf{t} = (\sigma_{j_1}, \sigma_{j_2}, \{\sigma_{j_2}\}_{j_1}) \in TS(I)$ and $\mathbf{u} = (\sigma_{\hat{j}_1}, \sigma_{\hat{j}_2}, \{\sigma_{\hat{j}_4}\}_{\hat{j}_3}) \in TS(\Delta)$, two cases have to be considered. On the one hand we have the correlation $j_1 = j_2$ for $\mathbf{t}$ where we get the only possible correlation $\hat{j}_1 = \hat{j}_2 = \hat{j}_3 = \hat{j}_4$ for $\mathbf{u}$, i.e. $\beta(\mathbf{t}) = 1$. On the other hand we have $j_1 \neq j_2$ as a correlation for $\mathbf{t}$ allowing us two different correlations $\hat{j}_1 = \hat{j}_3 \neq \hat{j}_2 = \hat{j}_4$ and $\hat{j}_2 = \hat{j}_3 \neq \hat{j}_1 = \hat{j}_4$ for $\mathbf{u}$. Thus we get $\beta(\mathbf{t}) = 2$ in the latter case.

Here, we note that for every tree $\mathbf{t} \in TS(*)$ with variable indices $j_1, \ldots, j_{s(\mathbf{t})/2}$ there exists a tree $\mathbf{u} \in TS(\Delta)$ with $|\mathbf{u}| \sim |\mathbf{t}|$ and variable indices $\hat{j}_1, \ldots, \hat{j}_{s(\mathbf{u})}$ such that for some suitable correlation of type $\hat{j}_k = \hat{j}_l$ or $\hat{j}_k \neq \hat{j}_l$, $1 \leq k < l \leq s(\mathbf{u})$, we have $\mathbf{t} \sim \mathbf{u}$ and thus $\mathbf{u} \in TS(*)$ with $\alpha_*(\mathbf{u}) = \alpha_*(\mathbf{t})$ for $* \in \{I,S\}$. However, we have $\alpha_*(\mathbf{u}) = 0$ for all $\mathbf{u} \in TS(\Delta) \setminus TS(*)$ for $* \in \{I,S\}$. The following theorem for the class of SRK methods (3) yields order conditions for the coefficients and the random variables of the method such that convergence with some order p in the weak sense is assured (see Thm 6.4 in [10] or Thm 4.4.3 in [12]).

Theorem 4.2 *Let $p \in \mathbb{N}$ and $f, a^i, \tilde{a}^i, b^{i,j} \in C_P^{2(p+1)}(\mathbb{R}^d, \mathbb{R})$ for $i = 1, \ldots, d$ and $j = 1, \ldots, m$. Then the approximation $(Y(t))_{t \in \mathcal{I}_h}$ by the SRK method (3) with maximum step size h is of weak order p, if for all $\boldsymbol{t} \in TS(\Delta)$ with*

$\rho(\boldsymbol{t}) \leq p + \frac{1}{2}$ and all correlations of type $j_k = j_l$ or $j_k \neq j_l$, $1 \leq k < l \leq s(\boldsymbol{t})$, between the indices $j_1, \ldots, j_{s(\boldsymbol{t})} \in \{1, \ldots, m\}$ of $\boldsymbol{t}$ the equations

$$\frac{\alpha_*(\boldsymbol{t}) \cdot h^{\rho(\boldsymbol{t})}}{2^{s(\boldsymbol{t})/2} \cdot \rho(\boldsymbol{t})!} = \frac{\alpha_\Delta(\boldsymbol{t}) \cdot \beta(\boldsymbol{t}) \cdot \gamma(\boldsymbol{t}) \cdot E(\Phi_S(\boldsymbol{t}))}{(l(\boldsymbol{t}) - 1)!} \tag{11}$$

hold for $* = I$ in case of the Itô SDE (1) and $* = S$ in case of the Stratonovich SDE (1), provided that (4), (5) and (10) hold.

We note that Theorem 4.2 provides uniform weak convergence with order p in the case of a non-random time discretization $\mathcal{I}_h$ [10].

5 Stochastic Runge–Kutta Methods for Itô SDE Systems

The aim of this section is to deduce conditions for the coefficients of the SRK method (3) from Theorem 4.2, such that convergence with order 2.0 to the solution of the Itô SDE (1) in the weak sense is guaranteed. Therefore, we choose $\mathcal{M} = \{\{j_1\}, \{j_1, j_2\} : 1 \leq j_1, j_2 \leq m\}$ and

$$\begin{aligned} z_i^{(0,0)} &= \alpha_i \, h_n & z_i^{(k,l)} &= \gamma_i^{(k)\,(k,l)} \, \hat{I}_{(k)} + \gamma_i^{(k,l)\,(k,l)} \, \frac{\hat{I}_{(k,l)}}{\sqrt{h_n}} \\ Z_{ij}^{(0,0),(0,0)} &= A_{ij}^{(0,0),(0,0)} \, h_n & Z_{ij}^{(0,0),(r,s)} &= B_{ij}^{(r)\,(0,0),(r,s)} \, \hat{I}_{(r)} \\ Z_{ij}^{(k,l),(0,0)} &= A_{ij}^{(k,l),(0,0)} \, h_n & Z_{ij}^{(k,l),(r,s)} &= B_{ij}^{(0)\,(k,l),(r,s)} \sqrt{h_n} \end{aligned}$$

for $1 \leq k, l, r, s \leq m$. Further, we define $B_{ij}^{(r)\,(0,0),(r,s)} = 0$ in the case of $r \neq s$ and $B_{ij}^{(0)\,(k,l),(r,s)} = 0$ in the case of $l \neq r$ or $l \neq s$ for $i, j = 1, \ldots, s$.

The random variables are defined by $\hat{I}_{(r)} = \Delta \hat{W}_n^r$ and $\hat{I}_{(k,l)} = \frac{1}{2}(\Delta \hat{W}_n^k \, \Delta \hat{W}_n^l + V_{k,l})$. Here, the $\Delta \hat{W}_n^r$ could be independent $N(0, h_n)$ Gaussian or three-point distributed random variables with $P(\Delta \hat{W}_n^r = \pm\sqrt{3\, h_n}) = \frac{1}{6}$ and $P(\Delta \hat{W}_n^r = 0) = \frac{2}{3}$. The $V_{k,l}$ are independent two-point distributed random variables with $P(V_{k,l} = \pm h_n) = \frac{1}{2}$ for $l = 1, \ldots, k-1$, $V_{k,k} = -h_n$ and $V_{l,k} = -V_{k,l}$ for $l = k+1, \ldots, m$ and $k = 1, \ldots, m$ [5].

Then the d-dimensional approximation process Y is given by an stochastic

Runge–Kutta method with s stages which is defined by $Y_0 = x_0$ and

$$\begin{aligned} Y_{n+1} = Y_n &+ \sum_{i=1}^{s} \alpha_i \, a(t_n + c_i^{(0,0)} h_n, H_i^{(0,0)}) \, h_n \\ &+ \sum_{i=1}^{s} \sum_{k,l=1}^{m} \gamma_i^{(k)(k,l)} \, b^k(t_n + c_i^{(k,l)} h_n, H_i^{(k,l)}) \, \hat{I}_{(k)} \\ &+ \sum_{i=1}^{s} \sum_{k,l=1}^{m} \gamma_i^{(k,l)(k,l)} \, b^k(t_n + c_i^{(k,l)} h_n, H_i^{(k,l)}) \, \frac{\hat{I}_{(k,l)}}{\sqrt{h_n}} \end{aligned} \tag{12}$$

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

$$\begin{aligned} H_i^{(0,0)} &= Y_n + \sum_{j=1}^{s} A_{ij}^{(0,0)(0,0)} \, a(t_n + c_i^{(0,0)} h_n, H_j^{(0,0)}) \, h_n \\ &\quad + \sum_{j=1}^{s} \sum_{r=1}^{m} B_{ij}^{(r)(0,0)(r,r)} \, b^r(t_n + c_i^{(r,r)} h_n, H_j^{(r,r)}) \, \hat{I}_{(r)} \\ H_i^{(k,l)} &= Y_n + \sum_{j=1}^{s} A_{ij}^{(k,l)(0,0)} \, a(t_n + c_i^{(0,0)} h_n, H_j^{(0,0)}) \, h_n \\ &\quad + \sum_{j=1}^{s} B_{ij}^{(0)(k,l)(l,l)} \, b^l(t_n + c_i^{(l,l)} h_n, H_j^{(l,l)}) \sqrt{h_n} \end{aligned}$$

for $i = 1, \ldots, s$ and $k, l = 1, \ldots, m$.

Clearly, the SRK method (12) may be implicit or explicit, which depends on the choice of the coefficients for the method. It turns out that the number of coefficients for the SRK method (12) can be reduced by just differing from the two cases $k = l$ in contrast to $k \neq l$. Thus we can restrict our considerations to the three different support values $H_i^{(0,0)}$, $H_i^{(k,k)}$ and $H_i^{(k,l)}$ for $k \neq l$.

As we are looking for coefficients, such that the proposed SRK method (12) converges with order 2.0 in the weak sense, we apply Theorem 4.2. Since $E({z^{(k,l)}}^T e) = 0$ holds for $1 \leq k, l \leq m$, Y is uniformly bounded and the following theorem gives general conditions for the coefficients of the stochastic Runge–Kutta method (12) (see also Thm 5.1.1 in [12]).

Theorem 5.1 *If the coefficients of the stochastic Runge–Kutta method (12) fulfill for $k, l = 1, \ldots, m$ the equations*

1. $\alpha^T e = 1$
2. $({\gamma^{(k)(k,k)}}^T e)^2 = 1$
3. ${\gamma^{(k,l)(k,l)}}^T e = 0$
4. ${\gamma^{(k)(k,l)}}^T B^{(0)(k,l)(l,l)} e = 0$
5. ${\gamma^{(k)(k,l)}}^T e = 0$ *for* $k \neq l$

then the method converges with order 1.0 in the weak sense to the solution of *the Itô SDE (1). If in addition for $k, l = 1, \ldots, m$ the equations*

6. $\alpha^T A^{(0,0)(0,0)} e = \frac{1}{2}$
7. ${\gamma^{(k,l)}}^{(k,l)^T} A^{(k,l)(0,0)} e = 0$
8. ${\gamma^{(k,l)}}^{(k,l)^T} ({B^{(0)}}^{(k,l)(l,l)} e)^2 = 0$
9. $\alpha^T ({B^{(k)}}^{(0,0)(k,k)} e)^2 = \frac{1}{2}$
10. ${\gamma^{(k,l)}}^{(k,l)^T} {B^{(0)}}^{(k,l)(l,l)} e = 1$
11. ${\gamma^{(k)}}^{(k,l)^T} ({B^{(0)}}^{(k,l)(l,l)} e)^3 = 0$

12. $({\gamma^{(k)}}^{(k,k)^T} e)(\alpha^T {B^{(k)}}^{(0,0)(k,k)} e) = \frac{1}{2}$
13. ${\gamma^{(k)}}^{(k,l)^T} ({B^{(0)}}^{(k,l)(l,l)} (A^{(l,l)(0,0)} e)) = 0$
14. ${\gamma^{(k)}}^{(k,l)^T} ({B^{(0)}}^{(k,l)(l,l)} ({B^{(0)}}^{(l,l)(l,l)} ({B^{(0)}}^{(l,l)(l,l)} e))) = 0$
15. ${\gamma^{(k)}}^{(k,l)^T} (A^{(k,l)(0,0)} ({B^{(k)}}^{(0,0)(k,k)} e)) = 0$
16. $\alpha^T ({B^{(k)}}^{(0,0)(k,k)} ({B^{(0)}}^{(k,k)(k,k)} e)) = 0$
17. $({\gamma^{(k)}}^{(k,k)^T} e)({\gamma^{(k)}}^{(k,k)^T} A^{(k,k)(0,0)} e) = \frac{1}{2}$
18. ${\gamma^{(k)}}^{(k,l)^T} A^{(k,l)(0,0)} e = 0$ *for* $k \neq l$
19. ${\gamma^{(k)}}^{(k,l)^T} (({B^{(0)}}^{(k,l)(l,l)} e)(A^{(k,l)(0,0)} e)) = 0$
20. ${\gamma^{(k)}}^{(k,l)^T} ({B^{(0)}}^{(k,l)(l,l)} ({B^{(0)}}^{(l,l)(l,l)} e)) = 0$
21. $({\gamma^{(k)}}^{(k,k)^T} e)({\gamma^{(k)}}^{(k,l)^T} ({B^{(0)}}^{(k,l)(l,l)} e)^2) = \frac{1}{2}$
22. ${\gamma^{(k,l)}}^{(k,l)^T} ({B^{(0)}}^{(k,l)(l,l)} ({B^{(0)}}^{(l,l)(l,l)} e)) = 0$
23. ${\gamma^{(k)}}^{(k,l)^T} ({B^{(0)}}^{(k,l)(l,l)} (A^{(l,l)(0,0)} ({B^{(k)}}^{(0,0)(k,k)} e))) = 0$
24. $\alpha^T (({B^{(k)}}^{(0,0)(k,k)} e)({B^{(k)}}^{(0,0)(k,k)} ({B^{(0)}}^{(k,k)(k,k)} e))) = 0$
25. ${\gamma^{(k,l)}}^{(k,l)^T} (A^{(k,l)(0,0)} (({B^{(k)}}^{(0,0)(k,k)} e)({B^{(l)}}^{(0,0)(l,l)} e))) = 0$
26. ${\gamma^{(k)}}^{(k,l)^T} (({B^{(0)}}^{(k,l)(l,l)} e)(A^{(k,l)(0,0)} ({B^{(k)}}^{(0,0)(k,k)} e))) = 0$
27. ${\gamma^{(k)}}^{(k,l)^T} (A^{(k,l)(0,0)} ({B^{(k)}}^{(0,0)(k,k)} ({B^{(0)}}^{(k,k)(k,k)} e))) = 0$
28. ${\gamma^{(k)}}^{(k,l)^T} ({B^{(0)}}^{(k,l)(l,l)} (({B^{(0)}}^{(l,l)(l,l)} e)^2)) = 0$
29. ${\gamma^{(k)}}^{(k,l)^T} (({B^{(0)}}^{(k,l)(l,l} e)({B^{(0)}}^{(k,l)(l,l)} ({B^{(0)}}^{(l,l)(l,l)} e))) = 0$
30. ${\gamma^{(k,l)}}^{(k,l)^T} (A^{(k,l)(0,0)} ({B^{(l)}}^{(0,0)(l,l)} e)) = 0$

are fulfilled then the stochastic Runge–Kutta method (12) converges with order 2.0 in the weak sense to the solution of the Itô SDE (1).

Remark 5.2 *The conditions for the coefficients presented in Theorem 5.1 have to be fulfilled for the two cases $k = l$ and $k \neq l$. Thus we have to solve 50 equations for $m > 1$. However in case of $m = 1$ the 50 conditions reduce* *to 28 conditions (see also [12]). For explicit SRK methods $s \geq 3$ is needed.*

Solving the order conditions of Theorem 5.1, it turns out that we can choose $A_{ij}^{(k,l),(0,0)} = 0$ in the case of $k \neq l$ for $i,j = 1, \ldots, s$. Thus, we can characterize the SRK method (3) by the following Butcher array for $1 \leq k,l \leq m$ with $k \neq l$:

$$
\begin{array}{c|c|c|c}
c^{(0,0)} & A^{(0,0),(0,0)} & B^{(k)\,(0,0),(k,k)} & \\
\hline
c^{(k,k)} & A^{(k,k),(0,0)} & B^{(0)\,(k,k),(k,k)} & B^{(0)\,(k,l),(l,l)} \\
\hline
 & \alpha^T & \gamma^{(k)\,(k,k)^T} & \gamma^{(k,k)\,(k,k)^T} \\
\cline{2-4}
 & & \gamma^{(k)\,(k,l)^T} & \gamma^{(k,l)\,(k,l)^T}
\end{array}
$$

Taking into account the order conditions 1.–5. of Theorem 5.1, we can easily calculate SRK methods converging with order 1.0 in the weak sense. As an example the well known Euler-Maruyama scheme (see, e.g., [5]) belongs to the introduced class of SRK methods having order 1.0 with $s = 1$ stage. Considering in addition the conditions 6.–30. of Theorem 5.1, we calculate some coefficients for an explicit SRK method (12) of weak order two. Due to some degrees of freedom in choosing the coefficients for the deterministic part, it is possible to calculate a SRK scheme converging with order three if it is applied to deterministic ordinary differential equations. Therefore, if the weights α_i and the coefficients $A_{ij}^{(0,0),(0,0)}$ are chosen such that condition 1. and 6. of Theorem 5.1 and additionally the conditions $\alpha^T(A^{(0,0),(0,0)}(A^{(0,0),(0,0)}e)) = \frac{1}{6}$ and $\alpha^T(A^{(0,0),(0,0)}e)^2 = \frac{1}{3}$ are fulfilled (see, e.g. [3]), then the SRK scheme is of order three in the case of $b \equiv 0$.

In the following let (p_D, p_S) with $p_D \geq p_S$ denote the order of convergence of the SRK scheme if it is applied to a deterministic or stochastic differential equation, respectively. Thus, the scheme converges at least with order $p = p_S$ in the weak sense and we suppose better convergence for schemes with $p_D > p_S$, particularly for SDEs with small noise. The schemes RI1WM,

$$
\begin{array}{c|ccc|ccc|ccc}
0 & & & & & & & & & \\
\frac{2}{3} & \frac{2}{3} & & & 1 & & & & & \\
\frac{2}{3} & -\frac{1}{3} & 1 & & 0 & 0 & & & & \\
\hline
0 & & & & & & & & & \\
1 & 1 & & & 1 & & & 1 & & \\
1 & 1 & 0 & & -1 & 0 & & -1 & 0 & \\
\hline
 & \frac{1}{4} & \frac{1}{2} & \frac{1}{4} & \frac{1}{2} & \frac{1}{4} & \frac{1}{4} & 0 & \frac{1}{2} & -\frac{1}{2} \\
\cline{2-10}
 & & & & -\frac{1}{2} & \frac{1}{4} & \frac{1}{4} & 0 & \frac{1}{2} & -\frac{1}{2}
\end{array}
\qquad
\begin{array}{c|ccc|ccc|ccc}
0 & & & & & & & & & \\
\frac{2}{3} & \frac{2}{3} & & & \frac{1}{3} & & & & & \\
\frac{2}{3} & -\frac{1}{3} & 1 & & \frac{4}{3} & 0 & & & & \\
\hline
0 & & & & & & & & & \\
1 & 1 & & & 1 & & & 1 & & \\
1 & 1 & 0 & & -1 & 0 & & -1 & 0 & \\
\hline
 & \frac{1}{4} & \frac{1}{2} & \frac{1}{4} & \frac{1}{2} & \frac{1}{4} & \frac{1}{4} & 0 & \frac{1}{2} & -\frac{1}{2} \\
\cline{2-10}
 & & & & -\frac{1}{2} & \frac{1}{4} & \frac{1}{4} & 0 & \frac{1}{2} & -\frac{1}{2}
\end{array}
$$

Table 1
SRK scheme RI1WM and RI2WM of order $p_D = 3.0$ and $p_S = 2.0$.

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

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

Table 2
SRK scheme RI4WM and RI5WM of order $p_D = 3.0$ and $p_S = 2.0$.

RI2WM, RI4WM and RI5WM presented in Table 1 and Table 2, respectively, are of order $p_S = 2.0$ and $p_D = 3.0$ in the weak sense. The explicit weak order 2.0 scheme proposed by Platen [5] of order $p_D = p_S = 2.0$ is also of the considered class of SRK methods. For coefficients of this scheme and further coefficients we refer to [12].

6 Stochastic Runge–Kutta Methods for Stratonovich SDE Systems with Commutative Noise

We consider SRK schemes for the approximation of Stratonovich SDE systems with commutative noise. The main advantage of such schemes is that they do not need the simulation of correlated random variables which saves computational effort. Therefore, we assume that the diffusion b of the Stratonovich SDE (1) satisfies the commutativity condition

$$\sum_{i=1}^{d} b^{i,j_1} \frac{\partial b^{k,j_2}}{\partial x^i} = \sum_{i=1}^{d} b^{i,j_2} \frac{\partial b^{k,j_1}}{\partial x^i} \qquad (13)$$

for every $j_1, j_2 = 1, \ldots, m$, $k = 1, \ldots, d$. SDEs satisfying (13) are called SDEs with commutative noise. For a SRK method for SDEs with commutative noise (see also [11,12]) we choose $\mathcal{M} = \{j : 1 \leq j \leq m\}$ and

$$z_i^{(0,0)} = \alpha_i \, h_n \qquad\qquad z_i^{(l,l)} = \gamma_i^{(l)\,(l,l)} \, \hat{I}_{(l)}$$
$$Z_{ij}^{(k,k),(0,0)} = A_{ij}^{(k,k),(0,0)} \, h_n \qquad\qquad Z_{ij}^{(k,k),(l,l)} = B_{ij}^{(l)\,(k,k),(l,l)} \, \hat{I}_{(l)}$$

for $k \in \mathcal{M} \cup \{0\}$, $l \in \mathcal{M}$ and $i, j = 1, \ldots, s$. The coefficients of such a method can be represented by the Butcher array taking for $k, l \in \mathcal{M}$ with $k \neq l$ the form

$$
\begin{array}{c|c|c|c}
c^{(0,0)} & A^{(0,0),(0,0)} & B^{(k)\,(0,0),(k,k)} & \\
\hline
c^{(k,k)} & A^{(k,k),(0,0)} & B^{(k)\,(k,k),(k,k)} & B^{(l)\,(k,k),(l,l)} \\
\hline
 & \alpha^T & {\gamma^{(k)\,(k,k)}}^T &
\end{array}
$$

Then the d-dimensional approximation process Y is given by a SRK method with s stages defined by $Y_0 = x_0$ and

$$
\begin{aligned}
Y_{n+1} = Y_n &+ \sum_{i=1}^{s} \alpha_i \, a(t_n + c_i^{(0,0)} h_n, H_i^{(0,0)}) \, h_n \\
&+ \sum_{i=1}^{s} \sum_{k=1}^{m} {\gamma_i^{(k)}}^{(k,k)} \, b^k(t_n + c_i^{(k,k)} h_n, H_i^{(k,k)}) \, \hat{I}_{(k)}
\end{aligned}
\tag{14}
$$

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

$$
\begin{aligned}
H_i^{(0,0)} = Y_n &+ \sum_{j=1}^{s} A_{ij}^{(0,0)(0,0)} \, a(t_n + c_i^{(0,0)} h_n, H_j^{(0,0)}) \, h_n \\
&+ \sum_{j=1}^{s} \sum_{r=1}^{m} {B_{ij}^{(r)}}^{(0,0)(r,r)} \, b^r(t_n + c_i^{(r,r)} h_n, H_j^{(r,r)}) \, \hat{I}_{(r)} \\
H_i^{(k,k)} = Y_n &+ \sum_{j=1}^{s} A_{ij}^{(k,k)(0,0)} \, a(t_n + c_i^{(0,0)} h_n, H_j^{(0,0)}) \, h_n \\
&+ \sum_{j=1}^{s} \sum_{r=1}^{m} {B_{ij}^{(r)}}^{(k,k)(r,r)} \, b^l(t_n + c_i^{(r,r)} h_n, H_j^{(r,r)}) \, \hat{I}_{(r)}
\end{aligned}
$$

for $i = 1, \ldots, s$ and $k = 1, \ldots, m$.

The commutativity condition can be illustrated in the light of rooted trees

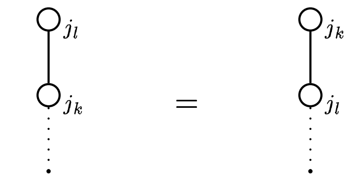


Fig. 2. Two equivalent endings of a tree in case of commutative noise.

as follows: If the commutativity condition (13) holds, then the endings $\{\sigma_{j_k}\}_{j_l}$ and $\{\sigma_{j_l}\}_{j_k}$ of a rooted tree, presented in Figure 2, are equal and we can substitute one of the two endings of a rooted tree by the other one. This is a direct consequence from the corresponding elementary differentials. By the use of this item it is possible to calculate conditions for the coefficients of the SRK method (see [11,12]).

Theorem 6.1 *If the coefficients of the stochastic Runge–Kutta method (14) fulfill for $k = 1, \ldots, m$ the equations*

$$1.\quad \alpha^T e = 1 \qquad 2.\quad ({\gamma^{(k)(k,k)}}^T e)^2 = 1 \qquad 3.\quad {\gamma^{(k)(k,k)}}^T B^{(k)(k,k)(k,k)} e = \tfrac{1}{2}$$

then the method converges with order 1.0 in the weak sense to the solution of the Stratonovich SDE (1). If in addition for $k, l = 1, \ldots, m$ with $k \neq l$ the equations

$$4.\quad \alpha^T A^{(0,0)(0,0)} e = \tfrac{1}{2} \qquad 7.\quad ({\gamma^{(k)(k,k)}}^T e)(\alpha^T B^{(k)(0,0)(k,k)} e) = \tfrac{1}{2}$$

$$5.\quad \alpha^T (B^{(k)(0,0)(k,k)} e)^2 = \tfrac{1}{2} \qquad 8.\quad \alpha^T (B^{(k)(0,0)(k,k)} (B^{(k)(k,k)(k,k)} e)) = \tfrac{1}{4}$$

$$6.\quad {\gamma^{(k)(k,k)}}^T B^{(l)(k,k)(l,l)} e = \tfrac{1}{2} \qquad 9.\quad {\gamma^{(k)(k,k)}}^T (B^{(k)(k,k)(k,k)} e)^3 = \tfrac{1}{4}$$

$$10.\quad ({\gamma^{(k)(k,k)}}^T e)({\gamma^{(k)(k,k)}}^T A^{(k,k)(0,0)} e) = \tfrac{1}{2}$$

$$11.\quad {\gamma^{(k)(k,k)}}^T (B^{(k)(k,k)(k,k)} (A^{(k,k)(0,0)} e)) = \tfrac{1}{4}$$

$$12.\quad {\gamma^{(k)(k,k)}}^T ((A^{(k,k)(0,0)} e)(B^{(k)(k,k)(k,k)} e)) = \tfrac{1}{4}$$

$$13.\quad ({\gamma^{(k)(k,k)}}^T e)({\gamma^{(k)(k,k)}}^T (B^{(k)(k,k)(k,k)} e)^2) = \tfrac{1}{3}$$

$$14.\quad ({\gamma^{(k)(k,k)}}^T e)({\gamma^{(k)(k,k)}}^T (B^{(l)(k,k)(l,l)} e)^2) = \tfrac{1}{2}$$

$$15.\quad ({\gamma^{(l)(l,l)}}^T e)({\gamma^{(k)(k,k)}}^T ((B^{(l)(k,k)(l,l)} e)(B^{(k)(k,k)(k,k)} e))) = \tfrac{1}{4}$$

$$16.\quad ({\gamma^{(k)(k,k)}}^T e)({\gamma^{(k)(k,k)}}^T (B^{(k)(k,k)(k,k)} (B^{(k)(k,k)(k,k)} e))) = \tfrac{1}{6}$$

$$17.\quad ({\gamma^{(k)(k,k)}}^T e)({\gamma^{(k)(k,k)}}^T (B^{(l)(k,k)(l,l)} (B^{(l)(l,l)(l,l)} e))) = \tfrac{1}{4}$$

$$18.\quad ({\gamma^{(k)(k,k)}}^T e)({\gamma^{(l)(l,l)}}^T (B^{(k)(l,l)(k,k)} (B^{(l)(k,k)(l,l)} e))) = 0$$

$$19.\quad ({\gamma^{(k)(k,k)}}^T e)({\gamma^{(l)(l,l)}}^T (B^{(l)(l,l)(l,l)} (B^{(k)(l,l)(k,k)} e))) = \tfrac{1}{4}$$

$$20.\quad {\gamma^{(k)(k,k)}}^T ((B^{(k)(k,k)(k,k)} e)(B^{(l)(k,k)(l,l)} e)^2) = \tfrac{1}{4}$$

$$21.\quad {\gamma^{(k)(k,k)}}^T ((B^{(k)(k,k)(k,k)} e)(B^{(k)(k,k)(k,k)} (B^{(k)(k,k)(k,k)} e))) = \tfrac{1}{8}$$

$$22.\quad {\gamma^{(k)(k,k)}}^T ((B^{(l)(k,k)(l,l)} e)(B^{(k)(k,k)(k,k)} (B^{(l)(k,k)(l,l)} e))) = \tfrac{1}{4}$$

$$23.\quad {\gamma^{(k)(k,k)}}^T ((B^{(l)(k,k)(l,l)} e)(B^{(l)(k,k)(l,l)} (B^{(k)(l,l)(k,k)} e))) = 0$$

$$24.\quad {\gamma^{(k)(k,k)}}^T ((B^{(k)(k,k)(k,k)} e)(B^{(l)(k,k)(l,l)} (B^{(l)(l,l)(l,l)} e))) = \tfrac{1}{8}$$

$$25.\quad {\gamma^{(k)(k,k)}}^T (B^{(k)(k,k)(k,k)} (B^{(k)(k,k)(k,k)} e)^2) = \tfrac{1}{12}$$

$$26.\quad {\gamma^{(k)(k,k)}}^T (B^{(k)(k,k)(k,k)} (B^{(l)(k,k)(l,l)} e)^2) = \tfrac{1}{4}$$

$$27.\quad {\gamma^{(k)(k,k)}}^T (B^{(l)(k,k)(l,l)} ((B^{(k)(l,l)(k,k)} e)(B^{(l)(l,l)(l,l)} e))) = 0$$

$$28.\quad {\gamma^{(k)(k,k)}}^T (B^{(k)(k,k)(k,k)} (B^{(k)(k,k)(k,k)} (B^{(k)(k,k)(k,k)} e))) = \tfrac{1}{24}$$

29. $\gamma^{(k)(k,k)^T}(B^{(k)(k,k)(k,k)}(B^{(l)(k,k)(l,l)}(B^{(l)(l,l)(l,l)}e))) = \frac{1}{8}$

30. $\gamma^{(k)(k,k)^T}(B^{(l)(k,k)(l,l)}(B^{(k)(l,l)(k,k)}(B^{(l)(k,k)(l,l)}e))) = 0$

31. $\gamma^{(k)(k,k)^T}(B^{(l)(k,k)(l,l)}(B^{(l)(l,l)(l,l)}(B^{(k)(l,l)(k,k)}e))) = 0$

32. $\gamma^{(k)(k,k)^T}(A^{(k,k)(0,0)}(B^{(k)(0,0)(k,k)}e)) = 0$

are fulfilled then the stochastic Runge–Kutta method (14) converges with order 2.0 in the weak sense to the solution of the Stratonovich SDE (1).

Remark 6.2 *The 32 conditions of Theorem 6.1 reduce to 17 conditions in the case of* $m = 1$*. Further we need* $s \geq 4$ *for an explicit SRK method (see [11,12]).*

Some explicit SRK schemes of order 2.0 are given by RS1WM and RS2WM with the coefficients presented in Table 3. Due to $s = 4$ stages needed for

0												
0	0				0							
1	1	0			$-\frac{3}{4}$	$\frac{3}{4}$						
0	0	0	0		1	0	0					
0												
0	0				$\frac{2}{3}$				0			
1	1	0			$\frac{1}{12}$	$\frac{1}{4}$			$\frac{1}{4}$	$\frac{3}{4}$		
1	1	0	0		$-\frac{5}{4}$	$\frac{1}{4}$	2		$\frac{1}{4}$	$\frac{3}{4}$	0	
	0	0	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$				

0												
$\frac{2}{3}$	$\frac{2}{3}$				0							
$\frac{2}{3}$	$\frac{1}{6}$	$\frac{1}{2}$			$\frac{1}{4}$	$\frac{3}{4}$						
0	0	0	0		0	0	0					
0												
0	0				$\frac{2}{3}$				0			
1	1	0			$\frac{1}{12}$	$\frac{1}{4}$			$\frac{1}{4}$	$\frac{3}{4}$		
1	1	0	0		$-\frac{5}{4}$	$\frac{1}{4}$	2		$\frac{1}{4}$	$\frac{3}{4}$	0	
	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{2}$	0	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$				

Table 3
SRK scheme RS1WM with $p_D = p_S = 2.0$ and RS2WM with $p_D = 3.0$ and $p_S = 2.0$.

the Stratonovich SRK methods, it is possible to calculate schemes of a higher order of convergence for the deterministic part, represented by the coefficients $A^{(0,0)(0,0)}$ and α. For example, RS2WM is of order $(3.0, 2.0)$ while RS1WM is of order $(2.0, 2.0)$. Clearly, for $s = 4$ it is also possible to calculate coefficients for a scheme of order $(4.0, 2.0)$ (see [12]).

7 Numerical Results

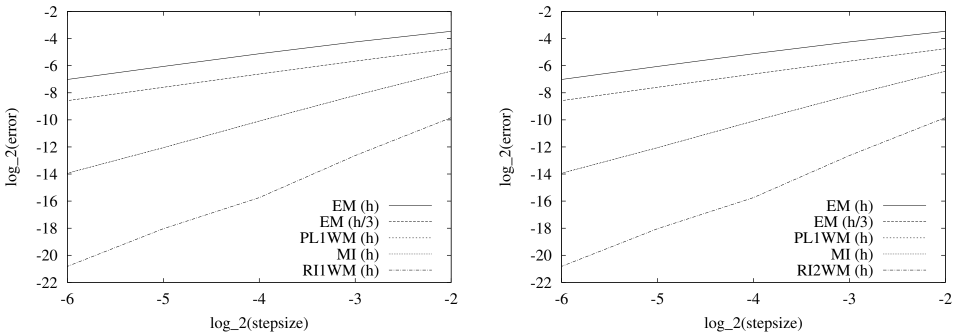
The SRK schemes for Itô SDEs are compared with the order 1.0 Euler-Maruyama scheme (EM), the order 2.0 scheme (PL1WM) due to Platen [5] and a scheme proposed by Milstein (MI) [8] of order 2.0 which is not derivative free.

In order to have nearly the same computational effort, the Euler-Maruyama scheme is applied with step size $h/3$. The test equation is the linear SDE

$$\mathrm{d}X_t = aX_t\,\mathrm{d}t + bX_t\,\mathrm{d}W_t, \qquad X_0 = x_0, \tag{15}$$

with $a = 1.5$, $b = 0.1$, $f(x) = x$, $x_0 = 0.1$ and $T = 1$. The expectation of the solution at T is given by $E(X_T) = x_0 \cdot \exp(a\,T)$. We approximate $E(f(X_t))$ by a Monte Carlo simulation using the sample average $\frac{1}{M}\sum_{k=1}^{M} f(Y_T(\omega_k))$ of M independent simulated realizations of Y_T. Then, the mean error is denoted by $\hat{\mu}$ and the empirical variance of the mean error is denoted by $\hat{\sigma}_{\mu}^2$. Now, $M = 80\,000\,000$ trajectories are simulated with step sizes $2^{-2}, \dots, 2^{-6}$ and the error $|\hat{\mu}|$ is considered at time $T = 1$. The results are presented in Table 4 and plotted in Figure 3 with double logarithmic scale. Then, the empirical order of convergence is the slope of the printed lines. The first two lines with slope

Fig. 3. Order of confergence of scheme RI1WM and RI2WM.



of ≈ 1.0 correspond to the scheme EM with step size h and $h/3$, respectively. The third and fourth lines with a slope of ≈ 2.0 coincide and correspond to PL1WM and MI. The fifth line represents the results of the SRK scheme. The empirical order of convergence of the schemes RI1WM, RI2WM, RI4WM and RI5WM is significantly higher than 2.0 which is due to the determinsitic order $p_D = 3.0$ of the SRK schemes. Clearly, the better performance depends on the order of magnitude of the diffusion. Thus, the improved performance decreases as the coefficient b increases. However, also for higher values of b, the SRK schemes still perform better than the other schemes under consideration, as additional simulations revealed (see [12]). Similar results for the schemes RS1WM and RS2WM can be found in [11].

A Proofs

Proof of Theorem 5.1. Apply Theorem 4.2 for all trees $\mathbf{t} \in TS(\Delta)$ of order $\rho(\mathbf{t}) \leq 2.5$. All necessary trees are also specified in [10] and [12]. Leaving out the trees whose conditions are fulfilled for any choice of the coefficients,

Table 4
Numerical results for the orders of convergence with SDE (15).

	h	$\vert\hat{\mu}\vert$	$\hat{\sigma}_{\mu}^2$		h	$\vert\hat{\mu}\vert$	$\hat{\sigma}_{\mu}^2$
RI1WM	2^{-2}	1.093351e-03	1.633037e-09	RI2WM	2^{-2}	1.093384e-03	1.578671e-09
	2^{-3}	1.570481e-04	1.602722e-09		2^{-3}	1.570562e-04	1.586961e-09
	2^{-4}	1.827831e-05	1.368048e-09		2^{-4}	1.828247e-05	1.364314e-09
	2^{-5}	3.720047e-06	1.677291e-09		2^{-5}	3.719851e-06	1.676111e-09
	2^{-6}	5.398320e-07	2.090135e-09		2^{-6}	5.397594e-07	2.089757e-09
RI4WM	2^{-2}	1.093387e-03	1.603532e-09	RI5WM	2^{-2}	1.097347e-03	1.524668e-09
	2^{-3}	1.570634e-04	1.594753e-09		2^{-3}	1.581176e-04	1.620628e-09
	2^{-4}	1.827862e-05	1.365900e-09		2^{-4}	2.235743e-05	1.683908e-09
	2^{-5}	3.719930e-06	1.676535e-09		2^{-5}	5.568290e-06	2.133419e-09
	2^{-6}	5.394784e-07	2.089904e-09		2^{-6}	1.857481e-06	2.756024e-09
EM (h)	2^{-2}	9.072138e-02	5.872681e-10	EM (h/3)	2^{-2}	3.717815e-02	1.154897e-09
	2^{-3}	5.274061e-02	8.745436e-10		2^{-3}	1.972579e-02	1.282961e-09
	2^{-4}	2.870995e-02	1.275816e-09		2^{-4}	1.017250e-02	1.566993e-09
	2^{-5}	1.502008e-02	1.789500e-09		2^{-5}	5.165140e-03	1.985171e-09
	2^{-6}	7.693292e-03	2.628478e-09		2^{-6}	2.607984e-03	2.774675e-09
MI	2^{-2}	1.180486e-02	1.496554e-09	PL1WM	2^{-2}	1.180486e-02	1.496554e-09
	2^{-3}	3.414500e-03	1.517499e-09		2^{-3}	3.414500e-03	1.517499e-09
	2^{-4}	9.167149e-04	1.723751e-09		2^{-4}	9.167149e-04	1.723751e-09
	2^{-5}	2.348662e-04	2.086635e-09		2^{-5}	2.348662e-04	2.086635e-09
	2^{-6}	6.369378e-05	2.850095e-09		2^{-6}	6.369378e-05	2.850095e-09

i.e. trees which don't supply any new restrictions, we calculate the following conditions (see also [12] for details):
For order 1 trees: $\mathbf{t}_{1.1} = (\tau)$: $E({z^{(0,0)}}^T e) = h$ yields the condition $\alpha^T e = 1$.

$\mathbf{t}_{1.2} = (\sigma_{j_1}, \sigma_{j_1})$: $E((\sum_{l=1}^m {z^{(j_1,l)}}^T e)^2) = h$ results in

$$(\sum_{l=1}^m {\gamma^{(j_1)(j_1,l)}}^T e)^2 + \tfrac{1}{2} \sum_{l=1}^m ({\gamma^{(j_1,l)(j_1,l)}}^T e)^2 = 1.$$

For order 1.5 trees: $\mathbf{t}_{1.5.4} = (\sigma_{j_1}, \sigma_{j_2}, \sigma_{j_3})$ corresponds to

$$E\Big((\sum_{l=1}^m {z^{(j_1,l)}}^T e)(\sum_{l=1}^m {z^{(j_2,l)}}^T e)(\sum_{l=1}^m {z^{(j_3,l)}}^T e)\Big) = 0.$$

Since $E(I_{(j_1)} I_{(j_2)} I_{(j_3)}) = 0$ and due to the different values of $E(I_{(j_1)} I_{(j_2)} I_{(j_3,l)})$ and $E(I_{(j_1,l_1)} I_{(j_2,l_2)} I_{(j_3,l_3)})$ for any choice of the indices (see [5]), and taking into account the case $m = 1$, we determine the condition ${\gamma^{(k,l)(k,l)}}^T e = 0$ for $k, l = 1, \ldots, m$.

$\mathbf{t}_{1.5.5} = (\{\sigma_{j_2}\}_{j_1}, \sigma_{j_3})$: $E\Big((2\sum_{l=1}^m {z^{(j_1,l)}}^T (\sum_{s=1}^m Z^{(j_1,l)(j_2,s)} e))(\sum_{l=1}^m {z^{(j_3,l)}}^T e)\Big) = 0$.
Since we have ${B^{(0)}}^{(j_1,l)(j_2,s)} = 0$ for $s \neq j_2$ or for $l \neq j_2$ and keeping in mind that $E(I_{(j)} I_{(k,l)}) = 0$ and that $E(I_{(j_1)} I_{(j_3)})$ takes either the value h if $j_1 = j_3$

or the value 0 if $j_1 \neq j_3$ (see [5]), we get with the conditions for tree $\mathbf{t}_{1.5.4}$:

$$(\sum_{l=1}^{m} \gamma^{(j_1)\,(j_1,l)^T} e)(\gamma^{(j_1)\,(j_1,j_2)^T} B^{(0)\,(j_1,j_2)(j_2,j_2)} e) = 0.$$

For order 2 trees: $\mathbf{t}_{2.1} = ([\tau])$: $E(2z^{(0,0)^T} Z^{(0,0)(0,0)} e) = h^2 \Rightarrow \alpha^T A^{(0,0)(0,0)} e = \frac{1}{2}$.

$\mathbf{t}_{2.2} = (\tau, \tau)$: $E\Big((z^{(0,0)^T} e)^2\Big) = h^2$ implies the condition $(\alpha^T e)^2 = 1$.

$\mathbf{t}_{2.4} = ([\sigma_{j_1}, \sigma_{j_2}])$: Case A): $j_1 = j_2$ with $\mathbf{t}_{2.4} \in TS(I)$ yields

$$E\Big(3\, z^{(0,0)^T} (\sum_{s=1}^{m} Z^{(0,0)(j_1,s)} e)^2\Big) = \tfrac{3}{2} h^2 \quad \Rightarrow \quad \alpha^T ((B^{(j_1)\,(0,0)(j_1,j_1)} e)^2) = \tfrac{1}{2}.$$

Case B): $j_1 \neq j_2$ with $\mathbf{t}_{2.4} \notin TS(I)$ yields $E(\Phi_S(\mathbf{t}_{2.4})) = 0$.

$\mathbf{t}_{2.5} = (\sigma_{j_1}, [\sigma_{j_2}])$: Case A): $j_1 = j_2$ with $\mathbf{t}_{2.5} \in TS(I)$ yields

$$E\Big((\sum_{l=1}^{m} z^{(j_1,l)^T} e)(2\, z^{(0,0)^T} (\sum_{s=1}^{m} Z^{(0,0)(j_1,s)} e))\Big) = h^2.$$

As a result of this, we obtain $(\sum_{l=1}^{m} \gamma^{(j_1)\,(j_1,l)^T} e)(\alpha^T B^{(j_1)\,(0,0)(j_1,j_1)} e) = \frac{1}{2}$.
Case B): $j_1 \neq j_2$ with $\mathbf{t}_{2.5} \notin TS(I)$ results in $E(\Phi_S(\mathbf{t}_{2.5})) = 0$.

$\mathbf{t}_{2.7} = (\sigma_{j_1}, \sigma_{j_2}, \tau)$: Case A): $j_1 = j_2$ with $\mathbf{t}_{2.7} \in TS(I)$ yields

$$E\Big((\sum_{l=1}^{m} z^{(j_1,l)^T} e)^2 (z^{(0,0)^T} e)\Big) = h^2.$$

Thus, we calculate $((\sum_{l=1}^{m} \gamma^{(j_1)\,(j_1,l)^T} e)^2 + \frac{1}{2} \sum_{l=1}^{m} (\gamma^{(j_1,l)\,(j_1,l)^T} e)^2)(\alpha^T e) = 1$.
Case B): $j_1 \neq j_2$ with $\mathbf{t}_{2.7} \notin TS(I)$ implies $E(\Phi_S(\mathbf{t}_{2.7})) = 0$.

$\mathbf{t}_{2.8} = (\sigma_{j_1}, \{\tau\}_{j_2})$: Case A): $j_1 = j_2$ with $\mathbf{t}_{2.8} \in TS(I)$ provides

$$E\Big((\sum_{l=1}^{m} z^{(j_1,l)^T} e)(2 \sum_{l=1}^{m} z^{(j_1,l)^T} Z^{(j_1,l)(0,0)} e)\Big) = h^2.$$

Therefore, we determine the condition

$$\begin{aligned} &2(\sum_{l=1}^{m} \gamma^{(j_1)\,(j_1,l)^T} e)(\sum_{l=1}^{m} \gamma^{(j_1)\,(j_1,l)^T} A^{(j_1,l)(0,0)} e) \\ &+ \sum_{l=1}^{m} (\gamma^{(j_1,l)\,(j_1,l)^T} e)(\gamma^{(j_1,l)\,(j_1,l)^T} A^{(j_1,l)(0,0)} e) = 1. \end{aligned}$$

Case B): $j_1 \neq j_2$ with $\mathbf{t}_{2.8} \notin TS(I)$ implies $E(\Phi_S(\mathbf{t}_{2.8})) = 0$.

$\mathbf{t}_{2.11} = (\sigma_{j_1}, \sigma_{j_2}, \sigma_{j_3}, \sigma_{j_4})$: Case A): $j_1 = j_3$ and $j_2 = j_4$ with $\mathbf{t}_{2.11} \in TS(I)$ results in $E\Big((\sum_{l=1}^m z^{(j_1,l)^T} e)^2 (\sum_{l=1}^m z^{(j_2,l)^T} e)^2\Big)$. Using the conditions calculated for $\mathbf{t}_{1.5.4}$, two cases are considered: Case a): $j_1 = j_2$, $\alpha_I(\mathbf{t}_{2.11a}) = 1$, $\beta(\mathbf{t}_{2.11a}) = 1$ and $E(I_{j_1}^4) = 3h^2$ imply $(\sum_{l=1}^m \gamma^{(j_1)(j_1,l)^T} e)^4 = 1$. Case b): $j_1 \neq j_2$, $\alpha_I(\mathbf{t}_{2.11b}) = 1$, $\beta(\mathbf{t}_{2.11b}) = 3$, $E(I_{j_1}^2 I_{j_2}^2) = h^2$ yield $(\sum_{l=1}^m \gamma^{(j_1)(j_1,l)^T} e)^2 (\sum_{l=1}^m \gamma^{(j_2)(j_2,l)^T} e)^2 = 1$. For the special case $m = 1$ we obtain from case a) and b) with $j_1 = j_2 = 1$ that $(\gamma^{(j_1)(j_1,j_1)^T} e)^4 = 1$ and thus $\gamma^{(j_1)(j_1,l)^T} e = 0$ for $j_1 \neq l$.
Case B): All remaining cases with $\mathbf{t}_{2.11} \notin TS(I)$ imply $E(\Phi_S(\mathbf{t}_{2.11})) = 0$.

$\mathbf{t}_{2.12} = (\sigma_{j_1}, \sigma_{j_2}, \{\sigma_{j_4}\}_{j_3})$: Case A): For $\mathbf{t}_{2.12} \in TS(I)$ with the conditions of tree $\mathbf{t}_{1.5.4}$ and $\mathbf{t}_{2.11}$ we get

$$E\Big((\sum_{l=1}^m z^{(j_1,l)^T} e)(\sum_{l=1}^m z^{(j_2,l)^T} e)(2\sum_{l=1}^m z^{(j_3,l)^T} (\sum_{s=1}^m Z^{(j_3,l)(j_4,s)} e)\Big)$$

Case a): $j_2 = j_4 \neq j_1 = j_3$ (or $j_2 = j_3 \neq j_1 = j_4$) with $\alpha_I(\mathbf{t}_{2.12a}) = 4$ and $\beta(\mathbf{t}_{2.12a}) = 2$ yields $(\gamma^{(j_1)(j_1,j_1)^T} e)(\gamma^{(j_2)(j_2,j_2)^T} e)(\gamma^{(j_1,j_2)(j_1,j_2)^T} B^{(0)(j_1,j_2)(j_2,j_2)} e) = 1$. Case b): $j_1 = j_2 = j_3 = j_4$ with $\alpha_I(\mathbf{t}_{2.12b}) = 4$ and $\beta(\mathbf{t}_{2.12b}) = 1$ we specify the condition $(\gamma^{(j_1)(j_1,j_1)^T} e)^2 (\gamma^{(j_1,j_1)(j_1,j_1)^T} B^{(0)(j_1,j_1)(j_1,j_1)} e) = 1$.
Case B): All remaining cases with $\mathbf{t}_{2.12} \notin TS(I)$ imply $E(\Phi_S(\mathbf{t}_{2.12})) = 0$.

$\mathbf{t}_{2.13} = (\sigma_{j_1}, \{\sigma_{j_3}, \sigma_{j_4}\}_{j_2})$: Case A): $j_1 = j_2$ and $j_3 = j_4$ with $\mathbf{t}_{2.13} \in TS(I)$ supplies $E\Big((\sum_{l=1}^m z^{(j_1,l)^T} e)(3\sum_{l=1}^m z^{(j_1,l)^T} (\sum_{s=1}^m Z^{(j_1,l)(j_3,s)} e)^2)\Big) = \frac{3}{2}h^2$. Case a): $j_1 = j_2 = j_3 = j_4$ gives $(\gamma^{(j_1)(j_1,j_1)^T} e)(\gamma^{(j_1)(j_1,j_1)^T} (B^{(0)(j_1,j_1)(j_1,j_1)} e)^2) = \frac{1}{2}$. Case b): $j_1 = j_2 \neq j_3 = j_4$ yields $(\gamma^{(j_1)(j_1,j_1)^T} e)(\gamma^{(j_2)(j_2,j_3)^T} (B^{(0)(j_2,j_3)(j_3,j_3)} e)^2) = \frac{1}{2}$.
Case B): $j_1 \neq j_2$ or $j_3 \neq j_4$ with $\mathbf{t}_{2.13} \notin TS(I)$ imply $E(\Phi_S(\mathbf{t}_{2.13})) = 0$.

$\mathbf{t}_{2.14} = (\sigma_{j_1}, \{\{\sigma_{j_4}\}_{j_3}\}_{j_2})$:

$$E\Big((\sum_{l=1}^m z^{(j_1,l)^T} e)(3\sum_{l=1}^m z^{(j_2,l)^T} (2\sum_{s=1}^m Z^{(j_2,l)(j_3,s)} (\sum_{k=1}^m Z^{(j_3,s)(j_4,k)} e)))\Big) = 0$$

contributes the condition

$$(\gamma^{(j_1)(j_1,j_1)^T} e)(\gamma^{(j_1)(j_1,j_3)^T} (B^{(0)(j_1,j_3)(j_3,j_3)} (B^{(0)(j_3,j_3)(j_3,j_3)} e)) = 0.$$

$\mathbf{t}_{2.15} = (\{\sigma_{j_3}\}_{j_1}, \{\sigma_{j_4}\}_{j_2})$: Case A): $j_1 = j_2$ and $j_3 = j_4$ with $\mathbf{t}_{2.15} \in TS(I)$:

$$E\Big((2\sum_{l=1}^m z^{(j_1,l)^T} (\sum_{s=1}^m Z^{(j_1,l)(j_3,s)} e))(2\sum_{l=1}^m z^{(j_2,l)^T} (\sum_{s=1}^m Z^{(j_2,l)(j_4,s)} e))\Big) = 2h^2$$

Case a): $j_1 = j_2 = j_3 = j_4$ results in the condition

$$2(\gamma^{(j_1)(j_1,j_1)^T} B^{(0)(j_1,j_1)(j_1,j_1)} e)^2 + (\gamma^{(j_1,j_1)(j_1,j_1)^T} B^{(0)(j_1,j_1)(j_1,j_1)} e)^2 = 1.$$

Case b): $j_1 = j_2 \neq j_3 = j_4$ provides the condition

$$2(\gamma^{(j_1)(j_1,j_3)^T} B^{(0)(j_1,j_3)(j_3,j_3)} e)^2 + (\gamma^{(j_1,j_3)(j_1,j_3)^T} B^{(0)(j_1,j_3)(j_3,j_3)} e)^2 = 1.$$

Case B): $\mathbf{t}_{2.15} \notin TS(I)$. Case c): $j_1 \neq j_2$ implies $E(\Phi_S(\mathbf{t}_{2.15})) = 0$. Case d): For $j_3 \neq j_4$ holds $E(I_{(j_1,j_3)} I_{(j_2,j_4)}) = 0$ and $E(I_{(j_1)} I_{(j_2)}) = 0$ if $j_1 \neq j_2$. Thus we have to consider the case $j_1 = j_2$ where we calculate the condition

$$(\gamma^{(j_1)(j_1,j_3)^T} B^{(0)(j_1,j_3)(j_3,j_3)} e)(\gamma^{(j_1)(j_1,j_4)^T} B^{(0)(j_1,j_4)(j_4,j_4)} e) = 0.$$

$\mathbf{t}_{2.20} = (\{[\sigma_{j_2}]\}_{j_1})$: $E\Big(3 \sum_{l=1}^{m} z^{(j_1,l)^T} (2 Z^{(j_1,l)(0,0)} (\sum_{s=1}^{m} Z^{(0,0)(j_2,s)} e))\Big) = 0$. Since $E(I_{(j_1)} I_{(j_2)}) = 0$ for $j_1 \neq j_2$, only the case $j_1 = j_2$ has to be considered, and for $l = 1, \ldots, m$ the condition results in $\gamma^{(j_1)(j_1,l)^T} (A^{(j_1,l)(0,0)} (B^{(j_1)(0,0)(j_1,j_1)} e)) = 0$.

Finally, we have to consider all trees $\mathbf{t} \in TS(\Delta)$ with $\rho(\mathbf{t}) = 2.5$ for which the condition $E(\Phi_S(\mathbf{t})) = 0$ has to be fulfilled. As the calculations are analogous to the ones already performed, repetition is avoided. Now, we just have to summarize the calculated conditions in order to arrive at the conditions in Theorem 5.1. □

Proof of Theorem 6.1. Apply Theorem 4.2 for all trees $\mathbf{t} \in TS(\Delta)$ of order $\rho(\mathbf{t}) \leq 2.5$, which are specified in [10] and [12]. Since $E(\Phi_S(\mathbf{t})) = 0$ holds for all $\mathbf{t}$ with an odd number of stochastic nodes, trees of order 0.5, 1.5 and 2.5 can be ignored. The calculations are analogously to the ones performed in the proof of Theorem 5.1 (see also [11,12] for details). Therefore, we restrict our considerations to the case of tree $\mathbf{t}_{2.15}$ where the commutativity condition has to be applied.

$\mathbf{t}_{2.15} = (\{\sigma_{j_2}\}_{j_1}, \{\sigma_{j_4}\}_{j_3})$: $E\Big((2\, z^{(j_1,j_1)^T} Z^{(j_1,j_1)(j_2,j_2)} e)(2\, z^{(j_3,j_3)^T} Z^{(j_3,j_3)(j_4,j_4)} e)\Big)$.
Case A): $\mathbf{t}_{2.15} \in TS(S)$. Case a): $j_1 = j_2 = j_3 = j_4$ with $\alpha_S(\mathbf{t}_{2.15a}) = 3$ and $\beta(\mathbf{t}_{2.15a}) = 1$ results in $(\gamma^{(j_1)(j_1,j_1)^T} B^{(j_1)(j_1,j_1)(j_1,j_1)} e)^2 = \frac{1}{4}$.
Case b): $j_1 = j_2 \neq j_3 = j_4$ with $\alpha_S(\mathbf{t}_{2.15b}) = 1$ and $\beta(\mathbf{t}_{2.15b}) = 1$ yields

$$(\gamma^{(j_1)(j_1,j_1)^T} B^{(j_1)(j_1,j_1)(j_1,j_1)} e)(\gamma^{(j_3)(j_3,j_3)^T} B^{(j_3)(j_3,j_3)(j_3,j_3)} e) = \tfrac{1}{4}.$$

Case c): $j_1 = j_3 \neq j_2 = j_4$ with $\alpha_S(\mathbf{t}_{2.15c}) = 2$ and $\beta(\mathbf{t}_{2.15c}) = 1$ gives

$$(\gamma^{(j_1)(j_1,j_1)^T} B^{(j_2)(j_1,j_1)(j_2,j_2)} e)(\gamma^{(j_1)(j_1,j_1)^T} B^{(j_2)(j_1,j_1)(j_2,j_2)} e) = \tfrac{1}{2}.$$

Case B): $\mathbf{t}_{2.15} \notin TS(S)$. Case d): $j_1 = j_4 \neq j_2 = j_3$ contributes the condition

$$(\gamma^{(j_1)(j_1,j_1)^T} B^{(j_2)(j_1,j_1)(j_2,j_2)} e)(\gamma^{(j_2)(j_2,j_2)^T} B^{(j_1)(j_2,j_2)(j_1,j_1)} e) = 0.$$

Since there is a contradiction between the conditions of the cases c) and d), there exist no coefficients such that we get a SRK method with the proposed choice of random variables having order 2.0 in general. In order to overcome this problem, we make use of the *commutativity condition*, which gives us $\{\sigma_j\}_k = \{\sigma_k\}_j$. Then the cases c) and d) are recalculated as follows:

Case c'): $j_1 = j_3 \neq j_2 = j_4$ with $\alpha_S(\mathbf{t}_{2.15c'}) = 2$ and $\beta(\mathbf{t}_{2.15c'}) = 2$ due to the commutativity condition yields $(\gamma^{(j_1)(j_1,j_1)^T} B^{(j_2)(j_1,j_1)(j_2,j_2)} e)^2 = \frac{1}{4}$.
Case d'): $j_1 = j_4 \neq j_2 = j_3$ is the same as case c'). □

References

[1] K. Burrage and P. M. Burrage, *High strong order explicit Runge–Kutta methods for stochastic ordinary differential equations*, Appl. Numer. Math., Vol. 22, No. 1-3, 81-101, 1996.

[2] K. Burrage and P. M. Burrage, *Order conditions of stochastic Runge–Kutta methods by B-series*, SIAM J. Numer. Anal., Vol. 38, No. 5, 1626-1646, 2000.

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

[4] I. Karatzas and S. E. Shreve, *Brownian Motion and Stochastic Calculus*, Springer New-York, Berlin, 1999.

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

[6] Y. Komori, T. Mitsui and H. Sugiura, *Rooted tree analysis of the order conditions of ROW-type scheme for stochastic differential equations*, BIT, Vol. 37, No. 1, 43–66, 1997.

[7] D. Küpper, J. Lehn and A. Rössler, *A step size control algorithm for the weak approximation of stochastic differential equations*, Preprint No. 2411, TU Darmstadt, 2005. (submitted)

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

[9] N. J. Newton, *Asymptotically efficient Runge–Kutta methods for a class of Itô and Stratonovich equations*, SIAM J. Appl. Math., Vol. 51, No. 2, 542–567, 1991.

[10] A. RÖSSLER, *Rooted tree analysis for order conditions of stochastic Runge–Kutta methods for the weak approximation of stochastic differential equations*, Preprint No. 2368, TU Darmstadt, 2004. (To appear in Stochastic Anal. Appl.)

[11] A. RÖSSLER, *Runge–Kutta methods for Stratonovich stochastic differential equation systems with commutative noise*, J. Comput. Appl. Math., Vol. 164–165, 613–627, 2004.

[12] A. RÖSSLER, *Runge–Kutta Methods for the Numerical Solution of Stochastic Differential Equations*, Ph.D. thesis, Darmstadt University of Technology. Shaker Verlag, Aachen, 2003.

[13] A. RÖSSLER, *Stochastic Taylor Expansions for the Expectation of Functionals of Diffusion Processes*, Stochastic Anal. Appl., Vol. 22, No. 6, 1553-1576, 2004.

[14] W. RÜMELIN, *Numerical treatment of stochastic differential equations*, SIAM J. Numer. Anal., Vol. 19, No. 3, 604–613, 1982.

[15] D. TALAY, *Second-order discretization schemes of stochastic differential systems for the computation of the invariant law*, Stochastics and Stochastic Rep., Vol. 29, No. 1, 13–36, 1990.

[16] A. TOCINO AND J. VIGO-AGUIAR, *Weak second order conditions for stochastic Runge–Kutta methods*, SIAM J. Sci. Comput., Vol. 24, No. 2, 507–523, 2002.