BIT Numerical Mathematics (2007) 47: 657–680
DOI: 10.1007/s10543-007-0130-3
Published online: 12 May 2007 – © Springer 2007

SECOND ORDER RUNGE–KUTTA METHODS FOR STRATONOVICH STOCHASTIC DIFFERENTIAL EQUATIONS*

ANDREAS RÖSSLER[1]

[1] *Technische Universität Darmstadt, Fachbereich Mathematik, Schlossgartenstr. 7, D-64289 Darmstadt, Germany. email: roessler@mathematik.tu-darmstadt.de*

Abstract.

The weak approximation of the solution of a system of Stratonovich stochastic differential equations with a m–dimensional Wiener process is studied. Therefore, a new class of stochastic Runge–Kutta methods is introduced. As the main novelty, the number of stages does not depend on the dimension m of the driving Wiener process which reduces the computational effort significantly. The colored rooted tree analysis due to the author is applied to determine order conditions for the new stochastic Runge–Kutta methods assuring convergence with order two in the weak sense. Further, some coefficients for second order stochastic Runge–Kutta schemes are calculated explicitly.

AMS subject classification (2000): 65C30, 65L06, 60H35, 60H10.

Key words: stochastic Runge–Kutta method, stochastic differential equation, colored rooted tree analysis, weak approximation, numerical method.

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. [1, 2, 5, 9, 11, 16]. In contrast to this, special schemes have to be developed for the weak approximation of SDEs, see e.g. [5, 6, 9, 17]. Recently, stochastic Runge–Kutta (SRK) methods for the weak approximation have been studied by Kloeden and Platen [5], Komori et. al. [7, 8], Rößler [12, 13, 14] and Tocino and Vigo-Aguiar [19]. However, due to the knowledge of the author, all proposed second order SRK methods suffer from an inefficiency if they are applied to SDEs with a multi-dimensional Wiener process. Then, the number of stages and thus the number of evaluations of the diffusion function depends at least linearly on the dimension m of the driving Wiener process. This drawback becomes significant especially for high-dimensional problems. In [14], SRK methods for SDE systems with commutative noise have been introduced with four stages independent of the dimension m of the driving Wiener process. The aim of the

* Received October 17, 2006. Accepted March 6, 2007. Communicated by Anders Szepessy.

present paper is to carry over this item to the non-commutative noise case. Therefore, we expand the class of SRK methods in [14] to a new class now also applicable to SDEs with non-commutative noise. Essentially, these new SRK methods possess two advantages. On the one hand, the number of stages and thus the number of evaluations of the drift and the diffusion functions for each step is constant, i.e. independent of the dimension $m \geq 1$ of the driving Wiener process. On the other hand, the number of random variables that have to be simulated for each step is only $2m-1$. The paper is organized as follows: Firstly, in Sections 2–4 we briefly review the main results of the rooted tree analysis for weak approximation [12, 15]. In Section 5, the new class of SRK methods is introduced and order conditions are calculated by the rooted tree analysis given in Section 2–4. Further, some coefficients for explicit second order SRK schemes are presented. Finally, some numerical examples are presented in Section 6.

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 the d-dimensional Stratonovich SDE system

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

for d, $m \geq 1$ and $t \in \mathcal{I}$ with an m-dimensional Wiener process $(W)_{t\geq 0}$. Let $X_{t_0} = x_0 \in \mathbb{R}^d$ be $\mathcal{F}_{t_0}$-measurable with $\mathrm{E}(\|X_{t_0}\|^{2l}) < \infty$ for some $l \in \mathbb{N}$. Throughout the paper, suppose that $a, b^j : \mathcal{I} \times \mathbb{R}^d \to \mathbb{R}^d$ fulfill a Lipschitz and a linear growth condition w.r.t. the state variable x for $j = 1, \dots, m$ and fulfill the conditions of the Existence and Uniqueness Theorem [4]. Further, 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]. We say that g belongs to $C_P^{k,l}(\mathcal{I} \times \mathbb{R}^d, \mathbb{R})$ if $g \in C^{k,l}(\mathcal{I} \times \mathbb{R}^d, \mathbb{R})$ and $g(t, \cdot) \in C_P^l(\mathbb{R}^d, \mathbb{R})$ holds uniformly in $t \in \mathcal{I}$. Let a discretization $\mathcal{I}_h = \{t_0, t_1, \dots, t_N\}$ with $t_0 < t_1 < \cdots < t_N = T$ 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. *An approximation process Y converges weakly with order p to X at time T as $h \to 0$ if for each $f \in C_P^{2(p+1)}(\mathbb{R}^d, \mathbb{R})$ exist a constant C_f and a finite $\delta_0 > 0$ such that for each $h \in]0, \delta_0[$ holds:*

$$|\,\mathrm{E}(f(X_T)) - \mathrm{E}(f(Y(T)))| \leq C_f\, h^p. \tag{1.2}$$

2 A general class of stochastic Runge–Kutta methods.

We consider a very general class of stochastic Runge–Kutta methods which has been introduced in [12]: Let $\mathcal{M}$ be a 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 (1.1), a general class of s-stage stochastic Runge–Kutta methods is given by $Y_0 = x_0$ and

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

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

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

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

$$z_i^{(0,0)} = \alpha_i \, h_n \qquad z_i^{(k,\nu)} = \sum_{\iota \in \mathcal{M}} \gamma_i^{(\iota)^{(k,\nu)}} \, \theta_\iota(h_n)$$

$$Z_{ij}^{(k,\nu)(0,0)} = A_{ij}^{(k,\nu)(0,0)} \, h_n \qquad Z_{ij}^{(k,\nu)(r,\mu)} = \sum_{\iota \in \mathcal{M}} B_{ij}^{(\iota)^{(k,\nu)(r,\mu)}} \, \theta_\iota(h_n)$$

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. The weights can be defined by

$$c^{(k,\nu)} = A^{(k,\nu)(0,0)} e \tag{2.2}$$

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 (2.1) 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

$$\mathrm{E}\left(\theta_{\nu_1}^{p_1}(h_n) \cdot \ldots \cdot \theta_{\nu_\kappa}^{p_\kappa}(h_n)\right) = O\big(h_n^{(p_1 + \cdots + p_\kappa)/2}\big) \tag{2.3}$$

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 in the case of $b \equiv 0$, 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.

For the analysis of order conditions, without loss of generality we consider an d-dimensional autonomous Stratonovich SDE system w.r.t. a m-dimensional

Wiener process. Following the approach in [12, 15], we denote by $TS(\Delta)$ the set of colored rooted trees which have a root of type $\gamma = \otimes$ and which may consist of some further deterministic nodes of type $\tau = \bullet$ and stochastic nodes of type $\sigma_{j_k} = \bigcirc_{j_k}$ with a variable index $j_k \in \{1, \ldots, m\}$. The variable index j_k is associated with the j_kth component of the m-dimensional driving Wiener process of the considered SDE. If not stated otherwise, each stochastic node has its own variable index. So, if we have a tree with s stochastic nodes then to each stochastic node corresponds exactly one of the indices $j_1, \ldots, j_s$.

Every tree $\mathbf{t} \in TS(\Delta)$ 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. Here, the order of the subtrees $\mathbf{t}_1, \ldots, \mathbf{t}_k$ does not matter since any order leads to equivalent trees. Therefore, we obtain $\mathbf{t}_{2.4} = ([\bigcirc_{j_1}, \bigcirc_{j_2}]) = ([\sigma_{j_1}, \sigma_{j_2}])$, $\mathbf{t}_{2.6} = (\{\bigcirc_{j_2}\}_{j_1}, \bullet) = (\{\sigma_{j_2}\}_{j_1}, \tau)$ and $\mathbf{t}_{2.15} = (\{\bigcirc_{j_2}\}_{j_1}, \{\bigcirc_{j_4}\}_{j_3}) = (\{\sigma_{j_2}\}_{j_1}, \{\sigma_{j_4}\}_{j_3})$ for the trees in Figure 3.1.

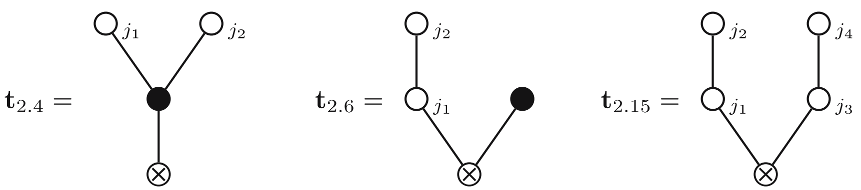

Figure 3.1: Some elements of $TS(\Delta)$ with $j_1, j_2, j_3, j_4 \in \{1, \ldots, m\}$.

In the following, let $l(\mathbf{t})$ be the number of nodes of $\mathbf{t} \in TS(\Delta)$. Then, we denote by $d(\mathbf{t})$ the number of deterministic nodes, by $s(\mathbf{t})$ the number of stochastic nodes of $\mathbf{t} \in TS(\Delta)$ and it holds $l(\mathbf{t}) = d(\mathbf{t}) + s(\mathbf{t}) + 1$. The order $\rho(\mathbf{t})$ of the tree $\mathbf{t} \in TS(\Delta)$ is defined as $\rho(\mathbf{t}) = d(\mathbf{t}) + \frac{1}{2} s(\mathbf{t})$ with $\rho(\gamma) = 0$. For example, it holds $\rho(\mathbf{t}_{2.4}) = \rho(\mathbf{t}_{2.6}) = \rho(\mathbf{t}_{2.15}) = 2$.

Now, let $LTS(\Delta)$ denote the set of monotonically labelled trees, i.e. where the nodes are monotonically numbered starting with number one at the root of the tree. Then, $\alpha_\Delta(\mathbf{t})$ is 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, for $\mathbf{t}_{1.2} = (\sigma_{j_1}, \sigma_{j_2})$ exists only one possible monotonically labelling $(\sigma_{j_1}^2, \sigma_{j_2}^3)^1$ and thus $\alpha_\Delta(\mathbf{t}_{1.2}) = 1$. In contrast to this, for $\mathbf{t}_{1.5.3} = (\tau, \sigma_{j_1})$ holds $\alpha_\Delta(\mathbf{t}_{1.5.3}) = 2$. Although $(\tau^2, \sigma_{j_1}^3)^1$ and $(\sigma_{j_1}^3, \tau^2)^1$ are equivalent trees, there exist two different labelled trees $(\tau^2, \sigma_{j_1}^3)^1$ and $(\tau^3, \sigma_{j_1}^2)^1$. So one has to distinguish between the labels of deterministic and stochastic nodes (see [12] for details). Further, it holds $\alpha_\Delta(\mathbf{t}_{2.4}) = 1$ and $\alpha_\Delta(\mathbf{t}_{2.6}) = \alpha_\Delta(\mathbf{t}_{2.15}) = 3$.

To each tree $\mathbf{t} \in TS(\Delta)$ we assign an elementary differential which is defined recursively by $F(\gamma)(x) = f(x)$, $F(\tau)(x) = a(x)$, $F(\sigma_j)(x) = b^j(x)$ and

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

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

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

where the components of vectors are denoted by superscript indices, which are chosen as capitals. Thus, we obtain for $\mathbf{t}_{2.6}$ the elementary differential

$$F(\mathbf{t}_{2.6}) = f''(b^{j_1\prime}(b^{j_2}), a) = \sum_{J_1,J_2=1}^{d} \frac{\partial^2 f}{\partial x^{J_1} \partial x^{J_2}} \left(\sum_{K_1=1}^{d} \frac{\partial b^{J_1,j_1}}{\partial x^{K_1}} b^{K_1,j_2} \cdot a^{J_2} \right).$$

DEFINITION 3.1. *Let $TS(S)$ denote the set of trees $\mathbf{t} \in TS(\Delta)$ with a root of type γ which can be build by finite many steps of the form*

a) adding a deterministic node of type τ, or
b) adding two stochastic nodes of type σ_{j_k}, both with the same new variable index j_k for some $k \in \mathbb{N}$.

Let $LTS(S)$ denote the set of labelled trees $\mathbf{t} \in TS(S)$ with the nodes labelled in the same order as they are added. Then, $\alpha_S(\mathbf{t})$ is the number of all possible different monotonically labels of $\mathbf{t} \in TS(S)$ with $\alpha_S(\mathbf{t}) = 0$ if $\mathbf{t} \notin TS(S)$.

The following Theorem holds due to Theorem 4.2 and Proposition 5.1 in [15].

THEOREM 3.1. *For $p \in \mathbb{N}_0$, $f, a^i \in C_P^{2p+2}(\mathbb{R}^d, \mathbb{R})$, $b^{i,j} \in C_P^{2p+3}(\mathbb{R}^d, \mathbb{R})$, $i = 1, \dots, d$, $j = 1, \dots, m$, and for $t \in [t_0, T]$ with $h = t - t_0$ the following truncated expansion holds:*

$$\text{(3.2)} \quad \mathrm{E}^{t_0,x_0}(f(X_t)) = \sum_{\substack{\mathbf{t} \in TS(S) \\ \rho(\mathbf{t}) \le p}} \sum_{j_1,\dots,j_{s(\mathbf{t})/2}=1}^{m} \frac{\alpha_S(\mathbf{t})\, F(\mathbf{t})(x_0)}{2^{s(\mathbf{t})/2}\, \rho(\mathbf{t})!} h^{\rho(\mathbf{t})} + O(h^{p+1}).$$

Next, we give an expansion for the approximation process $(Y(t))_{t \in \mathcal{I}_h}$ defined by the SRK method (2.1). For $\mathbf{t} \in TS(\Delta)$ let the density $\gamma(\mathbf{t})$ be defined recursively by $\gamma(\mathbf{t}) = 1$ if $l(\mathbf{t}) = 1$ and

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

Since the expansion for $(Y(t))_{t \in \mathcal{I}_h}$ contains the coefficients and the random variables of the SRK method, we define a coefficient function Φ_S which assigns to every tree $\mathbf{t} \in TS(\Delta)$ an elementary weight:

$$\Phi_S(\mathbf{t}) = \begin{cases} \prod_{i=1}^{\lambda} \Phi_S(\mathbf{t}_i) & \text{if } \mathbf{t} = (\mathbf{t}_1, \dots, \mathbf{t}_\lambda) \\ z^{(0,0)^T} \prod_{i=1}^{\lambda} \Psi^{(0,0)}(\mathbf{t}_i) & \text{if } \mathbf{t} = [\mathbf{t}_1, \dots, \mathbf{t}_\lambda] \\ \sum_{\nu \in \mathcal{M}} z^{(k,\nu)^T} \prod_{i=1}^{\lambda} \Psi^{(k,\nu)}(\mathbf{t}_i) & \text{if } \mathbf{t} = \{\mathbf{t}_1, \dots, \mathbf{t}_\lambda\}_k \end{cases} \tag{3.3}$$

where $\Phi_S(\gamma) = 1$, $\Psi^{(k,\nu)}(\emptyset) = e$ with $\tau = [\emptyset]$, $\sigma_k = \{\emptyset\}_k$ and

$$\Psi^{(k,\nu)}(\mathbf{t}) = \begin{cases} Z^{(k,\nu)(0,0)} \prod_{i=1}^{\lambda} \Psi^{(0,0)}(\mathbf{t}_i) & \text{if } \mathbf{t} = [\mathbf{t}_1, \dots, \mathbf{t}_\lambda] \\ \sum_{\mu \in \mathcal{M}} Z^{(k,\nu)(r,\mu)} \prod_{i=1}^{\lambda} \Psi^{(r,\mu)}(\mathbf{t}_i) & \text{if } \mathbf{t} = \{\mathbf{t}_1, \dots, \mathbf{t}_\lambda\}_r \end{cases} . \tag{3.4}$$

Here $e = (1, \dots, 1)^T$ and the product of vectors in (3.4) is defined by component-wise multiplication, i.e. $(a_1, \dots, a_n) \cdot (b_1, \dots, b_n) = (a_1 b_1, \dots, a_n b_n)$. Remark that $TS(S) \subset TS(\Delta)$. Further, each tree $\mathbf{t} \in TS(\Delta)$ has $s(\mathbf{t})$ different variable indices $j_1, \dots, j_{s(\mathbf{t})}$ while a tree $\mathbf{u} \in TS(S)$ has only $s(\mathbf{u})/2$ different variable indices. Then Proposition 6.1 in [12] holds:

PROPOSITION 3.2. *Let $(Y(t))_{t \in \mathcal{I}_h}$ be defined by the SRK method* (2.1). *Assume that for the random variables holds $\theta_\iota(h) = \sqrt{h} \cdot \vartheta_\iota$ for $\iota \in \mathcal{M}$ with some bounded random variables ϑ_ι. Then for $p \in \mathbb{N}_0$, $f, a^i, b^{i,j} \in C_P^{2(p+1)}(\mathbb{R}^d, \mathbb{R})$ for $i = 1, \dots, d$, $j = 1, \dots, m$ and for $t \in [t_0, T]$ with $h = t - t_0$ holds:*

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

4 Order conditions for stochastic Runge–Kutta methods.

Now, we apply the rooted tree expansions of the solution and the approximation processes in order to yield order conditions for the SRK method (2.1).

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

Note that in case of $m = 1$ we have $\beta(\mathbf{t}) = 1$ for all $\mathbf{t} \in TS(S)$. For example, for $\mathbf{t} = (\sigma_{j_1}, \sigma_{j_2}, \{\sigma_{j_2}\}_{j_1}) \in TS(S)$ 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.

The following theorem yields conditions for the coefficients and the random variables of the SRK method (2.1) such that convergence with some order p in the weak sense is assured (see Theorem 6.4 in [12]).

THEOREM 4.1. *For $p \in \mathbb{N}$ let $a^i \in C_P^{p+1,2p+2}(\mathcal{I} \times \mathbb{R}^d, \mathbb{R})$ and let $b^{i,j} \in C_P^{p+1,2p+3}(\mathcal{I} \times \mathbb{R}^d, \mathbb{R})$ for $i = 1, \dots, d$, $j = 1, \dots, m$. Then the SRK method* (2.1) *with step size h is of weak order p, if for all $\mathbf{t} \in TS(\Delta)$ with $\rho(\mathbf{t}) \leq p + \frac{1}{2}$ and all correlations of type $j_k = j_l$ or $j_k \neq j_l$, $1 \leq k < l \leq s(\mathbf{t})$, between the indices $j_1, \dots, j_{s(\mathbf{t})} \in \{1, \dots, m\}$ of $\mathbf{t}$ holds*

$$\mathrm{E}(\Phi_S(\mathbf{t})) = \frac{\alpha_S(\mathbf{t}) \cdot (l(\mathbf{t}) - 1)! \cdot h^{\rho(\mathbf{t})}}{\alpha_\Delta(\mathbf{t}) \cdot \beta(\mathbf{t}) \cdot \gamma(\mathbf{t}) \cdot 2^{s(\mathbf{t})/2} \cdot \rho(\mathbf{t})!} \tag{4.1}$$

provided (2.2) *and* (2.3) *hold and if the approximation Y has uniformly bounded moments w.r.t. the number N of steps.*

REMARK 4.1. The approximation Y by the SRK method (2.1) has uniformly bounded moments if bounded random variables are used by the method, if (2.3) is fulfilled and if $\mathrm{E}({z^{(k,\nu)}}^T e) = 0$ holds for $1 \leq k \leq m$ and $\nu \in \mathcal{M}$ (see [12]). Further, Theorem 4.1 provides uniform weak convergence with order p in the case of a non-random time discretization $\mathcal{I}_h$ [12].

5 Order two stochastic Runge–Kutta methods.

In the present section, we consider second order SRK schemes for the weak approximation of the solution of Stratonovich SDEs (1.1). Therefore, we consider a new class of SRK methods where the number of stages is independent of the dimension m of the driving Wiener process. Thus, we define the d-dimensional approximation process Y with $Y_n = Y(t_n)$ for $t_n \in \mathcal{I}_h$ by the following SRK method of s stages with $Y_0 = x_0$ and

$$\begin{aligned} Y_{n+1} = Y_n &+ \sum_{i=1}^{s} \alpha_i a\big(t_n + c_i^{(0)} h_n, H_i^{(0)}\big) h_n \\ &+ \sum_{i=1}^{s} \sum_{k=1}^{m} \beta_i^{(1)} b^k \big(t_n + c_i^{(1)} h_n, H_i^{(k)}\big) \hat{I}_{(k)} \\ &+ \sum_{i=1}^{s} \sum_{k=1}^{m} \beta_i^{(2)} b^k \big(t_n + c_i^{(2)} h_n, \hat{H}_i^{(k)}\big) \sqrt{h_n} \end{aligned} \tag{5.1}$$

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

$$\begin{aligned}
H_i^{(0)} &= Y_n + \sum_{j=1}^{s} A_{ij}^{(0)}\, a\big(t_n + c_j^{(0)} h_n, H_j^{(0)}\big) h_n \\
&\quad + \sum_{j=1}^{s} \sum_{l=1}^{m} B_{ij}^{(0)}\, b^l\big(t_n + c_j^{(1)} h_n, H_j^{(l)}\big)\, \hat{I}_{(l)} \\
H_i^{(k)} &= Y_n + \sum_{j=1}^{s} A_{ij}^{(1)}\, a\big(t_n + c_j^{(0)} h_n, H_j^{(0)}\big) h_n \\
&\quad + \sum_{j=1}^{s} B_{ij}^{(1)}\, b^k\big(t_n + c_j^{(1)} h_n, H_j^{(k)}\big)\, \hat{I}_{(k)} \\
&\quad + \sum_{j=1}^{s} \sum_{\substack{l=1 \\ l \neq k}}^{m} B_{ij}^{(3)}\, b^l\big(t_n + c_j^{(1)} h_n, H_j^{(l)}\big)\, \hat{I}_{(l)} \\
\hat{H}_i^{(k)} &= Y_n + \sum_{j=1}^{s} A_{ij}^{(2)}\, a\big(t_n + c_j^{(0)} h_n, H_j^{(0)}\big) h_n \\
&\quad + \sum_{j=1}^{s} \sum_{\substack{l=1 \\ l \neq k}}^{m} B_{ij}^{(2)}\, b^l\big(t_n + c_j^{(1)} h_n, H_j^{(l)}\big)\, \frac{\hat{I}_{(k,l)}}{\sqrt{h_n}}
\end{aligned}$$

for $i = 1, \dots, s$ and $k = 1, \dots, m$. The random variables are defined by

$$\hat{I}_{(k,l)} = \begin{cases} \hat{I}_{(k)}\, \tilde{I}_{(l)} & \text{if } l < k \\ -\hat{I}_{(l)}\, \tilde{I}_{(k)} & \text{if } k < l \end{cases} \tag{5.2}$$

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

$$\mathrm{E}\left(\hat{I}_{(k)}^q\right) = \begin{cases} 0 & \text{for } q \in \{1,3,5\} \\ (q-1) h_n^{q/2} & \text{for } q \in \{2,4\} \\ \mathcal{O}(h_n^{q/2}) & \text{for } q \geq 6 \end{cases} \tag{5.3}$$

and $\tilde{I}_{(k)}$, $1 \leq k \leq m-1$, having the moments

$$\mathrm{E}\left(\tilde{I}_{(k)}^q\right) = \begin{cases} 0 & \text{for } q \in \{1,3\} \\ h_n & \text{for } q = 2 \\ \mathcal{O}(h_n^{q/2}) & \text{for } q \geq 4 \end{cases} . \tag{5.4}$$

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

The coefficients of the SRK method (5.1) can be represented by an extended Butcher array taking the form

$c^{(0)}$	$A^{(0)}$	$B^{(0)}$	
$c^{(1)}$	$A^{(1)}$	$B^{(1)}$	$B^{(3)}$
$c^{(2)}$	$A^{(2)}$	$B^{(2)}$	
	α^T	${\beta^{(1)}}^T$	${\beta^{(2)}}^T$

Applying the rooted tree analysis presented in Section 3 and Section 4, we can calculate order conditions for the SRK method (5.1).

THEOREM 5.1. *Let $a^i \in C_P^{2,4}(\mathcal{I} \times \mathbb{R}^d, \mathbb{R})$ and $b^{i,j} \in C_P^{2,5}(\mathcal{I} \times \mathbb{R}^d, \mathbb{R})$ for $i = 1, \dots, d$, $j = 1, \dots, m$. If the coefficients of the stochastic Runge–Kutta method* (5.1) *fulfill the equations*

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

then the method attains order 1.0 *for the weak approximation of the solution of the Stratonovich SDE* (1.1)*. Further, if $a^i \in C_P^{3,6}(\mathcal{I} \times \mathbb{R}^d, \mathbb{R})$ and $b^{i,j} \in C_P^{3,7}(\mathcal{I} \times \mathbb{R}^d, \mathbb{R})$ for $1 \le i \le d$, $1 \le j \le m$ and if in addition the equations*

7. $\alpha^T A^{(0)} e = \frac{1}{2}$
8. $\alpha^T (B^{(0)} (B^{(1)} e)) = \frac{1}{4}$
9. $\alpha^T (B^{(0)} e)^2 = \frac{1}{2}$
10. $({\beta^{(1)}}^T e)(\alpha^T B^{(0)} e) = \frac{1}{2}$
11. $({\beta^{(1)}}^T e)({\beta^{(1)}}^T A^{(1)} e) = \frac{1}{2}$
12. ${\beta^{(1)}}^T (B^{(1)} (A^{(1)} e)) = \frac{1}{4}$
13. ${\beta^{(1)}}^T ((B^{(1)} e)(A^{(1)} e)) = \frac{1}{4}$
14. ${\beta^{(1)}}^T B^{(3)} e = \frac{1}{2}$
15. $({\beta^{(1)}}^T e)({\beta^{(1)}}^T (B^{(1)} e)^2) = \frac{1}{3}$
16. $({\beta^{(1)}}^T e)({\beta^{(1)}}^T (B^{(3)} e)^2) = \frac{1}{2}$
17. ${\beta^{(1)}}^T (B^{(3)} (B^{(3)} e)) = 0$
18. $({\beta^{(2)}}^T B^{(2)} e)^2 = \frac{1}{4}$
19. ${\beta^{(1)}}^T (B^{(1)} e)^3 = \frac{1}{4}$
20. ${\beta^{(1)}}^T (B^{(1)} (B^{(1)} e)^2) = \frac{1}{12}$
21. ${\beta^{(1)}}^T (B^{(1)} (B^{(3)} e)^2) = \frac{1}{4}$
22. ${\beta^{(1)}}^T (A^{(1)} (B^{(0)} e)) = 0$
23. ${\beta^{(2)}}^T (A^{(2)} e)^2 = 0$
24. ${\beta^{(2)}}^T (A^{(2)} (A^{(0)} e)) = 0$
25. ${\beta^{(1)}}^T (B^{(1)} (B^{(1)} (B^{(1)} e))) = \frac{1}{24}$
26. ${\beta^{(2)}}^T (A^{(2)} (B^{(0)} e)) = 0$
27. ${\beta^{(2)}}^T (A^{(2)} (B^{(0)} e)^2) = 0$
28. ${\beta^{(2)}}^T (B^{(2)} e)^4 = 0$
29. ${\beta^{(2)}}^T (B^{(2)} (B^{(1)} e))^2 = 0$
30. ${\beta^{(2)}}^T (B^{(2)} (B^{(3)} e))^2 = 0$
31. ${\beta^{(1)}}^T ((B^{(1)} e)(B^{(3)} e)^2) = \frac{1}{4}$
32. ${\beta^{(2)}}^T ((A^{(2)} e)(B^{(2)} e)^2) = 0$
33. ${\beta^{(1)}}^T (B^{(1)} (B^{(3)} (B^{(1)} e))) = \frac{1}{8}$
34. ${\beta^{(1)}}^T (B^{(3)} (B^{(3)} (B^{(3)} e))) = 0$
35. ${\beta^{(1)}}^T (B^{(3)} (B^{(1)} (B^{(3)} e))) = 0$
36. ${\beta^{(2)}}^T (A^{(2)} (B^{(0)} (B^{(1)} e))) = 0$

37. $(\beta^{(1)^T}e)(\beta^{(1)^T}((B^{(3)}e)(B^{(1)}e))) = \frac{1}{4}$
38. $(\beta^{(1)^T}e)(\beta^{(1)^T}(B^{(1)}(B^{(1)}e))) = \frac{1}{6}$
39. $(\beta^{(1)^T}e)(\beta^{(1)^T}(B^{(3)}(B^{(1)}e))) = \frac{1}{4}$
40. $(\beta^{(1)^T}e)(\beta^{(1)^T}(B^{(1)}(B^{(3)}e))) = \frac{1}{4}$
41. $\beta^{(1)^T}((B^{(1)}e)(B^{(1)}(B^{(1)}e))) = \frac{1}{8}$
42. $\beta^{(1)^T}((B^{(1)}e)(B^{(3)}(B^{(1)}e))) = \frac{1}{8}$
43. $\beta^{(1)^T}((B^{(3)}e)(B^{(1)}(B^{(3)}e))) = \frac{1}{4}$
44. $\beta^{(1)^T}((B^{(3)}e)(B^{(3)}(B^{(3)}e))) = 0$
45. $\beta^{(1)^T}(B^{(3)}((B^{(3)}e)(B^{(1)}e))) = 0$
46. $\beta^{(2)^T}((B^{(2)}(A^{(1)}e))(B^{(2)}e)) = 0$
47. $\beta^{(2)^T}((B^{(2)}e)(B^{(2)}(B^{(1)}e))) = 0$
48. $\beta^{(2)^T}((B^{(2)}e)(B^{(2)}(B^{(3)}e))) = 0$
49. $\beta^{(2)^T}((B^{(2)}e)(B^{(2)}((B^{(1)}e)^2))) = 0$
50. $\beta^{(2)^T}((B^{(2)}e)(B^{(2)}((B^{(3)}e)^2))) = 0$
51. $\beta^{(2)^T}((B^{(2)}e)(B^{(2)}((B^{(1)}e)(B^{(3)}e)))) = 0$
52. $\beta^{(2)^T}((B^{(2)}e)(B^{(2)}(B^{(1)}(B^{(1)}e)))) = 0$
53. $\beta^{(2)^T}((B^{(2)}e)(B^{(2)}(B^{(3)}(B^{(1)}e)))) = 0$
54. $\beta^{(2)^T}((B^{(2)}e)(B^{(2)}(B^{(3)}(B^{(3)}e)))) = 0$
55. $\beta^{(2)^T}((B^{(2)}e)(B^{(2)}(B^{(1)}(B^{(3)}e)))) = 0$

are fulfilled and if $c^{(i)} = A^{(i)}e$ *for* $i = 0, 1, 2$*, then the stochastic Runge–Kutta method* (5.1) *attains order* 2.0 *for the weak approximation of the solution of the Stratonovich SDE* (1.1)*.*

REMARK 5.1. The 55 conditions of Theorem 5.1 reduce to 17 conditions for $m = 1$ since we do not need $\hat{H}_i^{(k)}$ if we choose $A_{ij}^{(2)} = 0$ and we do not need $B_{ij}^{(3)}$ as well. Further we need $s \geq 4$ for an explicit SRK method of order 2.0 (see [14]).

PROOF. Firstly, we show hat the SRK method (5.1) is contained in the general class (2.1). Therefore, we choose $\mathcal{M} = \{(k), (k,l) : 0 \leq k, l \leq m\}$ and

$$\gamma_i^{(\iota)^{(k,\nu)}}\theta_\iota(h_n) = \begin{cases} \beta_i^{(1)}\,\hat{I}_{(k)} & \text{if } 0 < \iota = k,\ \nu = 0 \\ \beta_i^{(2)}\,\sqrt{h_n} & \text{if } 0 = \iota < k,\ \nu = 1 \\ 0 & \text{otherwise} \end{cases}$$

$$A_{ij}^{(k,\nu)(0,0)} h_n = \begin{cases} A_{ij}^{(0)} h_n & \text{if } k=\nu=0 \\ A_{ij}^{(1)} h_n & \text{if } k>0,\, \nu=0 \\ A_{ij}^{(2)} h_n & \text{if } k>0,\, \nu=1 \\ 0 & \text{otherwise} \end{cases}$$

$$B_{ij}^{(\iota)}{}^{(k,\nu)(r,\mu)} \theta_\iota(h_n) = \begin{cases} B_{ij}^{(0)}\, \hat{I}_{(r)} & \text{if } k=\nu=\mu=0,\, \iota=r \\ B_{ij}^{(1)}\, \hat{I}_{(k)} & \text{if } k=r=\iota,\, \nu=\mu=0 \\ B_{ij}^{(2)}\, \frac{\hat{I}_{(k,r)}}{\sqrt{h_n}} & \text{if } \iota=(k,r),\, 0<k\neq r,\, \nu=1,\, \mu=0 \\ B_{ij}^{(3)}\, \hat{I}_{(r)} & \text{if } 0<k\neq r,\, \iota=r,\, \nu=\mu=0 \\ 0 & \text{otherwise} \end{cases}$$

for $k=0,1,\ldots,m$, $r=1,\ldots,m$ and $\iota,\nu,\mu\in\mathcal{M}$. As a result of this we have

$$z_i^{(0,0)} = \alpha_i\, h_n \qquad z_i^{(k,0)} = \beta_i^{(1)}\, \hat{I}_{(k)} \qquad z_i^{(k,1)} = \beta_i^{(2)}\, \sqrt{h_n}$$

$$Z_{ij}^{(0,0)(0,0)} = A_{ij}^{(0)}\, h_n \qquad Z_{ij}^{(k,0)(0,0)} = A_{ij}^{(1)}\, h_n \qquad Z_{ij}^{(k,1)(0,0)} = A_{ij}^{(2)}\, h_n$$

$$Z_{ij}^{(0,0)(k,0)} = B_{ij}^{(0)}\, \hat{I}_{(k)} \qquad Z_{ij}^{(k,0)(k,0)} = B_{ij}^{(1)}\, \hat{I}_{(k)} \qquad Z_{ij}^{(k,1)(l,0)} = B_{ij}^{(2)}\, \frac{\hat{I}_{(k,l)}}{\sqrt{h_n}}$$

$$Z_{ij}^{(k,0)(l,0)} = B_{ij}^{(3)}\, \hat{I}_{(l)}$$

for $1\leq k,l\leq m$ with $k\neq l$. Further, we have $H_i^{(0,0)} = H_i^{(0)}$, $H_i^{(k,0)} = H_i^{(k)}$ and $H_i^{(k,1)} = \hat{H}_i^{(k)}$. Now, apply Theorem 4.1 for all $\mathbf{t}\in TS(\Delta)$ with $\rho(\mathbf{t})\leq 2.5$. We refer to [12] for all necessary trees and corresponding parameters $\alpha_S(\mathbf{t})$, $\alpha_\Delta(\mathbf{t})$ or $\beta(\mathbf{t})$. Apart from (5.3) and (5.4), the following moments are helpful in the subsequent calculations: $\mathrm{E}(\hat{I}_{(k,l)}^2) = h^2$, $\mathrm{E}(\hat{I}_{(k,l)}\hat{I}_{(l,k)}) = -h^2$, $\mathrm{E}(\hat{I}_{(k)}\hat{I}_{(l)}\hat{I}_{(k,l)}) = 0$ and $\mathrm{E}(\hat{I}_{(k,l)}^q) = 0$ for $q=1,3$ and $k\neq l$. In the following, it holds $\beta(\mathbf{t})=1$ if not stated otherwise and we write $h=h_n$.

Order 0.5 trees.
$\mathbf{t}_{0.5.1} = (\sigma_{j_1})$: $\Phi_S(\mathbf{t}) = z^{(j_1,0)^T} e + z^{(j_1,1)^T} e$
With $\alpha_S(\mathbf{t})=0$ follows $\mathrm{E}(\Phi_S(\mathbf{t}))=0 \Leftrightarrow \beta^{(2)^T} e\sqrt{h}=0$.

In the following, we assume that Condition 3. of Theorem 5.1 holds.

Order 1.0 trees.
$\mathbf{t}_{1.1} = (\tau)$: $\Phi_S(\mathbf{t}) = z^{(0,0)^T} e$
With $\alpha_S(\mathbf{t}) = \alpha_\Delta(\mathbf{t}) = 1$ follows $\mathrm{E}(\Phi_S(\mathbf{t})) = h \Leftrightarrow \alpha^T e\, h = h$.

$\mathbf{t}_{1.2} = (\sigma_{j_1}, \sigma_{j_2})$: $\Phi_S(\mathbf{t}) = (z^{(j_1,0)^T} e + z^{(j_1,1)^T} e)(z^{(j_2,0)^T} e + z^{(j_2,1)^T} e)$
For $j_1=j_2$ with $\alpha_S(\mathbf{t}) = \alpha_\Delta(\mathbf{t}) = 1$ follows $\mathrm{E}(\Phi_S(\mathbf{t})) = h \Leftrightarrow$
$(\beta^{(1)^T} e)^2\ \mathrm{E}(\hat{I}_{(j_1)}^2) + (\beta^{(2)^T} e)^2\, h = h$.

$\mathbf{t}_{1.3} = (\{\sigma_{j_2}\}_{j_1})$: $\Phi_S(\mathbf{t}) = z^{(j_1,0)^T} Z^{(j_1,0)(j_2,0)} e + z^{(j_1,1)^T} Z^{(j_1,1)(j_2,0)} e$
For $j_1 = j_2$ with $\alpha_S(\mathbf{t}) = \alpha_\Delta(\mathbf{t}) = 1$ follows $\mathrm{E}(\Phi_S(\mathbf{t})) = \frac{1}{2}h \Leftrightarrow$
$\beta^{(1)^T} B^{(1)} e \; \mathrm{E}(\hat{I}^2_{(j_1)}) = \frac{1}{2}h$.

Now, we additionally assume that Conditions 1., 2. and 4. of Theorem 5.1 hold.

Order 1.5 trees.
$\mathbf{t}_{1.5.2} = (\{\tau\}_{j_1})$: $\Phi_S(\mathbf{t}) = z^{(j_1,0)^T} Z^{(j_1,0)(0,0)} e + z^{(j_1,1)^T} Z^{(j_1,1)(0,0)} e$
With $\alpha_S(\mathbf{t}) = 0$ follows $\mathrm{E}(\Phi_S(\mathbf{t})) = 0 \Leftrightarrow \beta^{(2)^T} A^{(2)} e \sqrt{h}\, h = 0$.

$\mathbf{t}_{1.5.6} = (\{\sigma_{j_2}, \sigma_{j_3}\}_{j_1})$:

$$\begin{aligned} \Phi_S(\mathbf{t}) &= z^{(j_1,0)^T} ((Z^{(j_1,0)(j_2,0)} e)(Z^{(j_1,0)(j_3,0)} e)) \\ &\quad + z^{(j_1,1)^T} ((Z^{(j_1,1)(j_2,0)} e)(Z^{(j_1,1)(j_3,0)} e)) \end{aligned}$$

For $j_1 \neq j_2 = j_3$ with $\alpha_S(\mathbf{t}) = 0$ follows $\mathrm{E}(\Phi_S(\mathbf{t})) = 0 \Leftrightarrow$
$\beta^{(2)^T} (B^{(2)} e)^2 \; \mathrm{E}(\hat{I}^2_{(j_1,j_2)}) h^{-1/2} = 0$.

For the trees $\mathbf{t}_{1.5.1} = ([\sigma_{j_1}])$, $\mathbf{t}_{1.5.3} = (\tau, \sigma_{j_1})$, $\mathbf{t}_{1.5.4} = (\sigma_{j_1}, \sigma_{j_2}, \sigma_{j_3})$, $\mathbf{t}_{1.5.5} = (\{\sigma_{j_2}\}_{j_1}, \sigma_{j_3})$ and $\mathbf{t}_{1.5.7} = (\{\{\sigma_{j_3}\}_{j_2}\}_{j_1})$ holds $\alpha_S(\mathbf{t}) = 0$ and $\mathrm{E}(\Phi_S(\mathbf{t})) = 0$.

Order 2.0 trees.
$\mathbf{t}_{2.1} = ([\tau])$: $\Phi_S(\mathbf{t}) = z^{(0,0)^T} Z^{(0,0)(0,0)} e$
With $\alpha_S(\mathbf{t}) = \alpha_\Delta(\mathbf{t}) = 1$ follows $\mathrm{E}(\Phi_S(\mathbf{t})) = \frac{1}{2}h^2 \Leftrightarrow \alpha^T A^{(0)} e\, h^2 = \frac{1}{2}h^2$.

$\mathbf{t}_{2.2} = (\tau, \tau)$: $\Phi_S(\mathbf{t}) = (z^{(0,0)^T} e)^2$
With $\alpha_S(\mathbf{t}) = \alpha_\Delta(\mathbf{t}) = 1$ follows $\mathrm{E}(\Phi_S(\mathbf{t})) = h^2 \Leftrightarrow (\alpha^T e)^2\, h^2 = h^2$.

$\mathbf{t}_{2.3} = ([\{\sigma_{j_2}\}_{j_1}])$: $\Phi_S(\mathbf{t}) = z^{(0,0)^T} (Z^{(0,0)(j_1,0)} (Z^{(j_1,0)(j_2,0)} e))$
For $j_1 = j_2$ with $\alpha_S(\mathbf{t}) = \alpha_\Delta(\mathbf{t}) = 1$ follows $\mathrm{E}(\Phi_S(\mathbf{t})) = \frac{1}{4}h^2 \Leftrightarrow$
$\alpha^T (B^{(0)} (B^{(1)} e)) \; \mathrm{E}(\hat{I}^2_{(j_1)}) = \frac{1}{4}h^2$.

$\mathbf{t}_{2.4} = ([\sigma_{j_1}, \sigma_{j_2}])$: $\Phi_S(\mathbf{t}) = z^{(0,0)^T} ((Z^{(0,0)(j_1,0)} e)(Z^{(0,0)(j_2,0)} e))$
For $j_1 = j_2$ with $\alpha_S(\mathbf{t}) = \alpha_\Delta(\mathbf{t}) = 1$ follows $\mathrm{E}(\Phi_S(\mathbf{t})) = \frac{1}{2}h^2 \Leftrightarrow$
$\alpha^T (B^{(0)} e)^2 \; \mathrm{E}(\hat{I}^2_{(j_1)}) = \frac{1}{2}h^2$.

$\mathbf{t}_{2.5} = (\sigma_{j_1}, [\sigma_{j_2}])$: $\Phi_S(\mathbf{t}) = (z^{(j_1,0)^T} e + z^{(j_1,1)^T} e) \, (z^{(0,0)^T} Z^{(0,0)(j_2,0)} e)$
For $j_1 = j_2$ with $\alpha_S(\mathbf{t}) = 2$ and $\alpha_\Delta(\mathbf{t}) = 3$ follows $\mathrm{E}(\Phi_S(\mathbf{t})) = \frac{1}{2}h^2 \Leftrightarrow$
$(\beta^{(1)^T} e) \, (\alpha^T B^{(0)} e) \; \mathrm{E}(\hat{I}^2_{(j_1)}) = \frac{1}{2}h^2$.

$\mathbf{t}_{2.6} = (\{\sigma_{j_2}\}_{j_1}, \tau)$: $\Phi_S(\mathbf{t}) = (z^{(j_1,0)^T} Z^{(j_1,0)(j_2,0)} e + z^{(j_1,1)^T} Z^{(j_1,1)(j_2,0)} e) \, (z^{(0,0)^T} e)$
For $j_1 = j_2$ with $\alpha_S(\mathbf{t}) = 2$ and $\alpha_\Delta(\mathbf{t}) = 3$ follows $\mathrm{E}(\Phi_S(\mathbf{t})) = \frac{1}{2}h^2 \Leftrightarrow$
$(\beta^{(1)^T} B^{(1)} e)(\alpha^T e) \; \mathrm{E}(\hat{I}^2_{(j_1)}) h = \frac{1}{2}h^2$.

$\mathbf{t}_{2.7} = (\sigma_{j_1}, \sigma_{j_2}, \tau)$: $\Phi_S(\mathbf{t}) = (z^{(j_1,0)^T} e + z^{(j_1,1)^T} e)(z^{(j_2,0)^T} e + z^{(j_2,1)^T} e)(z^{(0,0)^T} e)$
For $j_1 = j_2$ with $\alpha_S(\mathbf{t}) = 2$ and $\alpha_\Delta(\mathbf{t}) = 3$ follows $\mathrm{E}(\Phi_S(\mathbf{t})) = h^2 \Leftrightarrow$
$(\beta^{(1)^T} e)^2 \; \mathrm{E}(\hat{I}^2_{(j_1)}) = h^2$.

$\mathbf{t}_{2.8} = (\sigma_{j_1}, \{\tau\}_{j_2})$:
$\Phi_S(\mathbf{t}) = ({z^{(j_1,0)}}^T e + {z^{(j_1,1)}}^T e)({z^{(j_2,0)}}^T Z^{(j_2,0)(0,0)} e + {z^{(j_2,1)}}^T Z^{(j_2,1)(0,0)} e)$
For $j_1 = j_2$ with $\alpha_S(\mathbf{t}) = 2$ and $\alpha_\Delta(\mathbf{t}) = 3$ follows $\mathrm{E}(\Phi_S(\mathbf{t})) = \frac{1}{2}h^2 \Leftrightarrow ({\beta^{(1)}}^T e)({\beta^{(1)}}^T A^{(1)} e)\ \mathrm{E}(\hat{I}^2_{(j_1)})h = \frac{1}{2}h^2$.

$\mathbf{t}_{2.9} = (\{\{\tau\}_{j_2}\}_{j_1})$:
$\Phi_S(\mathbf{t}) = {z^{(j_1,0)}}^T (Z^{(j_1,0)(j_2,0)}(Z^{(j_2,0)(0,0)} e)) + {z^{(j_1,1)}}^T (Z^{(j_1,1)(j_2,0)}(Z^{(j_2,0)(0,0)} e))$
For $j_1 = j_2$ with $\alpha_S(\mathbf{t}) = \alpha_\Delta(\mathbf{t}) = 1$ follows $\mathrm{E}(\Phi_S(\mathbf{t})) = \frac{1}{4}h^2 \Leftrightarrow$
${\beta^{(1)}}^T (B^{(1)}(A^{(1)} e))\ \mathrm{E}(\hat{I}^2_{(j_1)})h = \frac{1}{4}h^2$.

$\mathbf{t}_{2.10} = (\{\sigma_{j_2}, \tau\}_{j_1})$:

$$\Phi_S(\mathbf{t}) = {z^{(j_1,0)}}^T ((Z^{(j_1,0)(j_2,0)} e)(Z^{(j_1,0)(0,0)} e)) + {z^{(j_1,1)}}^T ((Z^{(j_1,1)(j_2,0)} e)(Z^{(j_1,1)(0,0)} e))$$

For $j_1 = j_2$ with $\alpha_S(\mathbf{t}) = 1$ and $\alpha_\Delta(\mathbf{t}) = 2$ follows $\mathrm{E}(\Phi_S(\mathbf{t})) = \frac{1}{4}h^2 \Leftrightarrow {\beta^{(1)}}^T ((B^{(1)} e)(A^{(1)} e))\ \mathrm{E}(\hat{I}^2_{(j_1)})h = \frac{1}{4}h^2$.

$\mathbf{t}_{2.11} = (\sigma_{j_1}, \sigma_{j_2}, \sigma_{j_3}, \sigma_{j_4})$:

$$\Phi_S(\mathbf{t}) = ({z^{(j_1,0)}}^T e + {z^{(j_1,1)}}^T e)({z^{(j_2,0)}}^T e + {z^{(j_2,1)}}^T e) \times ({z^{(j_3,0)}}^T e + {z^{(j_3,1)}}^T e)({z^{(j_4,0)}}^T e + {z^{(j_4,1)}}^T e)$$

Case A): For $j_1 = j_2 = j_3 = j_4$ with $\alpha_S(\mathbf{t}) = \alpha_\Delta(\mathbf{t}) = \beta(\mathbf{t}) = 1$ follows $\mathrm{E}(\Phi_S(\mathbf{t})) = 3h^2 \Leftrightarrow ({\beta^{(1)}}^T e)^4\ \mathrm{E}(\hat{I}^4_{(j_1)}) = 3h^2$.
Case B): For $j_1 = j_2 \neq j_3 = j_4$ with $\alpha_S(\mathbf{t}) = \alpha_\Delta(\mathbf{t}) = 1$ and $\beta(\mathbf{t}) = 3$ follows $\mathrm{E}(\Phi_S(\mathbf{t})) = h^2 \Leftrightarrow ({\beta^{(1)}}^T e)^2({\beta^{(1)}}^T e)^2\ \mathrm{E}(\hat{I}^2_{(j_1)}\hat{I}^2_{(j_3)}) = h^2$.

$\mathbf{t}_{2.12} = (\sigma_{j_1}, \sigma_{j_2}, \{\sigma_{j_4}\}_{j_3})$:

$$\Phi_S(\mathbf{t}) = ({z^{(j_1,0)}}^T e + {z^{(j_1,1)}}^T e)({z^{(j_2,0)}}^T e + {z^{(j_2,1)}}^T e) \times ({z^{(j_3,0)}}^T Z^{(j_3,0)(j_4,0)} e + {z^{(j_3,1)}}^T Z^{(j_3,1)(j_4,0)} e)$$

Case A): For $j_1 = j_2 = j_3 = j_4$ with $\alpha_S(\mathbf{t}) = \alpha_\Delta(\mathbf{t}) = 6$ and $\beta(\mathbf{t}) = 1$ follows $\mathrm{E}(\Phi_S(\mathbf{t})) = \frac{3}{2}h^2 \Leftrightarrow ({\beta^{(1)}}^T e)^2({\beta^{(1)}}^T B^{(1)} e)\ \mathrm{E}(\hat{I}^4_{(j_1)}) = \frac{3}{2}h^2$.
Case B): For $j_1 = j_3 \neq j_2 = j_4$ with $\alpha_S(\mathbf{t}) = 4$, $\alpha_\Delta(\mathbf{t}) = 6$ and $\beta(\mathbf{t}) = 2$ follows $\mathrm{E}(\Phi_S(\mathbf{t})) = \frac{1}{2}h^2 \Leftrightarrow ({\beta^{(1)}}^T e)^2({\beta^{(1)}}^T B^{(3)} e)\ \mathrm{E}(\hat{I}^2_{(j_1)}\hat{I}^2_{(j_2)}) = \frac{1}{2}h^2$.
Case C): For $j_1 = j_2 \neq j_3 = j_4$ with $\alpha_S(\mathbf{t}) = 2$, $\alpha_\Delta(\mathbf{t}) = 6$ and $\beta(\mathbf{t}) = 1$ follows $\mathrm{E}(\Phi_S(\mathbf{t})) = \frac{1}{2}h^2 \Leftrightarrow ({\beta^{(1)}}^T e)^2({\beta^{(1)}}^T B^{(1)} e)\ \mathrm{E}(\hat{I}^2_{(j_1)}\hat{I}^2_{(j_3)}) = \frac{1}{2}h^2$.

$\mathbf{t}_{2.13} = (\sigma_{j_1}, \{\sigma_{j_3}, \sigma_{j_4}\}_{j_2})$:

$$\Phi_S(\mathbf{t}) = ({z^{(j_1,0)}}^T e + {z^{(j_1,1)}}^T e)({z^{(j_2,0)}}^T ((Z^{(j_2,0)(j_3,0)} e)(Z^{(j_2,0)(j_4,0)} e)) + {z^{(j_2,1)}}^T ((Z^{(j_2,1)(j_3,0)} e)(Z^{(j_2,1)(j_4,0)} e)))$$

Case A): For $j_1 = j_2 = j_3 = j_4$ with $\alpha_S(\mathbf{t}) = \alpha_\Delta(\mathbf{t}) = 4$ and $\beta(\mathbf{t}) = 1$ follows $\mathrm{E}(\Phi_S(\mathbf{t})) = h^2 \Leftrightarrow ({\beta^{(1)}}^T e)({\beta^{(1)}}^T (B^{(1)}e)^2)\ \mathrm{E}(\hat{I}^4_{(j_1)}) = h^2$.

Case B): For $j_1 = j_2 \neq j_3 = j_4$ with $\alpha_S(\mathbf{t}) = 2$, $\alpha_\Delta(\mathbf{t}) = 4$ and $\beta(\mathbf{t}) = 1$ follows $\mathrm{E}(\Phi_S(\mathbf{t})) = \frac{1}{2}h^2 \Leftrightarrow ({\beta^{(1)}}^T e)({\beta^{(1)}}^T (B^{(3)}e)^2)\ \mathrm{E}(\hat{I}^2_{(j_1)}\hat{I}^2_{(j_3)}) = \frac{1}{2}h^2$.

Case C): For $j_1 = j_3 \neq j_2 = j_4$ with $\alpha_S(\mathbf{t}) = 2$, $\alpha_\Delta(\mathbf{t}) = 4$ and $\beta(\mathbf{t}) = 2$ follows $\mathrm{E}(\Phi_S(\mathbf{t})) = \frac{1}{4}h^2 \Leftrightarrow ({\beta^{(1)}}^T e)({\beta^{(1)}}^T ((B^{(3)}e)(B^{(1)}e)))\ \mathrm{E}(\hat{I}^2_{(j_1)}\hat{I}^2_{(j_2)}) = \frac{1}{4}h^2$.

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

$$\Phi_S(\mathbf{t}_{2.14}) = ({z^{(j_1,0)}}^T e + {z^{(j_1,1)}}^T e)({z^{(j_2,0)}}^T (Z^{(j_2,0)(j_3,0)}(Z^{(j_3,0)(j_4,0)}e)) + {z^{(j_2,1)}}^T (Z^{(j_2,1)(j_3,0)}(Z^{(j_3,0)(j_4,0)}e)))$$

Case A): For $j_1 = j_2 = j_3 = j_4$ with $\alpha_S(\mathbf{t}) = \alpha_\Delta(\mathbf{t}) = 4$ follows $\mathrm{E}(\Phi_S(\mathbf{t})) = \frac{1}{2}h^2$ $\Leftrightarrow ({\beta^{(1)}}^T e)({\beta^{(1)}}^T (B^{(1)}(B^{(1)}e)))\ \mathrm{E}(\hat{I}^4_{(j_1)}) = \frac{1}{2}h^2$.

Case B): For $j_1 = j_2 \neq j_3 = j_4$ with $\alpha_S(\mathbf{t}) = 2$ and $\alpha_\Delta(\mathbf{t}) = 4$ follows $\mathrm{E}(\Phi_S(\mathbf{t}))$ $= \frac{1}{4}h^2 \Leftrightarrow ({\beta^{(1)}}^T e)({\beta^{(1)}}^T (B^{(3)}(B^{(1)}e)))\ \mathrm{E}(\hat{I}^2_{(j_1)}\hat{I}^2_{(j_3)}) = \frac{1}{4}h^2$.

Case C): For $j_1 = j_4 \neq j_2 = j_3$ with $\alpha_S(\mathbf{t}) = 2$ and $\alpha_\Delta(\mathbf{t}) = 4$ follows $\mathrm{E}(\Phi_S(\mathbf{t}))$ $= \frac{1}{4}h^2 \Leftrightarrow ({\beta^{(1)}}^T e)({\beta^{(1)}}^T (B^{(1)}(B^{(3)}e)))\ \mathrm{E}(\hat{I}^2_{(j_1)}\hat{I}^2_{(j_2)}) = \frac{1}{4}h^2$.

Case D): For $j_1 = j_3 \neq j_2 = j_4$ with $\alpha_S(\mathbf{t}) = 0$ follows $\mathrm{E}(\Phi_S(\mathbf{t})) = 0 \Leftrightarrow$ $({\beta^{(1)}}^T e)({\beta^{(1)}}^T (B^{(3)}(B^{(3)}e)))\ \mathrm{E}(\hat{I}^2_{(j_1)}\hat{I}^2_{(j_2)}) = 0$.

$\mathbf{t}_{2.15} = (\{\sigma_{j_2}\}_{j_1}, \{\sigma_{j_4}\}_{j_3})$:

$$\Phi_S(\mathbf{t}) = ({z^{(j_1,0)}}^T Z^{(j_1,0)(j_2,0)}e + {z^{(j_1,1)}}^T Z^{(j_1,1)(j_2,0)}e)({z^{(j_3,0)}}^T Z^{(j_3,0)(j_4,0)}e + {z^{(j_3,1)}}^T Z^{(j_3,1)(j_4,0)}e)$$

Case A): For $j_1 = j_2 = j_3 = j_4$ with $\alpha_S(\mathbf{t}) = \alpha_\Delta(\mathbf{t}) = 3$ follows $\mathrm{E}(\Phi_S(\mathbf{t})) = \frac{3}{4}h^2$ $\Leftrightarrow ({\beta^{(1)}}^T B^{(1)}e)^2\ \mathrm{E}(\hat{I}^4_{(j_1)}) = \frac{3}{4}h^2$.

Case B): For $j_1 = j_3 \neq j_2 = j_4$ with $\alpha_S(\mathbf{t}) = 2$ and $\alpha_\Delta(\mathbf{t}) = 3$ follows $\mathrm{E}(\Phi_S(\mathbf{t}))$ $= \frac{1}{2}h^2 \Leftrightarrow ({\beta^{(1)}}^T B^{(3)}e)^2\ \mathrm{E}(\hat{I}^2_{(j_1)}\hat{I}^2_{(j_2)}) + ({\beta^{(2)}}^T B^{(2)}e)^2\ \mathrm{E}(\hat{I}^2_{(j_1,j_2)}) = \frac{1}{2}h^2$.

Case C): For $j_1 = j_2 \neq j_3 = j_4$ with $\alpha_S(\mathbf{t}) = 1$ and $\alpha_\Delta(\mathbf{t}) = 3$ follows $\mathrm{E}(\Phi_S(\mathbf{t}))$ $= \frac{1}{4}h^2 \Leftrightarrow ({\beta^{(1)}}^T B^{(1)}e)^2\ \mathrm{E}(\hat{I}^2_{(j_1)}\hat{I}^2_{(j_3)}) = \frac{1}{4}h^2$.

Case D): For $j_1 = j_4 \neq j_2 = j_3$ with $\alpha_S(\mathbf{t}) = 0$ follows $\mathrm{E}(\Phi_S(\mathbf{t})) = 0 \Leftrightarrow$ $({\beta^{(1)}}^T B^{(3)}e)^2\ \mathrm{E}(\hat{I}^2_{(j_1)}\hat{I}^2_{(j_2)}) + ({\beta^{(2)}}^T B^{(2)}e)^2\ \mathrm{E}(\hat{I}_{(j_1,j_2)}\hat{I}_{(j_2,j_1)}) = 0$.

$\mathbf{t}_{2.16} = (\{\sigma_{j_2}, \sigma_{j_3}, \sigma_{j_4}\}_{j_1})$:

$$\Phi_S(\mathbf{t}) = {z^{(j_1,0)}}^T ((Z^{(j_1,0)(j_2,0)}e)(Z^{(j_1,0)(j_3,0)}e)(Z^{(j_1,0)(j_4,0)}e)) + {z^{(j_1,1)}}^T ((Z^{(j_1,1)(j_2,0)}e)(Z^{(j_1,1)(j_3,0)}e)(Z^{(j_1,1)(j_4,0)}e))$$

Case A): For $j_1 = j_2 = j_3 = j_4$ with $\alpha_S(\mathbf{t}) = \alpha_\Delta(\mathbf{t}) = 1$ and $\beta(\mathbf{t}) = 1$ follows $\mathrm{E}(\Phi_S(\mathbf{t})) = \frac{3}{4}h^2 \Leftrightarrow {\beta^{(1)}}^T (B^{(1)}e)^3 \ \mathrm{E}(\hat{I}^4_{(j_1)}) = \frac{3}{4}h^2$.

Case B): For $j_1 = j_2 \neq j_3 = j_4$ with $\alpha_S(\mathbf{t}) = \alpha_\Delta(\mathbf{t}) = 1$ and $\beta(\mathbf{t}) = 3$ follows $\mathrm{E}(\Phi_S(\mathbf{t})) = \frac{1}{4}h^2 \Leftrightarrow {\beta^{(1)}}^T ((B^{(1)}e)(B^{(3)}e)^2) \ \mathrm{E}(\hat{I}^2_{(j_1)}\hat{I}^2_{(j_3)}) = \frac{1}{4}h^2$.

$\mathbf{t}_{2.17} = (\{\sigma_{j_2}, \{\sigma_{j_4}\}_{j_3}\}_{j_1})$:

$$\Phi_S(\mathbf{t}) = {z^{(j_1,0)}}^T ((Z^{(j_1,0)(j_2,0)}e)(Z^{(j_1,0)(j_3,0)}(Z^{(j_3,0)(j_4,0)}e))) + {z^{(j_1,1)}}^T ((Z^{(j_1,1)(j_2,0)}e)(Z^{(j_1,1)(j_3,0)}(Z^{(j_3,0)(j_4,0)}e)))$$

Case A): For $j_1 = j_2 = j_3 = j_4$ with $\alpha_S(\mathbf{t}) = \alpha_\Delta(\mathbf{t}) = 3$ follows $\mathrm{E}(\Phi_S(\mathbf{t})) = \frac{3}{8}h^2$ $\Leftrightarrow {\beta^{(1)}}^T ((B^{(1)}e)(B^{(1)}(B^{(1)}e))) \ \mathrm{E}(\hat{I}^4_{(j_1)}) = \frac{3}{8}h^2$.

Case B): For $j_1 = j_2 \neq j_3 = j_4$ with $\alpha_S(\mathbf{t}) = 1$ and $\alpha_\Delta(\mathbf{t}) = 3$ follows $\mathrm{E}(\Phi_S(\mathbf{t}))$ $= \frac{1}{8}h^2 \Leftrightarrow {\beta^{(1)}}^T ((B^{(1)}e)(B^{(3)}(B^{(1)}e))) \ \mathrm{E}(\hat{I}^2_{(j_1)}\hat{I}^2_{(j_3)}) = \frac{1}{8}h^2$.

Case C): For $j_1 = j_3 \neq j_2 = j_4$ with $\alpha_S(\mathbf{t}) = 2$ and $\alpha_\Delta(\mathbf{t}) = 3$ follows $\mathrm{E}(\Phi_S(\mathbf{t}))$ $= \frac{1}{4}h^2 \Leftrightarrow {\beta^{(1)}}^T ((B^{(3)}e)(B^{(1)}(B^{(3)}e))) \ \mathrm{E}(\hat{I}^2_{(j_1)}\hat{I}^2_{(j_2)}) = \frac{1}{4}h^2$.

Case D): For $j_1 = j_4 \neq j_2 = j_3$ with $\alpha_S(\mathbf{t}) = 0$ follows $\mathrm{E}(\Phi_S(\mathbf{t})) = 0 \Leftrightarrow$ ${\beta^{(1)}}^T ((B^{(3)}e)(B^{(3)}(B^{(3)}e))) \ \mathrm{E}(\hat{I}^2_{(j_1)}\hat{I}^2_{(j_2)}) = 0$.

$\mathbf{t}_{2.18} = (\{\{\sigma_{j_3}, \sigma_{j_4}\}_{j_2}\}_{j_1})$:

$$\Phi_S(\mathbf{t}) = {z^{(j_1,0)}}^T (Z^{(j_1,0)(j_2,0)}((Z^{(j_2,0)(j_3,0)}e)(Z^{(j_2,0)(j_4,0)}e))) + {z^{(j_1,1)}}^T (Z^{(j_1,1)(j_2,0)}((Z^{(j_2,0)(j_3,0)}e)(Z^{(j_2,0)(j_4,0)}e)))$$

Case A): For $j_1 = j_2 = j_3 = j_4$ with $\alpha_S(\mathbf{t}) = \alpha_\Delta(\mathbf{t}) = 1$ follows $\mathrm{E}(\Phi_S(\mathbf{t})) = \frac{1}{4}h^2$ $\Leftrightarrow {\beta^{(1)}}^T (B^{(1)}(B^{(1)}e)^2) \ \mathrm{E}(\hat{I}^4_{(j_1)}) = \frac{1}{4}h^2$.

Case B): For $j_1 = j_2 \neq j_3 = j_4$ with $\alpha_S(\mathbf{t}) = \alpha_\Delta(\mathbf{t}) = 1$ follows $\mathrm{E}(\Phi_S(\mathbf{t})) = \frac{1}{4}h^2$ $\Leftrightarrow {\beta^{(1)}}^T (B^{(1)}(B^{(3)}e)^2) \ \mathrm{E}(\hat{I}^2_{(j_1)}\hat{I}^2_{(j_3)}) = \frac{1}{4}h^2$.

Case C): For $j_1 = j_3 \neq j_2 = j_4$ with $\alpha_S(\mathbf{t}) = 0$ follows $\mathrm{E}(\Phi_S(\mathbf{t})) = 0 \Leftrightarrow$ ${\beta^{(1)}}^T (B^{(3)}((B^{(3)}e)(B^{(1)}e))) \ \mathrm{E}(\hat{I}^2_{(j_1)}\hat{I}^2_{(j_2)}) = 0$.

$\mathbf{t}_{2.19} = (\{\{\{\sigma_{j_4}\}_{j_3}\}_{j_2}\}_{j_1})$:

$$\Phi_S(\mathbf{t}) = {z^{(j_1,0)}}^T (Z^{(j_1,0)(j_2,0)}(Z^{(j_2,0)(j_3,0)}(Z^{(j_3,0)(j_4,0)}e))) + {z^{(j_1,1)}}^T (Z^{(j_1,1)(j_2,0)}(Z^{(j_2,0)(j_3,0)}(Z^{(j_3,0)(j_4,0)}e)))$$

Case A): For $j_1 = j_2 = j_3 = j_4$ with $\alpha_S(\mathbf{t}) = \alpha_\Delta(\mathbf{t}) = 1$ follows $\mathrm{E}(\Phi_S(\mathbf{t})) = \frac{1}{8}h^2$ $\Leftrightarrow {\beta^{(1)}}^T (B^{(1)}(B^{(1)}(B^{(1)}e))) \ \mathrm{E}(\hat{I}^4_{(j_1)}) = \frac{1}{8}h^2$.

Case B): For $j_1 = j_2 \neq j_3 = j_4$ with $\alpha_S(\mathbf{t}) = \alpha_\Delta(\mathbf{t}) = 1$ follows $\mathrm{E}(\Phi_S(\mathbf{t})) = \frac{1}{8}h^2$ $\Leftrightarrow {\beta^{(1)}}^T (B^{(1)}(B^{(3)}(B^{(1)}e))) \ \mathrm{E}(\hat{I}^2_{(j_1)}\hat{I}^2_{(j_3)}) = \frac{1}{8}h^2$.

Case C): For $j_1 = j_3 \neq j_2 = j_4$ with $\alpha_S(\mathbf{t}) = 0$ follows $\mathrm{E}(\Phi_S(\mathbf{t})) = 0 \Leftrightarrow$ ${\beta^{(1)}}^T(B^{(3)}(B^{(3)}(B^{(3)}e)))\ \mathrm{E}(\hat{I}^2_{(j_1)}\hat{I}^2_{(j_2)}) = 0$.

Case D): For $j_1 = j_4 \neq j_2 = j_3$ with $\alpha_S(\mathbf{t}) = 0$ follows $\mathrm{E}(\Phi_S(\mathbf{t})) = 0 \Leftrightarrow$ ${\beta^{(1)}}^T(B^{(3)}(B^{(1)}(B^{(3)}e)))\ \mathrm{E}(\hat{I}^2_{(j_1)}\hat{I}^2_{(j_2)}) = 0$.

$\mathbf{t}_{2.20} = (\{[\sigma_{j_2}]\}_{j_1})$:
$\Phi_S(\mathbf{t}) = {z^{(j_1,0)}}^T(Z^{(j_1,0)(0,0)}(Z^{(0,0)(j_2,0)}e)) + {z^{(j_1,1)}}^T(Z^{(j_1,1)(0,0)}(Z^{(0,0)(j_2,0)}e))$
For $j_1 = j_2$ with $\alpha_S(\mathbf{t}) = 0$ follows $\mathrm{E}(\Phi_S(\mathbf{t})) = 0 \Leftrightarrow$
${\beta^{(1)}}^T(A^{(1)}(B^{(0)}e))\ \mathrm{E}(\hat{I}^2_{(j_1)}) = 0$.

For all correlations between $j_1, \ldots, j_4$ which have not been considered explicitly holds $\alpha_I(\mathbf{t}) = 0$ and $\mathrm{E}(\Phi_S(\mathbf{t})) = 0$.

Finally, we have to consider all trees $\mathbf{t} \in TS(\Delta)$ with $\rho(\mathbf{t}) = 2.5$ for which due to $\alpha_S(\mathbf{t}) = 0$ the condition $\mathrm{E}(\Phi_S(\mathbf{t})) = 0$ has to be fulfilled. Since the calculations are analogous to the ones already performed, repetition is avoided (see [12] for all trees up to order 2.5). Leaving out the trees which do not supply any new restrictions, we calculate the following conditions:

Table 5.1: Conditions from $\mathbf{t} \in TS(\Delta)$ with $\rho(\mathbf{t}) = 2.5$.

$\mathbf{t}$	correlation	condition
$(\{\tau\}_{j_1}, \tau)$		5.
$(\{\tau, \tau\}_{j_1})$		23.
$(\{[\tau]\}_{j_1})$		24.
$(\tau, \{\sigma_{j_2}, \sigma_{j_3}\}_{j_1})$	$j_1 \neq j_2 = j_3$	6.
$(\{\tau\}_{j_1}, \sigma_{j_2}, \sigma_{j_3})$	$j_2 = j_3$	5.
$(\{\tau\}_{j_1}, \{\sigma_{j_3}\}_{j_2})$	$j_2 = j_3$	5.
$(\{[\sigma_{j_2}]\}_{j_1}, \sigma_{j_3})$	$j_2 = j_3$	26.
$(\{\tau, \sigma_{j_2}, \sigma_{j_3}\}_{j_1})$	$j_1 \neq j_2 = j_3$	32.
$(\{[\sigma_{j_2}, \sigma_{j_3}]\}_{j_1})$	$j_2 = j_3$	27.
$(\{[\{\sigma_{j_3}\}_{j_2}]\}_{j_1})$	$j_2 = j_3$	36.
$(\{\{\tau\}_{j_2}, \sigma_{j_3}\}_{j_1})$	$j_1 \neq j_2 = j_3$	46.
$(\sigma_{j_1}, \sigma_{j_2}, \{\sigma_{j_4}, \sigma_{j_5}\}_{j_3})$	$j_1 = j_2, j_3 \neq j_4 = j_5$	6.
	$j_1 = j_4 > j_3,\ j_2 = j_5 > j_3$	6.
$(\{\sigma_{j_2}\}_{j_1}, \{\sigma_{j_4}, \sigma_{j_5}\}_{j_3})$	$j_1 = j_2,\ j_3 \neq j_4 = j_5$	6.
	$j_1 = j_4 > j_3,\ j_2 = j_5 > j_3$	6.
$(\{\sigma_{j_2}, \{\sigma_{j_4}\}_{j_3}\}_{j_1}, \sigma_{j_5})$	$j_1 \neq j_2 = j_3 = j_4 = j_5$	47.
	$j_1 \neq j_2 = j_3,\ j_3 \neq j_4 = j_5$	48.
	$j_1 < j_2 = j_5,\ j_1 < j_3 = j_4,\ j_2 \neq j_3$	47.
	$j_1 < j_2 = j_4,\ j_1 < j_3 = j_5,\ j_2 \neq j_3$	48.
$(\{\sigma_{j_2}, \sigma_{j_3}, \sigma_{j_4}, \sigma_{j_5}\}_{j_1})$	$j_1 \neq j_2 = j_3,\ j_1 \neq j_4 = j_5$	28.
$(\{\sigma_{j_2}, \{\sigma_{j_4}, \sigma_{j_5}\}_{j_3}\}_{j_1})$	$j_1 \neq j_2 = j_3 = j_4 = j_5$	49.

$\mathbf{t}$	correlation	condition
$(\{\sigma_{j_2},\{\{\sigma_{j_5}\}_{j_4}\}_{j_3}\}_{j_1})$	$j_1 \neq j_2 = j_3,\ j_3 \neq j_4 = j_5$	50.
	$j_1 < j_2 = j_4,\ j_1 < j_3 = j_5,\ j_4 \neq j_5$	51.
	$j_1 \neq j_2 = j_3 = j_4 = j_5$	52.
	$j_1 \neq j_2 = j_3,\ j_3 \neq j_4 = j_5$	53.
$(\{\{\sigma_{j_3}\}_{j_2},\{\sigma_{j_5}\}_{j_4}\}_{j_1})$	$j_1 < j_2 = j_4,\ j_1 < j_3 = j_5,\ j_2 \neq j_3$	54.
	$j_1 < j_2 = j_5,\ j_1 < j_3 = j_4,\ j_2 \neq j_3$	55.
	$j_1 \neq j_2 = j_3 = j_4 = j_5$	29.
	$j_1 \neq j_2 = j_4,\ j_2 \neq j_3 = j_5$	30.
	$j_1 < j_2 = j_5,\ j_1 < j_4 = j_3$	30.

Now, we just have to summarize the calculated conditions in order to arrive at the conditions in Theorem 5.1. Finally, the approximation Y by the SRK method (5.1) has uniformly bounded moments due to $\mathrm{E}({z^{(k,\nu)}}^T e) = 0$ for $1 \leq k \leq m$ and $\nu \in \{0,1\}$. ∎

We remark that in the case of SDEs with commutative noise the SRK method (5.1) can be simplified to the one introduced in [14] by choosing $\beta_i^{(2)} = 0$ for $i = 1,\ldots,s$. In this case only m random variables $\hat{I}_{(j)}$, $j = 1,\ldots,m$, have to be simulated and the number of effective stages is reduced as well. The commutativity condition can be illustrated in the light of rooted trees as follows: If the commutativity condition [5, 14] holds, then the endings $\{\sigma_{j_k}\}_{j_l}$ and $\{\sigma_{j_l}\}_{j_k}$ of a rooted tree, presented in Figure 5.1, 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 particularly for the tree $\mathbf{t}_{2.15}$ in the proof of Theorem 5.1 and with $\beta^{(2)} \equiv 0$, it can be easily checked that already those order conditions of Theorem 5.1 containing only the coefficients α, $\beta^{(1)}$, $A^{(0)}$, $B^{(0)}$, $B^{(1)}$ and $B^{(3)}$ guarantee order 2.0 for the SRK method (5.1) with $\beta^{(2)} \equiv 0$ and coincide with the ones determined in [14] for commutative noise.

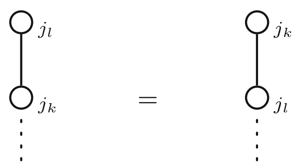

Figure 5.1: Two equivalent endings of a tree in case of commutative noise.

Some explicit SRK schemes of order 2.0 are given by RS1 and RS2 with the coefficients presented in Table 5.2 and Table 5.3. Due to $s = 4$ stages needed for the Stratonovich SRK methods, it is possible to calculate schemes of a higher deterministic order p_D than the stochastic order of convergence p_S. So the SRK method converges at least with order $p = p_S$ for SDEs, however it converges

Table 5.2: SRK scheme RS1 with order $p_D = p_S = 2.0$.

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

Table 5.3: SRK scheme RS2 with order $p_D = 3.0$ and $p_S = 2.0$.

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

with order $p_D \geq p_S$ in case of an ODE. The deterministic part is represented by the coefficients $A^{(0)}$ and α. RS2 is of order $p_D = 3.0$ and $p_S = 2.0$ while RS1 is of order $p_D = 2.0$ and $p_S = 2.0$. For $s = 4$ it is also possible to calculate coefficients for a scheme of order $p_D = 4.0$ and $p_S = 2.0$.

6 Numerical example.

We consider now some test equations in order to compare the SRK scheme (RS1) with the order one Euler–Maruyama scheme (EM), with the second order SRK scheme (PL1WM) due to Platen [5] which is also contained in the class of SRK methods proposed in [19] and with the extrapolated Euler–Maruyama scheme (ExEu) due to Talay and Tubaro [18] attaining order two. The extrapolated Euler–Maruyama approximation is given by $2\,\mathrm{E}(f(Z_T^{h/2})) - \mathrm{E}(f(Z_T^h))$ based on the Euler–Maruyama approximations $Z_T^{h/2}$ and Z_T^h calculated with step sizes h and $h/2$. Since the schemes (EM), (PL1WM) and (ExEu) are designed for Itô SDEs, we always apply them to the corresponding Itô SDEs in the following. The values $\mathrm{E}(f(X_T))$ are approximated for $f(x^1,x^2) = x^1$ and $f(x^1,x^2) = x^1 x^2$ or $f(x^1,x^2) = (x^1)^2$ by Monte Carlo simulation. Therefore, we estimate $\mathrm{E}(f(Y_T))$ by the sample average of M independent simulated realizations of the approximations $f(Y_{T,k})$, $k = 1,\dots,M$, with $Y_{T,k}$ calculated by the scheme under consideration. Then, the error is denoted by $\hat{\mu} = \mathrm{E}(f(X_T)) - \frac{1}{M}\sum_{k=1}^{M} f(Y_{T,k})$. The empirical variance $\hat{\sigma}_\mu^2$ of the error $\hat{\mu}$ is calculated following [5] based on M_1 batches with M_2 trajectories in each, i.e., $M = M_1 \cdot M_2$. In order to analyze the systematic error of the schemes under consideration, we minimize the statistical error by choosing M very large [5], i.e. much larger than usually necessary for approximations in practice. Then, the errors $\hat{\mu}$ at time $T = 1.0$ are plotted versus the corresponding step sizes h with double logarithmic scale in order to obtain the empirical order of convergence. Further, some reference lines with slope 1.0 and 2.0 are plotted for better comparison.

Firstly, we consider for $d = m = 2$ the linear SDE system with commutative noise

$$\mathrm{d}\begin{pmatrix} X_t^1 \\ X_t^2 \end{pmatrix} = \begin{pmatrix} \frac{299}{200} X_t^1 \\ \frac{299}{200} X_t^2 \end{pmatrix} \mathrm{d}t + \begin{pmatrix} \frac{1}{10} X_t^1 \\ 0 \end{pmatrix} \circ \mathrm{d}W_t^1 + \begin{pmatrix} 0 \\ \frac{1}{10} X_t^2 \end{pmatrix} \circ \mathrm{d}W_t^2, \tag{6.1}$$

with initial value $X_0 = (\frac{1}{10}, \frac{1}{10})^T$. Then, the moments of the solution are given as $\mathrm{E}(X_t^i) = \frac{1}{10}\exp(\frac{3}{2}t)$ and $\mathrm{E}((X_t^i)^2) = \frac{1}{100}\exp(\frac{301}{100}t)$ for $i = 1,2$. Here, we choose $M_1 = 20$ batches with $M_2 = 5 \times 10^6$ trajectories in each and consider the step sizes $2^0,\dots,2^{-7}$. The errors $|\hat{\mu}|$ and empirical variances $\hat{\sigma}_\mu^2$ with corresponding step sizes are presented in Figure 6.1 and Tables 6.1–6.2.

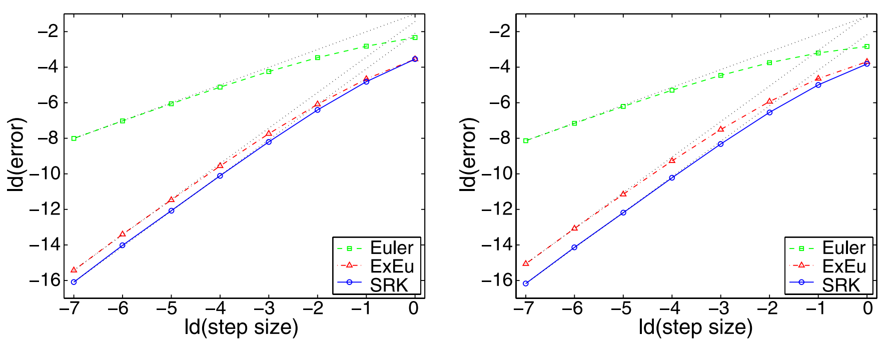

Figure 6.1: Step size vs. error for the approximation of $\mathrm{E}(X_T^1)$ in the left and of $\mathrm{E}((X_T^1)^2)$ in the right figure for SDE (6.1).

Table 6.1: Results for the approximation of $\mathrm{E}(X_t^1)$ for SDE (6.1).

h	RS1 $\lvert\hat{\mu}\rvert$	RS1 $\hat{\sigma}_\mu^2$	EM $\lvert\hat{\mu}\rvert$	EM $\hat{\sigma}_\mu^2$	ExEu $\lvert\hat{\mu}\rvert$	ExEu $\hat{\sigma}_\mu^2$	PL1WM $\lvert\hat{\mu}\rvert$	PL1WM $\hat{\sigma}_\mu^2$
2^{-0}	8.57e-02	1.20e-10	1.98e-1	1.95e-11	8.57e-2	1.19e-10	8.57e-02	1.21e-10
2^{-1}	3.56e-02	1.87e-10	1.42e-1	4.40e-11	3.95e-2	1.72e-10	3.56e-02	1.85e-10
2^{-2}	1.18e-02	3.44e-10	9.07e-2	1.33e-10	1.48e-2	3.27e-10	1.18e-02	3.43e-10
2^{-3}	3.40e-03	1.71e-10	5.27e-2	9.67e-11	4.67e-3	1.64e-10	3.40e-03	1.71e-10
2^{-4}	9.08e-04	4.42e-10	2.87e-2	3.25e-10	1.33e-3	4.35e-10	9.08e-04	4.41e-10
2^{-5}	2.32e-04	8.51e-10	1.50e-2	7.25e-10	3.52e-4	8.43e-10	2.32e-04	8.51e-10
2^{-6}	6.01e-05	6.18e-10	7.69e-3	5.71e-10	9.23e-5	6.16e-10	6.01e-05	6.18e-10
2^{-7}	1.43e-05	4.18e-10	3.89e-3	4.02e-10	2.27e-5	4.17e-10	1.43e-05	4.18e-10

Table 6.2: Results for the approximation of $\mathrm{E}((X_t^1)^2)$ for SDE (6.1).

h	RS1 $\lvert\hat{\mu}\rvert$	RS1 $\hat{\sigma}_\mu^2$	EM $\lvert\hat{\mu}\rvert$	EM $\hat{\sigma}_\mu^2$	ExEu $\lvert\hat{\mu}\rvert$	ExEu $\hat{\sigma}_\mu^2$	PL1WM $\lvert\hat{\mu}\rvert$	PL1WM $\hat{\sigma}_\mu^2$
2^{-0}	7.08e-02	6.32e-11	1.40e-1	4.84e-12	7.73e-2	5.13e-11	7.08e-02	6.32e-11
2^{-1}	3.14e-02	1.33e-10	1.09e-1	1.69e-11	4.01e-2	1.05e-10	3.14e-02	1.31e-10
2^{-2}	1.07e-02	2.74e-10	7.44e-2	6.97e-11	1.64e-2	2.40e-10	1.07e-02	2.72e-10
2^{-3}	3.12e-03	1.48e-10	4.54e-2	6.50e-11	5.50e-3	1.38e-10	3.12e-03	1.48e-10
2^{-4}	8.37e-04	3.74e-10	2.54e-2	2.40e-10	1.62e-3	3.65e-10	8.37e-04	3.74e-10
2^{-5}	2.14e-04	7.26e-10	1.35e-2	5.76e-10	4.39e-4	7.17e-10	2.14e-04	7.26e-10
2^{-6}	5.55e-05	5.06e-10	6.99e-3	4.51e-10	1.16e-4	5.04e-10	5.55e-05	5.06e-10
2^{-7}	1.36e-05	3.40e-10	3.55e-3	3.21e-10	2.92e-5	3.39e-10	1.36e-05	3.40e-10

As a second test equation, we consider for $d = m = 2$ the linear SDE system with non-commutative noise

(6.2)
$$\mathrm{d}\begin{pmatrix} X_t^1 \\ X_t^2 \end{pmatrix} = \begin{pmatrix} \frac{5}{4}X_t^2 - \frac{5}{4}X_t^1 \\ \frac{1}{4}X_t^1 - \frac{1}{4}X_t^2 \end{pmatrix} \mathrm{d}t + \begin{pmatrix} \frac{\sqrt{3}}{2}\left(X_t^1 - X_t^2\right) \\ 0 \end{pmatrix} \circ \mathrm{d}W_t^1 + \begin{pmatrix} \frac{1}{2}\left(X_t^1 + X_t^2\right) \\ X_t^1 \end{pmatrix} \circ \mathrm{d}W_t^2,$$

with initial value $X_0 = (\frac{1}{10}, \frac{1}{10})^T$. Then, we can calculate the first moment of X_t^i as $\mathrm{E}(X_t^i) = \frac{1}{10} e^{\frac{1}{2}t}$ and the second moment as well as the mixed second moment of X_t^1 and X_t^2 as $\mathrm{E}((X_t^i)^2) = \mathrm{E}(X_t^1\, X_t^2) = \frac{1}{100} e^{2t}$ for $i = 1, 2$. For the

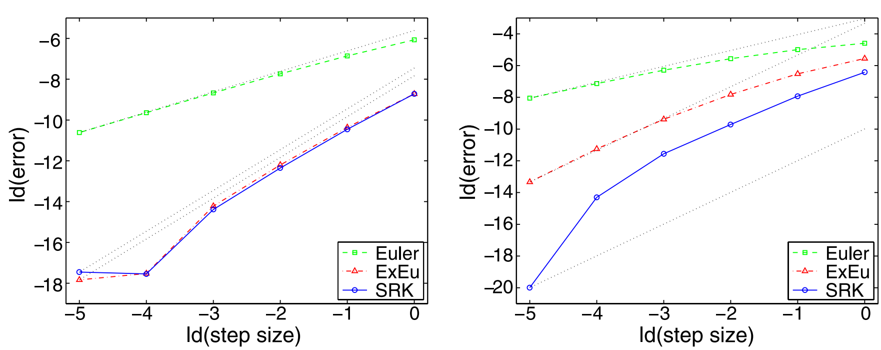

Figure 6.2: Step size vs. error for the approximation of $\mathrm{E}(X_T^1)$ in the left and of $\mathrm{E}(X_T^1\, X_T^2)$ in the right figure for SDE (6.2).

approximation, we choose $M_1 = 20$ and $M_2 = 5 \times 10^7$. The results for the step sizes $2^0, \dots, 2^{-5}$ are presented in Figure 6.2 and Tables 6.3–6.4.

Table 6.3: Results for the approximation of $\mathrm{E}(X_t^1)$ for SDE (6.2).

h	RS1 $\lvert\hat{\mu}\rvert$	RS1 $\hat{\sigma}_\mu^2$	EM $\lvert\hat{\mu}\rvert$	EM $\hat{\sigma}_\mu^2$	ExEu $\lvert\hat{\mu}\rvert$	ExEu $\hat{\sigma}_\mu^2$	PL1WM $\lvert\hat{\mu}\rvert$	PL1WM $\hat{\sigma}_\mu^2$
2^{-0}	2.37e-03	8.21e-10	1.49e-2	3.82e-10	2.37e-3	8.90e-10	2.37e-03	8.12e-10
2^{-1}	7.13e-04	6.54e-10	8.62e-3	4.57e-10	7.60e-4	4.19e-10	7.12e-04	5.48e-10
2^{-2}	1.91e-04	6.55e-10	4.69e-3	4.97e-10	2.12e-4	7.57e-10	1.91e-04	6.40e-10
2^{-3}	4.67e-05	1.08e-09	2.45e-3	8.89e-10	5.24e-5	1.10e-09	4.59e-05	1.08e-09
2^{-4}	5.25e-06	8.67e-10	1.25e-3	7.57e-10	5.32e-6	9.40e-10	5.15e-06	8.81e-10
2^{-5}	5.58e-06	1.25e-09	6.38e-4	1.12e-09	4.28e-6	1.38e-09	5.45e-06	1.24e-09

Table 6.4: Results for the approximation of $\mathrm{E}(X_t^1\, X_t^2)$ for SDE (6.2).

h	RS1 $\lvert\hat{\mu}\rvert$	RS1 $\hat{\sigma}_\mu^2$	EM $\lvert\hat{\mu}\rvert$	EM $\hat{\sigma}_\mu^2$	ExEu $\lvert\hat{\mu}\rvert$	ExEu $\hat{\sigma}_\mu^2$	PL1WM $\lvert\hat{\mu}\rvert$	PL1WM $\hat{\sigma}_\mu^2$
2^{-0}	1.17e-02	8.08e-10	4.14e-2	3.25e-11	2.13e-2	2.35e-10	2.00e-02	2.51e-10
2^{-1}	4.11e-03	4.44e-09	3.14e-2	5.53e-11	1.09e-2	3.84e-10	8.98e-03	7.43e-10
2^{-2}	1.19e-03	4.48e-09	2.11e-2	1.75e-10	4.43e-3	1.28e-09	3.14e-03	1.89e-09
2^{-3}	3.32e-04	5.21e-09	1.28e-2	1.24e-09	1.50e-3	4.01e-09	9.41e-04	4.41e-09
2^{-4}	4.95e-05	3.91e-09	7.12e-3	1.46e-09	4.10e-4	3.90e-09	2.21e-04	3.70e-09
2^{-5}	9.61e-07	7.87e-09	3.78e-3	4.33e-09	9.69e-5	7.61e-09	4.68e-05	7.70e-09

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

(6.3)
$$\mathrm{d}\begin{pmatrix} X_t^1 \\ X_t^2 \end{pmatrix} = \begin{pmatrix} -\frac{5}{4}X_t^1 + \frac{9}{4}X_t^2 \\ \frac{9}{4}X_t^1 - \frac{5}{4}X_t^2 \end{pmatrix} \mathrm{d}t + \begin{pmatrix} \sqrt{\frac{3}{4}\left(X_t^1\right)^2 - \frac{3}{2}X_t^1 X_t^2 + \frac{3}{4}\left(X_t^2\right)^2 + \frac{3}{20}} \\ 0 \end{pmatrix} \circ \mathrm{d}W_t^1$$
$$+ \begin{pmatrix} -\sqrt{\frac{1}{4}\left(X_t^1\right)^2 - \frac{1}{2}X_t^1 X_t^2 + \frac{1}{4}\left(X_t^2\right)^2 + \frac{1}{20}} \\ \sqrt{\left(X_t^1\right)^2 - 2X_t^1 X_t^2 + \left(X_t^2\right)^2 + \frac{1}{5}} \end{pmatrix} \circ \mathrm{d}W_t^2,$$

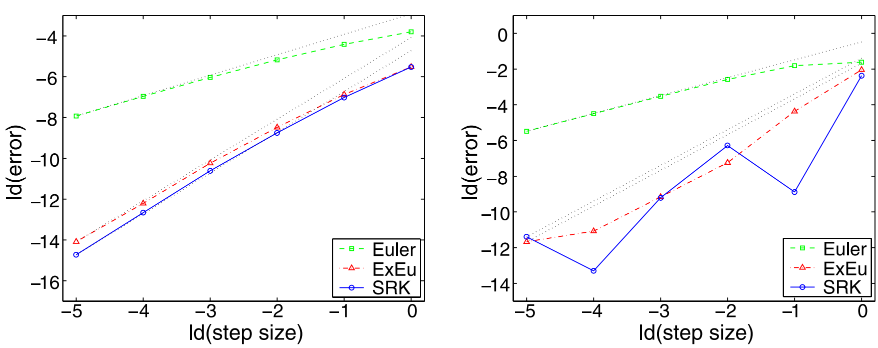

Figure 6.3: Step size vs. error for the approximation of $\mathrm{E}(X_T^1)$ in the left and of $\mathrm{E}(X_T^1\, X_T^2)$ in the right figure for SDE (6.3).

with initial value $X_0 = (\frac{1}{10}, \frac{1}{10})^T$. Then, the corresponding moments of the solution are $\mathrm{E}(X_t^i) = \frac{1}{10}\exp(t)$, $\mathrm{E}((X_t^i)^2) = \frac{3}{50}\exp(2t) - \frac{1}{10}\exp(-t) + \frac{1}{20}$ and $\mathrm{E}(X_t^1\, X_t^2) = \frac{3}{50}\exp(2t) + \frac{1}{5}\exp(-t) - \frac{1}{4}$ for $i = 1, 2$. Here, we choose $M_1 = 20$ and $M_2 = 5 \times 10^7$. The results are presented in Figure 6.3 and Tables 6.5–6.6.

Table 6.5: Results for the approximation of $\mathrm{E}(X_t^1)$ for SDE (6.3).

	RS1		EM		ExEu		PL1WM	
h	$\lvert\hat{\mu}\rvert$	$\hat{\sigma}_\mu^2$	$\lvert\hat{\mu}\rvert$	$\hat{\sigma}_\mu^2$	$\lvert\hat{\mu}\rvert$	$\hat{\sigma}_\mu^2$	$\lvert\hat{\mu}\rvert$	$\hat{\sigma}_\mu^2$
2^{-0}	2.18e-02	1.58e-08	7.18e-2	2.98e-09	2.18e-2	1.31e-08	2.18e-02	4.57e-09
2^{-1}	7.72e-03	1.32e-08	4.68e-2	6.80e-09	8.53e-3	1.33e-08	7.76e-03	1.05e-08
2^{-2}	2.32e-03	1.50e-08	2.77e-2	1.07e-08	2.80e-3	1.64e-08	2.33e-03	1.50e-08
2^{-3}	6.40e-04	5.91e-09	1.53e-2	5.71e-09	8.28e-4	6.13e-09	6.38e-04	7.05e-09
2^{-4}	1.54e-04	7.43e-09	8.01e-3	6.95e-09	2.11e-4	7.79e-09	1.49e-04	7.53e-09
2^{-5}	3.70e-05	4.19e-09	4.12e-3	4.79e-09	5.77e-5	5.29e-09	3.16e-05	4.50e-09

Table 6.6: Results for the approximation of $\mathrm{E}(X_t^1\, X_t^2)$ for SDE (6.3).

	RS1		EM		ExEu		PL1WM	
h	$\lvert\hat{\mu}\rvert$	$\hat{\sigma}_\mu^2$	$\lvert\hat{\mu}\rvert$	$\hat{\sigma}_\mu^2$	$\lvert\hat{\mu}\rvert$	$\hat{\sigma}_\mu^2$	$\lvert\hat{\mu}\rvert$	$\hat{\sigma}_\mu^2$
2^{-0}	1.94e-01	4.81e-07	3.27e-1	8.15e-10	2.43e-1	6.46e-09	3.92e-02	2.83e-09
2^{-1}	2.12e-03	1.09e-06	2.85e-1	4.33e-09	4.85e-2	1.45e-08	2.73e-02	5.34e-09
2^{-2}	1.28e-02	9.48e-07	1.67e-1	6.23e-09	6.60e-3	2.84e-08	1.12e-02	9.88e-09
2^{-3}	1.70e-03	1.46e-07	8.67e-2	2.19e-08	1.75e-3	1.05e-07	3.36e-03	8.05e-08
2^{-4}	9.91e-05	2.67e-07	4.42e-2	3.46e-08	4.63e-4	1.85e-07	8.18e-04	1.86e-07
2^{-5}	3.71e-04	2.83e-07	2.24e-2	1.31e-07	3.07e-4	3.08e-07	4.05e-04	3.07e-07

7 Conclusion.

In the present paper, a class of SRK methods applicable to non–commutative Stratonovich SDE systems with a m–dimensional driving Wiener process is introduced. Compared to other SRK schemes, the main advantage of the introduced class of SRK methods (5.1) is the significant reduction of the computational costs. If we consider the order two SRK schemes proposed in [5, p. 486] and in [19] for Itô SDE systems, then at least 2 evaluations of the drift function a and $2m + 1$ evaluations of each diffusion function b^j, $j = 1, \ldots, m$, are necessary for each step. Further, $m(m + 1)/2$ independent random variables have to be simulated for the schemes in [5, 19] for each step. Furthermore, these SRK methods are not directly applicable to Stratonovich SDEs. On the other hand, Komori [7] proposed a SRK family for non–commutative Stratonovich SDEs. However, the computational effort of his SRK family depends also linearly on the dimension of the driving Wiener process in each step. Although the SRK schemes proposed in [7] need only the simulation of $2m - 1$ independent random variables for each step, these SRK schemes need 4 evaluations of the drift a and $3m+1$ evaluations of each diffusion function b^j, $j = 1, \ldots, m$, in each step. Thus, all known SRK methods seem not to have much relevance in practice, especially for high dimensional problems, due to the dependence of the computational costs on the dimension m of the driving Wiener process.

In contrast to this, the new SRK scheme RS1 needs 2 evaluations of the drift a and only 6 evaluations of the diffusion functions b^j, $j = 1, \ldots, m$, for each step due to $s = 4$, $H_1^{(k)} = \hat{H}_1^{(k)}$ and $\beta_4^{(2)} = 0$. Further, also only $2m - 1$ independent random variables need to be simulated for each step. As a result of this, the introduced SRK methods are appropriate also for high dimensional problems with large values for d and m. Further, the computational effort of the new SRK schemes is comparable to that of the extrapolated Euler–Maruyama scheme for Itô SDEs where 3 evaluations of the drift a and 3 evaluations of the diffusion functions b^j, $j = 1, \ldots, m$, are necessary, and which needs the simulation of $2m$ independent random variables at each step. However, if we transform a Stratonovich SDE to an Itô SDE with the same solution process, then the drift function contains also the diffusion functions b^j and their first derivatives [5]. Therefore, the extrapolated Euler–Maruyama scheme applied to a Stratonovich SDE needs finally 3 additional evaluations of the first derivatives of b^j, $j = 1, \ldots, m$, for each step. As a result of this, the introduced new class of SRK methods is of considerable importance, especially for high dimensional problems like e.g. in mathematical finance or physics. Further, embedded SRK schemes may be applied for step size control, see also [10]. For future research, the calculation of coefficients for implicit SRK methods and an analysis of the stability attributes for the SRK methods may be of particular interest.

REFERENCES

1. K. Burrage and P. M. Burrage, *High strong order explicit Runge–Kutta methods for stochastic ordinary differential equations*, Appl. Numer. Math., 22(1–3) (1996), pp. 81–101.
2. K. Burrage and P. M. Burrage, *Order conditions of stochastic Runge–Kutta methods by B-series*, SIAM J. Numer. Anal., 38(5) (2000), pp. 1626–1646.
3. E. Hairer, S. P. Nørsett, and G. Wanner, *Solving Ordinary Differential Equations I*, Springer, 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, Berlin, 1999.
6. P. E. Kloeden, E. Platen, and N. Hofmann, *Extrapolation methods for the weak approximation of Itô diffusions*, SIAM J. Numer. Anal., 32(5) (1995), pp. 1519–1534.
7. Y. Komori, *Weak second-order stochastic Runge–Kutta methods for non-commutative stochastic differential equations*, J. Comput. Appl. Math., in press, doi:10.1016/j.cam.2006.06.006.
8. Y. Komori, T. Mitsui, and H. Sugiura, *Rooted tree analysis of the order conditions of ROW-type scheme for stochastic differential equations*, BIT, 37(1) (1997), pp. 43–66.
9. G. N. Milstein and M. V. Tretyakov, *Stochastic Numerics for Mathematical Physics*, Springer, Berlin, 2004.
10. K.-S. Moon, A. Szepessy, R. Tempone, and G. E. Zouraris, *Convergence rates for adaptive weak approximation of stochastic differential equations*, Stochastic Anal. Appl., 23 (2005), pp. 511–558.
11. N. J. Newton, *Asymptotically efficient Runge–Kutta methods for a class of Itô and Stratonovich equations*, SIAM J. Appl. Math., 51(2) (1991), pp. 542–567.

12. A. Rößler, *Rooted tree analysis for order conditions of stochastic Runge–Kutta methods for the weak approximation of stochastic differential equations*, Stochastic Anal. Appl., 24(1) (2006), pp. 97–134.

13. A. Rößler, *Runge–Kutta methods for Itô stochastic differential equations with scalar noise*, BIT, 46(1) (2006), pp. 97–110.

14. A. Rößler, *Runge–Kutta methods for Stratonovich stochastic differential equation systems with commutative noise*, J. Comput. Appl. Math., 164–165 (2004), pp. 613–627.

15. A. Rößler, *Stochastic Taylor Expansions for the Expectation of Functionals of Diffusion Processes*, Stochastic Anal. Appl., 22(6) (2004), pp. 1553–1576.

16. W. Rümelin, *Numerical treatment of stochastic differential equations*, SIAM J. Numer. Anal., 19(3) (1982), pp. 604–613.

17. D. Talay, *Second-order discretization schemes of stochastic differential systems for the computation of the invariant law*, Stochastics Stochastics Rep., 29(1) (1990), pp. 13–36.

18. D. Talay and L. Tubaro, *Expansion of the global error for numerical schemes solving stochastic differential equations*, Stochastic Anal. Appl., 8(4) (1990), pp. 483–509.

19. A. Tocino and J. Vigo-Aguiar, *Weak second order conditions for stochastic Runge–Kutta methods*, SIAM J. Sci. Comput., 24(2) (2002), pp. 507–523.