

B-series for SDEs with application to exponential integrators for non-autonomous semi-linear problems

Alemayehu Adugna Arara^a, Kristian Debrabant^b, Anne Kværnø^c

^a*Hawassa University, Department of Mathematics, P.O. Box 05, Hawassa, Ethiopia*

^b*University of Southern Denmark, Department of Mathematics and Computer Science, 5230 Odense M, Denmark*

^c*Norwegian University of Science and Technology - NTNU, Department of Mathematical Sciences, NO-7491 Trondheim, Norway*

Abstract

In this paper a set of previous general results for the development of B-series for a broad class of stochastic differential equations has been collected. The applicability of these results is demonstrated by the derivation of B-series for non-autonomous semi-linear SDEs and exponential Runge-Kutta methods applied to this class of SDEs, which is a significant generalization of existing theory on such methods.

Keywords: Stochastic differential equations, stochastic B-series, exponential integrators

Dedicated to Brynjulf Owren and Hans Munthe-Kaas on the occasion of their 60th birthdays

1. Introduction

The idea of B-series, as described here, is simply about describing the solution of a differential equation (or approximating numerical scheme) as a formal series in time (or in the step size), allowing then to represent the local error, i.e. the difference between the exact solution and the numerical approximation after one time step, as a B-series as well and thereby easily analyse the order of convergence of the numerical scheme. These series were initially derived for ordinary differential equations (ODEs) by Butcher [8], and the idea was further extended by Hairer and Wanner in [15], see [14] and references therein. Owren and Munthe-Kaas have both given several contributions to the development of B-series, in particular in the context of geometrical integration, see [22, 4, 26, 9, 24, 23]. B-series for stochastic differential equations (SDEs) have been developed in different contexts by several authors, [5, 7, 6, 19, 18, 27, 28]. A rather general framework for developing B-series for SDEs was developed and extended in [10, 12, 1]. This approach is independent on whether the SDE is Itô or Stratonovich and whether weak or strong convergence is considered.

Email addresses: alemayehuadugna@hu.edu.et (Alemayehu Adugna Arara),
debrabant@imada.sdu.dk (Kristian Debrabant), anne.kvarno@ntnu.no (Anne Kværnø)

The aim of this paper is to collect the main tools from this approach, which is done in section Section 2, and then apply the theory to non-autonomous semi-linear SDEs and a class of exponential Runge–Kutta (ERK) methods. A similar work has very recently been published by Yang et.al. [31], however, their results only apply to Stratonovich SDEs with a commutative linear part. The results which will be presented in the current paper do not have these limitations.

The outline is as follows: The main ideas and results on stochastic B–series for autonomous SDEs are presented in Section 2. In Section 3 these are used for a generalisation to non-autonomous SDEs, and non-autonomous semi-linear SDEs are discussed in Section 4.

2. Stochastic B–series: Definitions and main results

In this section the basic definitions and results to be used in the later sections are summarised.

Consider a system of stochastic partitioned differential equations with Q partitions and M diffusion terms that potentially are split,

$$X^{(q)}(t) = x_0^{(q)} + \sum_{m=0}^M \int_0^t \left(\sum_{r=1}^{R_{m,q}} g_m^{(q,r)}(X^{(1)}(s), X^{(2)}(s), \dots, X^{(Q)}(s)) \right) \star dW_m(s) \quad (1)$$

for $q = 1, \dots, Q$, for which we will also use the abbreviated form

$$dX^{(q)}(t) = \sum_{m=0}^M \left(\sum_{r=1}^{R_{m,q}} g_m^{(q,r)}(X^{(1)}(t), X^{(2)}(t), \dots, X^{(Q)}(t)) \right) \star dW_m(t), \quad X^{(q)}(0) = x_0^{(q)}. \quad (2)$$

To simplify the notation the deterministic terms are represented by $m = 0$, such that $dW_0(s) = ds$, while W_m , $m = 1, \dots, M$, denote one-dimensional and pairwise independent Wiener processes. The integrals w.r.t. the Wiener processes are interpreted as either Itô integrals, $\star dW_m(s) = dW_m(s)$, or Stratonovich integrals, $\star dW_m(s) = \circ dW_m(s)$. We also define the vector of initial values, $x_0 = [x_0^{(1)\top}, x_0^{(2)\top}, \dots, x_0^{(Q)\top}]^\top$. Furthermore, we assume that the coefficients $g_m^{(q,r)} : \mathbb{R}^{d_1} \times \dots \times \mathbb{R}^{d_Q} \rightarrow \mathbb{R}^{d_q}$ are sufficiently smooth, and that the conditions of the existence and uniqueness theorem [25] are satisfied. The systems are considered to be autonomous. If $R_{m,q} = 1$ for $m = 0, \dots, M$ and $q = 1, \dots, Q$ then the splitting is horizontal, if $Q = 1$ the splitting is vertical, see [3]. In those cases, the corresponding indices may be omitted.

Example 1. Consider the equation system

$$\begin{aligned} \dot{r} &= v, \\ \dot{v} &= f_d(r, t) - \alpha(t)v + f_s(r, t)\beta(t) \end{aligned}$$

which describes the evolution of a particle with unit mass, coordinate $r(t)$ and velocity $v(t)$. The particle is affected by three forces: an external force $f_d(t, r)$, a time-dependent

friction force $\alpha(t) \geq 0$ and a thermal white noise term with a time and space dependent intensity. This can be rewritten as a proper SDE of the form (2),

$$d \underbrace{\begin{pmatrix} R(t) \\ V(t) \end{pmatrix}}_{X^{(1)}(t)} = \underbrace{\begin{pmatrix} 0 \\ f_d(R(t), t) \end{pmatrix}}_{g_0^{(1,1)}} dt + \underbrace{\begin{pmatrix} V(t) \\ -\alpha(t)V(t) \end{pmatrix}}_{g_0^{(1,2)}} dt + \underbrace{\begin{pmatrix} 0 \\ f_s(R(t), V(t), t) \end{pmatrix}}_{g_1^{(1,1)}} dW_1(t) \quad (3a)$$

$$d \underbrace{t}_{X^{(2)}(t)} = \underbrace{1}_{g_0^{(2,1)}} dt. \quad (3b)$$

This is of course only one of many possible splittings of this problem, but it is the one which will be used for demonstration purposes throughout this paper. Other splittings have been discussed in [1], Example 1.1.

Partitioned systems appear naturally in many situations, e.g. for Hamiltonian problems. Note that (2) also covers the case of matrix-valued SDEs, in which case $X^{(q)}$ would correspond to column q of the solution.

In this paper, horizontal splitting will be used to derive B-series for non-autonomous problems, vertical splitting will be used for semi-linear problems.

The main idea is to express the exact solutions $X^{(q)}(h)$ as B-series $B^{(q)}(\phi, x_0; h)$:

$$B^{(q)}(\phi, x_0; h) = \sum_{\tau \in T_q} \alpha(\tau) \cdot \phi(\tau)(h) \cdot F(\tau)(x_0),$$

where T_q is the set of shaped, colored, rooted trees as defined below. The terms $\alpha(\tau)$ are combinatoric terms. The elementary weight functions $\phi(\tau)(h)$ are stochastic integrals or random variables, and $F(\tau)(x_0)$ are the elementary differentials. To simplify the presentation, we assume that all elementary differentials exist and all considered B-series converge. Otherwise, one has to consider truncated B-series and discuss the remainder term [27].

The following theory on partitioned SDEs is taken from [1], where also all proofs can be found.

Definition 1 (Trees and combinatorial coefficients). *The set of shaped, rooted trees*

$$T = T_1 \cup T_2 \cup \dots \cup T_Q$$

where

$$T_{q,m} = T_{q,0,m} \cup T_{q,1,m} \cup \dots \cup T_{q,R_{m,q},m} \quad \text{and} \quad T_q = \{\emptyset_q\} \cup T_{q,0} \cup T_{q,1} \cup \dots \cup T_{q,M}$$

for $q = 1, \dots, Q$ and $m = 0, \dots, M$ is recursively defined as follows for $r = 1, \dots, R_{m,q}$:

1. The graph $\bullet_{q,r,m}$ with only one vertex of shape (q, r) and color m belongs to $T_{q,r,m}$.
2. If $\tau_1, \tau_2, \dots, \tau_\kappa \in T \setminus \{\emptyset_1, \dots, \emptyset_Q\}$, then also $[\tau_1, \tau_2, \dots, \tau_\kappa]_{q,r,m} \in T_{q,r,m}$, where $[\tau_1, \tau_2, \dots, \tau_\kappa]_{q,r,m}$ denotes the tree formed by joining the subtrees $\tau_1, \tau_2, \dots, \tau_\kappa$ each by a single branch to a common root of shape (q, r) and color m .

Further, we define $\alpha(\tau)$ as

$$\alpha(\emptyset_q) = 1, \quad \alpha(\bullet_{q,r,m}) = 1, \quad \alpha(\tau = [\tau_1, \dots, \tau_\kappa]_{q,r,m}) = \frac{1}{v_1! v_2! \dots v_\kappa!} \prod_{k=1}^{\kappa} \alpha(\tau_k),$$

where $v_1, v_2, \dots, v_\kappa$ count equal trees among $\tau_1, \tau_2, \dots, \tau_\kappa$.

In addition to the bracket notation used in Definition 1, rooted trees can be illustrated as graphs, see Example 2, where vertices $\bullet_{q,r,0}$ with color $m = 0$ are represented as black nodes (so-called deterministic nodes) with indices q, r , while vertices $\bullet_{q,r,m}$ with color $m > 0$ are represented by white nodes (so-called stochastic nodes) with indices q, r, m .

Definition 2 (Elementary differentials). *For a tree $\tau \in T$ the elementary differential is a mapping $F(\tau): \mathbb{R}^{d_1} \times \dots \times \mathbb{R}^{d_Q} \rightarrow \mathbb{R}^d$ defined recursively by*

1. $F(\emptyset_q)(x_0) = x_0^{(q)}$,
2. $F(\bullet_{q,r,m})(x_0) = g_m^{(q,r)}(x_0)$,
3. if $\tau = [\tau_1, \tau_2, \dots, \tau_\kappa]_{q,r,m} \in T_{q,r,m}$, then

$$F(\tau)(x_0) = (D_\tau g_m^{(q,r)})(x_0)(F(\tau_1)(x_0), F(\tau_2)(x_0), \dots, F(\tau_\kappa)(x_0))$$

where $D_\tau = \frac{\partial^\kappa}{\partial x^{(q_1)} \dots \partial x^{(q_\kappa)}}$ for (q_k, r_k) being the shape of τ_k , $k = 1, \dots, \kappa$.

Fundamental for this work is the following lemma which says that if some functions $V^{(q)}(h)$ can be written as a B-series, then $f(V^{(1)}(h), \dots, V^{(Q)}(h))$ can also be written as a B-series. This is a trivial extension of the lemma found in [10].

Lemma 1. *If $V^{(q)}(h) = B^{(q)}(\phi, x_0; h)$, $q = 1, \dots, Q$, are some B-series with $\phi(\emptyset_q) \equiv 1$ and $f \in C^\infty(\mathbb{R}^{d_1} \times \dots \times \mathbb{R}^{d_Q}, \mathbb{R}^d)$, then $f(V^{(1)}(h), \dots, V^{(Q)}(h))$ can be written as a formal series of the form*

$$f(V^{(1)}(h), \dots, V^{(Q)}(h)) = \sum_{u \in U_f} \beta(u) \cdot \psi_\phi(u)(h) \cdot G(u)(x_0) \quad (4)$$

where

1. U_f is a set of trees derived from T , by
 - (a) $\bullet_f \in U_f$,
 - (b) if $\tau_1, \tau_2, \dots, \tau_\kappa \in T \setminus \{\emptyset_1, \dots, \emptyset_Q\}$, then $[\tau_1, \tau_2, \dots, \tau_\kappa]_f \in U_f$.
2. $G(\bullet_f)(x_0) = f(x_0)$ and

$$G(\tau = [\tau_1, \tau_2, \dots, \tau_\kappa]_f)(x_0) = (D_\tau f)(x_0)(F(\tau_1)(x_0), \dots, F(\tau_\kappa)(x_0)),$$

where D_τ is defined in analogy to Definition 2.

3. $\beta(\bullet_f) = 1$ and

$$\beta([\tau_1, \tau_2, \dots, \tau_\kappa]_f) = \frac{1}{v_1! v_2! \dots v_\kappa!} \prod_{k=1}^{\kappa} \alpha(\tau_k),$$

where $v_1, v_2, \dots, v_\kappa$ count equal trees among $\tau_1, \tau_2, \dots, \tau_\kappa$.

4. $\psi_\phi(\bullet_f)(h) = 1$ and $\psi_\phi([\tau_1, \tau_2, \dots, \tau_\kappa]_f)(h) = \prod_{k=1}^{\kappa} \phi(\tau_k)(h)$.

By the use of this lemma, the following result was proved in [1].

Theorem 1. *The exact solutions $X^{(q)}(h)$ of (1), $q = 1, \dots, Q$, can be written as B-series $B^{(q)}(\varphi, x_0; h)$ with*

$$\begin{aligned}\varphi(\emptyset_q)(h) &= 1, & \varphi(\bullet_{q,r,m})(h) &= \int_0^h dW_m(s) = W_m(h), \\ \varphi([\tau_1, \tau_2, \dots, \tau_\kappa]_{q,r,m})(h) &= \int_0^h \prod_{j=1}^{\kappa} \varphi(\tau_j(s)) \star dW_m(s),\end{aligned}$$

for all $[\tau_1, \tau_2, \dots, \tau_\kappa]_{q,r,m} \in T_{q,m}$, $r = 1, \dots, R_{m,q}$, $q = 1, \dots, Q$, $m = 0, 1, \dots, M$.

Theorem 1 represents the solution of (1) as B-series. By also deriving a B-series representation of the numerical approximation, one can represent the local error, i.e. their difference, as well as B-series. Comparing the weight functions in the B-series representation of the exact solution and numerical approximation allows thus to analyze the (strong or mean-square) consistency of a numerical method. For weak approximations, where function evaluations of the exact solution and numerical approximation need to be compared, additionally Lemma 1 is used.

To characterize the contribution of (the weight function of) a tree to the order of the error, the following definition is needed.

Definition 3 (Tree order). *The order ρ of a tree $\tau \in T$ is defined by*

$$\begin{aligned}\rho(\emptyset_q) &= 1, & \rho(\bullet_{q,r,m}) &= \begin{cases} 1 & \text{for } m = 0, \\ \frac{1}{2} & \text{otherwise,} \end{cases} \\ \rho([\tau_1, \tau_2, \dots, \tau_\kappa]_{q,r,m}) &= \sum_{k=1}^{\kappa} \rho(\tau_k) + \begin{cases} 1 & \text{for } m = 0, \\ \frac{1}{2} & \text{otherwise.} \end{cases}\end{aligned}$$

Example 2. *Let*

$$\tau = \begin{array}{c} \bullet_{21} \quad \bullet_{21} \\ \quad \diagdown \quad \diagup \\ \bullet_{12} \quad \bullet_{11} \\ \quad \diagdown \quad \diagup \\ \bullet_{111} \quad \bullet_{21} \\ \quad \diagdown \quad \diagup \\ \bullet_{12} \end{array} = [[[\bullet_{2,1,0}^2]_{1,2,0}, \bullet_{1,1,0}]_{1,1,1}, \bullet_{2,1,0}]_{1,2,0},$$

in which vertices $\bullet_{q,r,0}$ with color $m = 0$ are represented as black nodes with indices q, r , while vertices $\bullet_{q,r,1}$ with color $m = 1$ are represented by white nodes with indices $q, r, 1$. The corresponding terms are

$$\alpha(\tau) = 1/2,$$

$$\rho(\tau) = 6 \frac{1}{2},$$

$$F(\tau) = D_{1,2} g_0^{(1,2)} \left(D_{1,1} g_1^{(1,1)} \left(D_{2,2} g_0^{(1,2)} (g_0^{(2,1)}, g_0^{(2,1)}), g_0^{(1,1)} \right), g_0^{(2,1)} \right),$$

$$\varphi(\tau) = \int_0^h \int_0^{s_1} \int_0^{s_2} s_3^2 ds_3 s_2 \star dW_1(s_2) s_1 ds_1 = \frac{1}{3} \int_0^h \int_0^{s_1} s_2^4 \star dW_1(s_2) s_1 ds_1$$

where $D_{ij} = \frac{\partial^2}{\partial x^{(i)} \partial x^{(j)}}$. How the elementary differentials would look like for the SDE introduced in Example 1 is presented in Appendix A (the weight function $\varphi(\tau)$ and coefficient α only depend on the tree, not on the concrete SDE).

Stochastic B-series evaluated in a B-series can (similar to the deterministic case) be written as a B-series. This result is useful for the construction of implicit Taylor methods (see e.g. [17, 29, 21, 11, 12]), splitting methods and composition methods. Note that for implicit methods, it might be necessary to use truncated random variables to ensure the solvability of the implicit equations, see [21].

Let $ST(\tau)$ be the set of all possible subtrees of τ together with the corresponding remainder multiset ω , that is, for each $\tau \in T \setminus \{\emptyset_1, \dots, \emptyset_Q\}$ we have

$$\begin{aligned} ST(\bullet_{q,r,m}) &= \{(\emptyset_q, \{\bullet_{q,r,m}\}), (\bullet_{q,r,m}, \{\emptyset_q\})\}, \\ ST(\tau = [\tau_1, \dots, \tau_\kappa]_{q,r,m}) &= \{(\vartheta, \omega) : \vartheta = [\vartheta_1, \dots, \vartheta_\kappa]_{q,r,m}, \quad \omega = \{\omega_1, \dots, \omega_\kappa\}, \\ &\quad (\vartheta_i, \omega_i) \in ST(\tau_i), \quad i = 1, \dots, \kappa\} \cup (\emptyset, \{\tau\}). \end{aligned}$$

We also have to take care of possible equal terms in the formula presented below. This is done as follows: For a given triple $(\tau, \vartheta, \omega)$ write first $\vartheta = [\vartheta_1, \dots, \vartheta_{\kappa_\vartheta}]_{q,r,m} = [\bar{\vartheta}_1^{v_1}, \dots, \bar{\vartheta}_K^{v_K}]_{q,r,m}$, where the latter only expresses that ϑ is composed by K different nonempty trees, each appearing v_i times, hence $\sum_{i=1}^K v_i = \kappa_\vartheta$. Let $\tau = [\tau_1, \dots, \tau_\kappa]_{q,r,m}$. For $i = 1, \dots, K$, each $\bar{\vartheta}_i$ is a subtree of some of the τ_j 's, with corresponding remainder multisets ω_j . Assume that there are exactly p_i different such triples $(\bar{\tau}_{ik}, \bar{\vartheta}_i, \bar{\omega}_{ik})$ each appearing exactly v_{ik} times so that $\sum_{k=1}^{p_i} v_{ik} = v_i$. Finally, let $\bar{\delta}_k \in \omega$ be the distinct trees with multiplicity s_k , $k = 1, \dots, p_0$, of the remainder multiset which are directly connected to the root of τ . Then, τ can be written as

$$\tau = [\bar{\delta}_1^{s_1}, \dots, \bar{\delta}_{p_0}^{s_{p_0}}, \bar{\tau}_{11}^{v_{11}}, \dots, \bar{\tau}_{1p_1}^{v_{1p_1}}, \dots, \bar{\tau}_{K1}^{v_{K1}}, \dots, \bar{\tau}_{Kp_K}^{v_{Kp_K}}]_{q,r,m} = [\bar{\tau}_1^{R_1}, \dots, \bar{\tau}_V^{R_V}]_{q,r,m}, \quad (5)$$

where the rightmost expression above indicates that τ is composed by V different trees $\bar{\tau}_1, \dots, \bar{\tau}_V$, with $\bar{\tau}_i$ appearing R_i times, $i = 1, \dots, V$.

With these definitions, we can state the following theorem, whose proof is similar to the one for the nonpartitioned case given in [12], see also [11]:

Theorem 2 (Composition of B-series). *Let $\phi_x(\emptyset_q) \equiv 1$ for $q = 1, \dots, Q$. Then for $q = 1, \dots, Q$, the B-series $B^{(q)}(\phi_y, \cdot; h)$ evaluated at*

$$B(\phi_x, x_0; h) := [B^{(1)}(\phi_x, x_0; h)^\top, \dots, B^{(Q)}(\phi_x, x_0; h)^\top]^\top$$

is again a B-series,

$$B^{(q)}(\phi_y, B(\phi_x, x_0; h); h) = B^{(q)}(\phi_x \circ \phi_y, x_0; h),$$

where

$$(\phi_x \circ \phi_y)(\tau) = \sum_{(\vartheta, \omega) \in ST(\tau)} \gamma(\tau, \vartheta, \omega) \left(\phi_y(\vartheta) \prod_{\delta \in \omega} \phi_x(\delta) \right)$$

with $\gamma(\emptyset_q, \emptyset_q, \{\emptyset_q\}) = 1$ and

$$\gamma(\tau, \vartheta, \omega) = \frac{R_1! \cdots R_V!}{s_1! \cdots s_{p_0}! v_{11}! \cdots v_{Kp_K}!} \prod_{i=1}^{\kappa_\vartheta} \gamma(\tau_i, \vartheta_i, \omega_i).$$

for $\tau \neq \emptyset_q$.

Finally, the following lemma, proved in [11], is the key to derive B-series for exponential integrators for SDEs in Section 4.

Lemma 2. *If $\phi_x(\emptyset_q) \equiv 0$ for $q = 1, \dots, Q$, then we have for $q = 1, \dots, Q$*

$$\partial_2 B^{(q)}(\phi_y, x_0; h) B(\phi_x, x_0; h) = B^{(q)}(\phi_x * \phi_y, x_0; h),$$

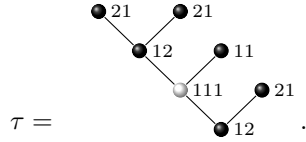
where the bi-linear operator $*$ is given by

$$(\phi_x * \phi_y)(\tau) = \begin{cases} 0 & \text{if } \tau = \emptyset, \\ \sum_{(\vartheta, \{\delta\}) \in SP(\tau)} \gamma(\tau, \vartheta, \{\delta\}) \cdot \phi_y(\vartheta) \phi_x(\delta) & \text{otherwise,} \end{cases} \quad (6)$$

with

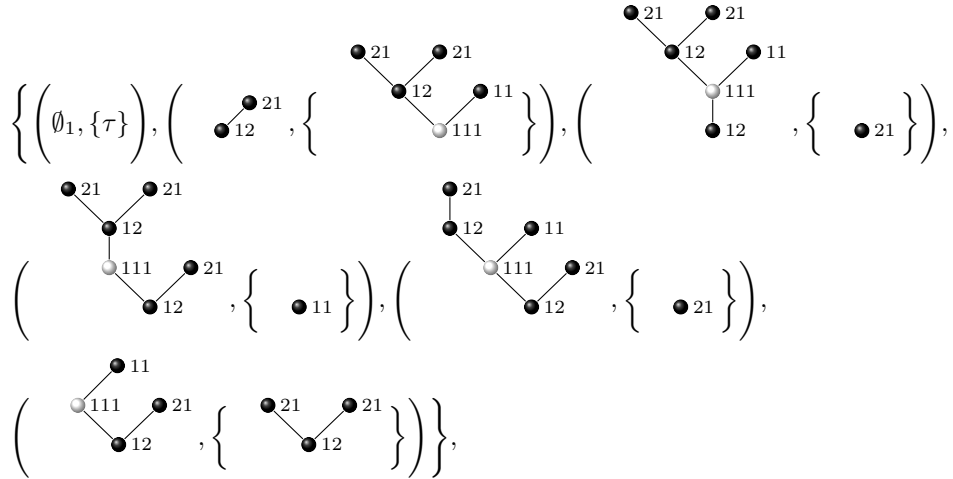
$$SP(\tau) = \{(\vartheta, \omega) \in ST(\tau) : \omega \text{ contains exactly one element } \delta\}.$$

Example 3. *Consider the tree τ from Example 2,*



Then

$$SP(\tau) =$$



which will be further used in Example 6.

For general SDEs, the number of terms in the B-series is rather overwhelming, so whenever applied to a particular class of SDEs, the main issue is to identify all trees for which the elementary differentials automatically become zero, and remove those from the definition of trees for this class. This idea was demonstrated for several problems in [1] and will again be applied here.

3. Non-autonomous SDEs

B-series are usually expressed for autonomous versions of the equations, with the assumption, which also will be used here, that a non-autonomous system easily can be converted to an autonomous one. This is of particular interest in the rather common case of SDEs with additive but time-dependent noise, for example:

$$dX(t) = g_0(X(t))dt + \sum_{m=1}^M g_m(t) \star dW_m(t). \quad (7)$$

In this section, we will derive B-series for SDEs for which also some or all of the Wiener processes are allowed to enter explicitly into the coefficient functions,

$$dX(t) = \sum_{m=0}^M \sum_{r=1}^{R_m} g_m^{(r)}(X(t), W(t)) \star dW_m(t), \quad X(0) = x_0, \quad (8)$$

where $W(t) = [W_0(t), W_1(t), \dots, W_l(t)]^\top$ with $l \leq m$ and we allow for vertical splitting.

Example 4 (Lawson - or the integrating factor method). *Consider the semi-linear equation*

$$dX(t) = \sum_{m=0}^M (A_m X(t) + g_m(X(t))) \star dW_m(t) \quad (9)$$

with constant, commutative linear terms, that is

$$[A_{m_1}, A_{m_2}] = A_{m_1} A_{m_2} - A_{m_2} A_{m_1} = 0, \quad m_1, m_2 = 1, \dots, M.$$

Using the Lawson transformation

$$V(t) = e^{-L(t)} X(t), \quad L(t) = (A_0 - \gamma^* \sum_{m=1}^M A_m^2) t + \sum_{m=1}^M A_m W_m(t)$$

with $\gamma^* = 1/2$ in the Itô case and $\gamma^* = 0$ in the Stratonovich case, the transformed system becomes

$$dV(t) = \sum_{m=0}^M \hat{g}_m(V(t), W(t)) \star dW_m(t), \quad \hat{g}_m(x, W(t)) = e^{-L(t)} g_m(e^{L(t)} x)$$

which is a non-autonomous SDE of the form (8). See [13] and references therein.

Following the well known approach from ODEs, let $\tilde{X}^{(m+2)}(t) = W_m(t)$, $m = 0, \dots, l$, such that the SDE can be written as a horizontal (and vertical) split system:

$$\begin{aligned} d\tilde{X}^{(1)}(t) &= \sum_{m=0}^M \sum_{r=1}^{R_m} g_m^{(1,r)}(\tilde{X}(t)) \star dW_m(t), & \tilde{X}^{(1)}(0) &= x_0, \\ d\tilde{X}^{(2)}(t) &= g_0^{(2,1)}(\tilde{X}(t)) \star dW_0(t) = 1 \cdot dW_0(t), & \tilde{X}^{(2)}(0) &= t_0, \\ &\vdots & &\vdots \\ d\tilde{X}^{(l+2)}(t) &= g_l^{(l+2,1)}(\tilde{X}(t)) \star dW_l(t) = 1 \cdot dW_l(t), & \tilde{X}^{(l+2)}(0) &= W_l(t_0), \end{aligned}$$

where $\tilde{X}(t) = [X(t + t_0)^\top, W_0(t + t_0), \dots, W_l(t + t_0)]^\top$ and $g_m^{(1,r)} = g_m^{(r)}$. To simplify the notation, we will in the following write $\bullet_{r,m}$ instead of $\bullet_{1,r,m}$, and denote the nodes corresponding to the W_i 's by \bullet_{W_i} or also simply \bullet_t in case of $i = 0$. For their elementary differentials it holds that $F(\bullet_{W_i})(X(t)) = 1$, and all further differentials of these are 0, and their corresponding trees can be omitted from the set of trees. With these considerations we conclude:

Corollary 1. *The solution $X(t)$ of (8) can be written as a B-series for which:*

1. *The set of trees T is defined by*

$$\emptyset \in T \text{ and } \bullet_{r,m} \in T, \\ \tau = [\tau_1, \dots, \tau_\kappa]_{r,m} \in T \text{ for all } \tau_1, \dots, \tau_\kappa \in T \cup \{\bullet_{W_0}, \dots, \bullet_{W_M}\},$$

where $r = 1, \dots, R_m$, $m = 0, \dots, M$.

2. *The elementary differentials are given by*

$$F(\bullet_{r,m})(\tilde{x}_0) = g_m^{(r)}(\tilde{x}_0), \quad F(\bullet_{W_m})(\tilde{x}_0) = 1, \\ F(\tau = [\tau_1, \tau_2, \dots, \tau_\kappa]_{r,m})(\tilde{x}_0) = \left(D_\tau g_m^{(r)}(\tilde{x}_0) \right) (F(\tau_1)(\tilde{x}_0), \dots, F(\tau_\kappa)(\tilde{x}_0)).$$

4. Semi-linear SDEs and exponential integrators

Semi-linear problems have over the years been given quite some attention, and several exponential methods have been developed both for ODEs and SDEs to exploit the special structure of such systems. B-series for semi-linear ODEs have been discussed e.g. in [4, 16]. This has been extended to exponential integrators in [2] and more recently in [31].

Consider a semi-linear non-autonomous SDE of the form

$$dX(t) = A(t)X(t)dt + \sum_{m=0}^M g_m(X(t), t) \star dW_m(t). \quad (10)$$

Example 5. *The SDE presented in Example 1 can be written in this form:*

$$d \begin{pmatrix} R(t) \\ V(t) \end{pmatrix} = \underbrace{\begin{pmatrix} 0 & 1 \\ 0 & -\alpha(t) \end{pmatrix}}_{A(t)} \begin{pmatrix} R(t) \\ V(t) \end{pmatrix} + \underbrace{\begin{pmatrix} 0 \\ f_d(R(t), t) \end{pmatrix}}_{g_0} dt + \underbrace{\begin{pmatrix} 0 \\ f_s(R(t), V(t), t) \end{pmatrix}}_{g_1} dW_1(t).$$

The following class of exponential integrators is considered:

$$H_i = Z_{i0}(A; t_n, h)Y_n + \sum_{m=0}^M \sum_{j=1}^s Z_{ij}^{(m)}(A; t_n, h)g_m(H_j, t_n + c_j h), \quad i = 1, \dots, s, \quad (11a)$$

$$Y_{n+1} = z_0(A; t_n, h)Y_n + \sum_{m=0}^M \sum_{i=1}^s z_i^{(m)}(A; t_n, h)g_m(H_i, t_n + c_i h). \quad (11b)$$

The Magnus type methods proposed by Yang et al. [30] for noncommutative SDEs (9) belong to this class of exponential integrators. For these methods, construction of the random coefficients is discussed in [20].

Yang et al. [31] derived B-series for the exact and numerical solution of this problem, given that the matrix function $A(t)$ is commutative, by expanding exponentials of the form $e^{\int_{t_n}^{t_{n+h}} A(s)ds}$. In the following, it will be demonstrated how these series can be derived significantly simpler by the theory derived above, and without the assumptions on commutativity of $A(t)$.

Let us start with discussing the B-series for the exact solution of (10). This is already a non-autonomous, vertically split system, so the results of Section 3 hold. Let \bullet_m represent g_m , $m = 0, 1, \dots, M$, \bullet_A represent the term $A(t)X(t)$ that is linear with respect to $X(t)$ and \bullet_t be the node corresponding to t . Clearly, the term $A(t)X$ can only be differentiated once with respect to X , thus $F(\tau = [\tau_1, \dots, \tau_\kappa]_A) = 0$ whenever there is more than one $\tau_k \in \cup_{m=0}^M T_m \cup T_A$, that is, a node \bullet_A can have maximum one child which is not in T_t . The following result is thus a consequence of Corollary 1.

Corollary 2. *The solution $X(t)$ of (10) can be written as a B-series where*

1. *The set of trees $T = T_g \cup T_A$ is defined by*

$$\begin{aligned} \bullet_m &\in T_g \text{ for } m = 0, \dots, M \text{ and } \bullet_A \in T_A, \\ \tau = [\tau_1, \dots, \tau_\kappa]_m &\in T_g \text{ for all } m = 0, \dots, M \text{ if } \tau_1, \tau_2, \dots, \tau_\kappa \in T \cup \{\bullet_t\}, \\ \tau = [\tau_1, \bullet_t^{\kappa-1}]_A &\in T_A \text{ if } \tau_1 \in T \cup \{\bullet_t\}. \end{aligned}$$

2. *The elementary differentials are given by*

$$\begin{aligned} F(\bullet_m)(x_0, t_0) &= g_m(x_0, t_0), & F(\bullet_A)(x_0, t_0) &= A(t_0)x_0, & F(\bullet_t)(x_0, t_0) &= 1, \\ F(\tau = [\tau_1, \tau_2, \dots, \tau_\kappa]_m)(x_0, t_0) &= (D_\tau g_m(x_0, t_0)) (F(\tau_1)(x_0, t_0), \dots, F(\tau_\kappa)(x_0, t_0)), \\ F(\tau = [\tau_1, \bullet_t^{\kappa-1}]_A)(x_0, t_0) &= \begin{cases} A^{(\kappa)}(t_0) \cdot x_0 & \text{if } \tau_1 = \bullet_t, \\ A^{(\kappa-1)}(t_0) \cdot F(\tau_1)(x_0, t_0) & \text{otherwise.} \end{cases} \end{aligned}$$

We will now move our attention to the B-series of the exponential RK-method given in (11). In the case of a constant A , the coefficients are usually represented as some kind of functions of Ah or approximations of those, and the order conditions are found by the series expansions of these functions. In the present case, with a time-dependent A , we will assume that the coefficients $Z_{ij}^{(m)}$, $i = 1, \dots, s$, $j = 0, \dots, s$, and $z_i^{(m)}$, $i = 0, \dots, s$, can be written in terms of a B-series

$$Z(A; t_0, h)x_0 = \sum_{\tau \in \bar{T}_A} \alpha(\tau) \cdot Z(\tau)(h) \cdot F(\tau)(x_0, t_0) = B(Z, x_0, t_0; h) \quad (12)$$

where $\bar{T}_A \subset T$ is the subset of T consisting of trees with no \bullet_m nodes, $m = 0, \dots, M$.

Theorem 3. *Assume that the method coefficients $Z_{i0}(A; t_0, h)$, $Z_{ij}^{(m)}(A; t_0, h)$, $z_0(A; t_0, h)$ and $z_i^{(m)}(A; t_0, h)$ all can be written as B-series of the form (12), with $Z_{i0}(\emptyset) = z_0(\emptyset) \equiv$*

1. Then the stage values H_i and the numerical solution after one step Y_i can both be written as B -series

$$H_i = B(\Phi_i, x_0, t_0; h), \quad i = 1, \dots, s, \quad Y_1 = B(\Phi, x_0, t_0; h),$$

where trees and elementary differentials are defined in Corollary 2 and with

$$\begin{aligned} \Phi_i(\emptyset)(h) &= \Phi(\emptyset)(h) = 1, \\ \Phi_i(\bullet_m)(h) &= \sum_{j=1}^s Z_{ij}^{(m)}(\emptyset)(h), \quad \Phi_i(\bullet_t)(h) = c_i h, \quad i = 1, \dots, s, \\ \Phi(\bullet_m)(h) &= \sum_{i=1}^s z_i^{(m)}(\emptyset)(h), \\ \Phi_i(\tau)(h) &= \begin{cases} Z_{i0}(\tau)(h) & \text{if } \tau \in \bar{T}_A, \\ \sum_{j=1}^s Z_{ij}^{(m)}(\vartheta)(h) \cdot \prod_{k=1}^{\kappa} \Phi_j(\tau_k)(h) & \text{otherwise,} \end{cases} \\ \Phi(\tau)(h) &= \begin{cases} z_0(\tau)(h) & \text{if } \tau \in \bar{T}_A, \\ \sum_{i=1}^s z_i^{(m)}(\vartheta)(h) \cdot \prod_{k=1}^{\kappa} \Phi_i(\tau_k)(h) & \text{otherwise,} \end{cases} \end{aligned}$$

where $(\vartheta, \{\delta\})$ is the only non-zero element of $SP(\tau)$ for which $\vartheta \in \bar{T}_A$ and $\delta = [\tau_1, \dots, \tau_\kappa]_m \in T_g$.

Example 6. Let

$$\tau = [[[\bullet_t^2]_A, \bullet_0]_1, \bullet_t]_A =$$

in which the stochastic node with color $m = 1$ is represented by a white node with index 1, the deterministic ones with shape 0 or t are represented by black nodes with the shape as index, and the deterministic nodes with shape A are represented by red nodes with index A . This is the same tree as presented in Example 2, where the mapping of the corresponding node indices is given by $110 \rightarrow 0$, $111 \rightarrow 1$, $120 \rightarrow A$ and $210 \rightarrow t$. The elementary differential $F(\tau)$ can in this case be directly derived from Corollary 2, $F(\tau) = \dot{A}D_{\tau_1}g_1(\dot{A}, g_0)$ where $\tau_1 = [[\bullet_t^2]_A, \bullet_0]_1$, $\dot{A} = dA/dt$ and $\ddot{A} = d^2A/dt^2$. The weight function for the exact solution $\varphi(\tau)$ has not changed from Example 2,

$$\varphi(\tau) = \frac{1}{3} \int_0^h \int_0^{s_1} s_2^4 \star dW_1(s_2) s_1 ds_1,$$

and the numerical weight function can be derived from Theorem 3 by noting that according to Example 3, the only element $(\vartheta, \{\delta\})$ in $SP(\tau)$ with $\vartheta \in \bar{T}_A$ and $\delta \in T_g$ is given by

$$(\vartheta, \{\delta\}) = \left(\begin{array}{c} \bullet_{21} \\ \diagup \quad \diagdown \\ \bullet_{12} \end{array}, \left\{ \begin{array}{c} \bullet_{21} \quad \bullet_{21} \\ \diagup \quad \diagdown \quad \diagup \quad \diagdown \\ \bullet_{12} \quad \bullet_{11} \\ \diagup \quad \diagdown \\ \bullet_{111} \end{array} \right\} \right) = \left(\begin{array}{c} \bullet_t \\ \diagup \quad \diagdown \\ \bullet_A \end{array}, \left\{ \begin{array}{c} \bullet_t \quad \bullet_t \\ \diagup \quad \diagdown \\ \bullet_A \quad \bullet_0 \\ \diagup \quad \diagdown \\ \bullet_1 \end{array} \right\} \right).$$

$$\Phi(\tau) = \sum_{i=1}^s z_i^{(1)}([\bullet_t]_A) \cdot \Phi_i([\bullet_t^2]_A) \cdot \Phi_i(\bullet_0) = \sum_{i,j=1}^s z_i^{(1)}([\bullet_t]_A) \cdot Z_{i0}([\bullet_t^2]_A) \cdot Z_{ij}^{(0)}(\emptyset).$$

Proof of Theorem 3. From Lemma 1 the following can be deduced if $\phi(\emptyset) \equiv 1$:

$$g_m(B(\phi, x_0, t_0; h), t_0 + ch) = B(\psi_\phi^{(m)}, x_0, t_0; h)$$

where

$$\psi_\phi^{(m)}(\tau)(h) = \begin{cases} \prod_{k=1}^{\kappa} \phi(\tau_k)(h) & \text{if } \tau = [\tau_1, \dots, \tau_\kappa]_m, \\ 0 & \text{otherwise} \end{cases}$$

with the definition $\phi(\bullet_t) = ch$. Further, by (12), $Z(A; t_0, h) = \partial_2 B(Z, x_0, t_0; h)$, so Lemma 2 applies to $Z(A; t_0, h)g_m(B(\phi, x_0, t_0; h), t_0 + ch)$.

Assuming that $H_i = B(\Phi_i, x_0, t_0; h)$ and $Y_1 = B(\Phi, x_0, t_0; h)$, the method (11) can be written in terms of B-series

$$B(\Phi_i, x_0, t_0; h) = B(Z_{i,0}, x_0, t_0; h) + \sum_{m=0}^M \sum_{j=1}^s \partial_2 B(Z_{ij}^{(m)}, x_0, t_0; h) \cdot B(\psi_{\Phi_j}^{(m)}, x_0, t_0; h),$$

$$B(\Phi, x_0, t_0; h) = B(z_0, x_0, t_0; h) + \sum_{m=0}^M \sum_{i=1}^s \partial_2 B(z_i^{(m)}, x_0, t_0; h) \cdot B(\psi_{\Phi_i}^{(m)}, x_0, t_0; h).$$

Now, by comparing term by term and applying Lemma 2, the statement of the theorem follows by induction over the height of the trees, starting with the induction hypothesis $\Phi(\emptyset) = \Phi_i(\emptyset) \equiv 1$. In addition, since each \bullet_A -node can have maximum one \bullet_m -node as root of a child tree, for each $\tau \in T$ there is maximum one non-zero element $(\vartheta, \{\delta\})$ in $SP(\tau)$ with $\vartheta \in \bar{T}_A$ and $\delta \in T_g$. \square

We will conclude the paper with an example from Yang et.al. [31], in which low order terms are found and mean square convergence of order 1 for Stratonovich SDEs is proved. We do not intend to reproduce their results, but we will only use this example as a demonstration on how our theory can be applied.

Example 7. *Given a semi-linear Stratonovich SDE with one-dimensional noise:*

$$dX = A(t)X(t)dt + g_0(X(t), t)dt + g_1(X(t), t) \circ dW.$$

The following exponential midpoint rule was proposed by Yang et.al. [31]:

$$H_1 = e^{\int_{t_n}^{t_n+h/2} A(s)ds} x_0 + \frac{h}{2} g_0(H_1, t_n + \frac{h}{2}) + \frac{\Delta W_n}{2} g_1(H_1, t_n + \frac{h}{2}),$$

$$Y_{n+1} = e^{\int_{t_n}^{t_n+h} A(s)ds} x_0 + e^{\int_{t_n}^{t_n+h/2} A(s)ds} \left(h g_0(H_1, t_n + \frac{h}{2}) + \Delta W_n g_1(H_1, t_n + \frac{h}{2}) \right),$$

where $\Delta W_n = W(t_{n+1}) - W(t_n)$. Thus $s = 1$ and $c_1 = 1/2$. The method coefficients are

$$\begin{aligned} Z_{10}(A; t_n, h) &= e^{\int_{t_n}^{t_n+h/2} A(s) ds} \\ &= I + \frac{h}{2} A_n + \frac{h^2}{8} (\dot{A}_n + A_n^2) + \frac{h^3}{48} (\ddot{A}_n + \frac{3}{2} (\dot{A}_n A_n + A_n \dot{A}_n) + A_n^3) + \dots, \\ Z_{11}^{(0)}(A; t_n, h) &= \frac{h}{2}, \quad Z_{11}^{(1)} = \frac{\Delta W_n}{2}, \\ z_0(A; t_n, h) &= e^{\int_{t_n}^{t_n+h} A(s) ds} \\ &= I + h A_n + \frac{h^2}{2} (\dot{A}_n + A_n^2) + \frac{h^3}{6} (\ddot{A}_n + \frac{3}{2} (\dot{A}_n A_n + A_n \dot{A}_n) + A_n^3) + \dots, \\ z_1^{(0)}(A; t_n, h) &= e^{\int_{t_n}^{t_n+h/2} A(s) ds} h = hI + \frac{h^2}{2} A_n + \frac{h^3}{8} (3\dot{A}_n + A_n^2) + \dots, \\ z_1^{(1)}(A; t_n, h) &= e^{\int_{t_n}^{t_n+h/2} A(s) ds} \Delta W_n I = \Delta W_n + \frac{h \Delta W_n}{2} A_n + \frac{h^2 \Delta W_n}{8} (3\dot{A}_n + A_n^2) + \dots, \end{aligned}$$

where $A_n = A(t_n)$, $\dot{A}_n = \dot{A}(t_n)$, $\ddot{A}_n = \ddot{A}(t_n)$.

It is clear that all method coefficients can be written as B -series of the form (12) and fulfill the conditions required in Theorem 3.

This can now be used to find the weight function $\Phi(\tau)(h)$ in Example 6. As $s = 1$ and $z_1^{(1)}([\bullet_t]_A) = 3h^2 \Delta W_n / 8$, $Z_{11}^{(0)}(\emptyset) = h/2$, $Z_{10}([\bullet_t]_A) = h^3/48$ we get $\Phi(\tau) = 3h^6 \Delta W / 768$. Clearly $\Phi(\tau) \neq \varphi(\tau)$, so the order condition related to this particular tree is not satisfied, which could also not be expected, as the order of the tree is 6.5.

Note that although only SDEs with a commutative linear term are considered in [31], the above shows that this result also holds for the noncommutative case, as the Lie-bracket $\dot{A}A - A\dot{A}$ is a term of order $\mathcal{O}(h^3)$ and will thus have no effect on the order for a first order method.

Appendix A. Example 2 applied to the SDE in Example 1

We will here provide the details on how the result presented in Example 2 for the elementary differential looks like more concrete for the SDE introduced in Example 1. The partitioned SDE in Example 1 is of the form

$$\begin{aligned} dX^{(1)} &= g_0^{(1,1)}(X^{(1)}, X^{(2)})dt + g_0^{(1,2)}(X^{(1)}, X^{(2)})dt + g_1^{(1,1)}(X^{(1)}, X^{(2)})dW_1(t) \\ dX^{(2)} &= g_0^{(2,1)}(X^{(1)}, X^{(2)})dt \end{aligned}$$

with

$$X^{(1)} = \begin{pmatrix} R \\ V \end{pmatrix}, \quad X^{(2)} = t.$$

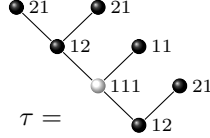
The elementary differential $F(\tau)$ is then step by step developed, starting with

$$\begin{aligned} F(\bullet_{11}) &= g_0^{(1,1)} = \begin{pmatrix} 0 \\ f_d(R, t) \end{pmatrix}, & F(\bullet_{12}) &= g_0^{(1,2)} = \begin{pmatrix} V \\ -\alpha(t)V \end{pmatrix}, \\ F(\circ_{111}) &= g_1^{(1,1)} = \begin{pmatrix} 0 \\ f_s(R, V, t) \end{pmatrix}, & F(\bullet_{21}) &= g_0^{(2,1)} = 1. \end{aligned}$$

From Definition 2 we derive

$$F\left(\begin{array}{c} \bullet_{21} \quad \bullet_{21} \\ \diagdown \quad / \\ \bullet_{12} \end{array}\right) = \begin{pmatrix} 0 \\ -\ddot{\alpha}V \end{pmatrix}, \quad F\left(\begin{array}{c} \bullet_{21} \quad \bullet_{21} \\ \diagdown \quad / \\ \bullet_{12} \\ \diagdown \quad / \\ \bullet_{111} \end{array}\right) = \begin{pmatrix} 0 \\ \frac{\partial^2 f_s}{\partial V^2} \cdot (-\ddot{\alpha}V) \cdot f_d \end{pmatrix},$$

and finally for the tree



of Example 2

$$F(\tau) = \begin{pmatrix} 0 \\ -\dot{\alpha} \cdot \frac{\partial^2 f_s}{\partial V^2} \cdot (-\ddot{\alpha}V) \cdot f_d \end{pmatrix}.$$

We notice that in the case when f_s only depends on the position R , the elementary differential for this tree is 0, thus even if the order condition is not satisfied for this particular tree, it will still not contribute to the total error.

Acknowledgements

The authors would like to thank the unknown referees for their very helpful comments.

References

- [1] Sverre Anmarkrud, Kristian Debrabant, and Anne Kværnø. General order conditions for stochastic partitioned Runge-Kutta methods. *BIT*, 58(2):257–280, 2018.
- [2] Alemayehu Adugna Arara, Kristian Debrabant, and Anne Kværnø. Stochastic B-series and order conditions for exponential integrators. In *Numerical Mathematics and Advanced Applications*, Lecture Notes in Computational Science and Engineering. Springer, 2019.
- [3] Uri M. Ascher, Steven J. Ruuth, and Raymond J. Spiteri. Implicit-explicit Runge-Kutta methods for time-dependent partial differential equations. *Appl. Numer. Math.*, 25(2-3):151–167, 1997. Special issue on time integration (Amsterdam, 1996).
- [4] Håvard Berland, Brynjulf Owren, and Bård Skaflestad. B-series and order conditions for exponential integrators. *SIAM J. Numer. Anal.*, 43(4):1715–1727, 2005.
- [5] Kevin Burrage and Pamela Marion Burrage. High strong order explicit Runge-Kutta methods for stochastic ordinary differential equations. *Appl. Numer. Math.*, 22(1-3):81–101, 1996. Special issue celebrating the centenary of Runge-Kutta methods.
- [6] Kevin Burrage and Pamela Marion Burrage. Order conditions of stochastic Runge-Kutta methods by B-series. *SIAM J. Numer. Anal.*, 38(5):1626–1646, 2000.
- [7] Pamela Marion Burrage. *Runge-Kutta methods for stochastic differential equations*. PhD thesis, The University of Queensland, Brisbane, 1999.
- [8] John C. Butcher. Coefficients for the study of Runge-Kutta integration processes. *J. Austral. Math. Soc.*, 3:185–201, 1963.
- [9] Elena Celledoni, Robert Iain McLachlan, Brynjulf Owren, and Gilles Reinout Willem Quispel. Energy-preserving integrators and the structure of B-series. *Found. Comput. Math.*, 10(6):673–693, 2010.
- [10] Kristian Debrabant and Anne Kværnø. B-series analysis of stochastic Runge-Kutta methods that use an iterative scheme to compute their internal stage values. *SIAM J. Numer. Anal.*, 47(1):181–203, 2008/09.

- [11] Kristian Debrabant and Anne Kværnø. B-series analysis of iterated Taylor methods. *BIT*, 51(3):529–553, 2011.
- [12] Kristian Debrabant and Anne Kværnø. Composition of stochastic B-series with applications to implicit Taylor methods. *Appl. Numer. Math.*, 61(4):501–511, 2011.
- [13] Kristian Debrabant, Anne Kværnø, and Nicky Cordua Mattsson. Runge-Kutta Lawson schemes for stochastic differential equations. *BIT*, 61(2):381–409, 2021.
- [14] Ernst Hairer, Christian Lubich, and Gerhard Wanner. *Geometric numerical integration*, volume 31 of *Springer Series in Computational Mathematics*. Springer, Heidelberg, 2010. Structure-preserving algorithms for ordinary differential equations, Reprint of the second (2006) edition.
- [15] Ernst Hairer and Gerhard Wanner. Multistep-multistage-multiderivative methods of ordinary differential equations. *Computing (Arch. Elektron. Rechnen)*, 11(3):287–303, 1973.
- [16] Marlis Hochbruck and Alexander Ostermann. Explicit exponential Runge-Kutta methods for semi-linear parabolic problems. *SIAM J. Numer. Anal.*, 43(3):1069–1090, 2005.
- [17] Peter Eris Kloeden and Eckhard Platen. *Numerical solution of stochastic differential equations*, volume 23 of *Applications of Mathematics*. Springer-Verlag, Berlin, 2 edition, 1999.
- [18] Yoshio Komori. Multi-colored rooted tree analysis of the weak order conditions of a stochastic Runge-Kutta family. *Appl. Numer. Math.*, 57(2):147–165, 2007.
- [19] Yoshio Komori, Taketomo Mitsui, and Hiroshi Sugiura. Rooted tree analysis of the order conditions of ROW-type scheme for stochastic differential equations. *BIT*, 37(1):43–66, 1997.
- [20] Yoshio Komori, Guoguo Yang, and Kevin Burrage. Formulae for mixed moments of wiener processes and a stochastic area integral. *SIAM Journal on Numerical Analysis*, 61(4):1716–1736, 2023.
- [21] Grigori N. Milstein, Yuri M. Repin, and Michael V. Tretyakov. Numerical methods for stochastic systems preserving symplectic structure. *SIAM J. Numer. Anal.*, 40(4):1583–1604, 2002.
- [22] Hans Munthe-Kaas. Lie-Butcher theory for Runge-Kutta methods. *BIT*, 35(4):572–587, 1995.
- [23] Hans Munthe-Kaas and Olivier Verdier. Aromatic Butcher series. *Found. Comput. Math.*, 16(1):183–215, 2016.
- [24] Hans Z. Munthe-Kaas and Alexander Lundervold. On post-Lie algebras, Lie-Butcher series and moving frames. *Found. Comput. Math.*, 13(4):583–613, 2013.
- [25] Bernt Øksendal. *Stochastic differential equations*. Universitext. Springer-Verlag, Berlin, sixth edition, 2003. An introduction with applications.
- [26] Brynjulf Owren. Order conditions for commutator-free Lie group methods. *J. Phys. A*, 39(19):5585–5599, 2006.
- [27] Andreas Rößler. Stochastic Taylor expansions for the expectation of functionals of diffusion processes. *Stochastic Anal. Appl.*, 22(6):1553–1576, 2004.
- [28] Andreas Rößler. Rooted tree analysis for order conditions of stochastic Runge-Kutta methods for the weak approximation of stochastic differential equations. *Stoch. Anal. Appl.*, 24(1):97–134, 2006.
- [29] Tianhai Tian and Kevin Burrage. Implicit Taylor methods for stiff stochastic differential equations. *Appl. Numer. Math.*, 38(1-2):167–185, 2001.
- [30] Guoguo Yang, Kevin Burrage, Yoshio Komori, Pamela Burrage, and Xiaohua Ding. A class of new Magnus-type methods for semi-linear non-commutative Itô stochastic differential equations. *Numer. Algorithms*, 88(4):1641–1665, 2021.
- [31] Guoguo Yang, Kevin Burrage, Yoshio Komori, and Xiaohua Ding. A new class of structure-preserving stochastic exponential Runge-Kutta integrators for stochastic differential equations. *BIT*, 62(4):1591–1623, 2022.