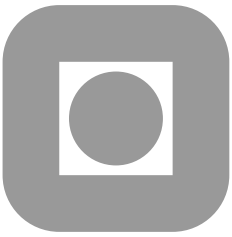
NORGES TEKNISK-NATURVITENSKAPELIGE
UNIVERSITET

Convergence of stochastic Runge-Kutta methods that use an iterative scheme to compute their internal stage values

by

Kristian Debrabant and Anne Kværnø

PREPRINT
NUMERICS NO. 4/2007



NORWEGIAN UNIVERSITY OF
SCIENCE AND TECHNOLOGY
TRONDHEIM, NORWAY

This report has URL
http://www.math.ntnu.no/preprint/numerics/2007/N4-2007.pdf
Address: Department of Mathematical Sciences, Norwegian University of Science and Technology, N-7491 Trondheim, Norway.

CONVERGENCE OF STOCHASTIC RUNGE-KUTTA METHODS THAT USE AN ITERATIVE SCHEME TO COMPUTE THEIR INTERNAL STAGE VALUES

KRISTIAN DEBRABANT[†] AND ANNE KVÆRNØ[‡]

Abstract. In the last years, implicit SRK methods have been developed both for strong and weak approximation. For these methods, the stage values are only given implicitly. However, in practice these implicit equations are solved by iterative schemes like simple iteration, modified Newton iteration or full Newton iteration. We employ a unifying approach for the construction of stochastic B-series which is valid both for Itô- and Stratonovich-SDEs and applicable both for weak and strong convergence to analyze the order of the iterated Runge-Kutta method. Moreover, the analytical techniques applied in this paper can be of use in many other similar contexts.

Key words. Stochastic Runge-Kutta method, stochastic differential equation, iterative scheme, order, internal stage values, Newton's method, weak approximation, strong approximation, growth functions, stochastic B-series

AMS subject classifications. 65C30, 60H35, 65C20, 68U20

1. Introduction. In many applications, e.g., in epidemiology and financial mathematics, taking stochastic effects into account when modeling dynamical systems often leads to stochastic differential equations (SDEs). Therefore, in recent years, the development of numerical methods for the approximation of SDEs has become a field of increasing interest, see e.g [16, 22] and references therein. Many stochastic schemes fall into the class of stochastic Runge-Kutta (SRK) methods. SRK methods have been studied both for strong approximation [1, 10, 11, 16], where one is interested in obtaining good pathwise solutions, and for weak approximation [7, 9, 16, 19, 21, 30], which focuses on the expectation of functionals of solutions. For solving SDEs which are stiff, implicit SRK methods have to be considered, which also has been done both for strong [4, 11, 12] and weak [8, 12, 17] approximation. For these methods, the stage values are only given implicitly. However, in practice these implicit equations are solved by iterative schemes like simple iteration, modified Newton iteration or full Newton iteration. For the numerical solution of ODEs such iterative schemes have been studied by the means of B-series and rooted trees [13, 14] and it was shown that certain growth functions defined on trees give a precise description of the development of the iterations. The aim of the present paper is to do a similar analysis in the SDE case. In particular, it will be shown that the growth functions describing the iterative schemes are the same in the ODE and the SDE case.

Let $(\Omega, \mathcal{A}, \mathcal{P})$ be a probability space. We denote by $(X(t))_{t\in I}$ the stochastic process which is the solution of either a d-dimensional Itô or Stratonovich SDE defined by

$$X(t) = x_0 + \int_{t_0}^{t} g_0(X(s))ds + \sum_{l=1}^{m} \int_{t_0}^{t} g_l(X(s)) \star dW_l(s) \tag{1.1}$$

with an m-dimensional Wiener process $(W(t))_{t\geq 0}$ and $I = [t_0, T]$. The integral w.r.t. the Wiener process has to be interpreted either as Itô integral with $\star dW_l(s) = dW_l(s)$

[†]Technische Universität Darmstadt, Fachbereich Mathematik, Schloßgartenstr.7, D-64289 Darmstadt, Germany, debrabant@mathematik.tu-darmstadt.de

[‡]Department of Mathematical Sciences, Norwegian University of Science and Technology,N-7491 Trondheim, Norway, anne@math.ntnu.no

or as Stratonovich integral with $\star dW_l(s) = \circ dW_l(s)$. We assume that the Borel-measurable coefficients $g_l : \mathbb{R}^d \to \mathbb{R}^d$ are sufficiently differentiable and satisfy a Lipschitz and a linear growth condition such that the Existence and Uniqueness Theorem [15] applies. To simplify the presentation, we define $W_0(s) = s$, so that (1.1) can be written as

$$X(t) = x_0 + \sum_{l=0}^{m} \int_{t_0}^{t} g_l(X(s)) \star dW_l(s). \tag{1.2}$$

In the following we denote by Ξ a set of families of measurable mappings,

$$\Xi := \big\{ \{\varphi(h)\}_{h\geq 0} : \ \varphi(h) : \ \Omega \to \mathbb{R} \text{ is } \mathcal{A}-\mathcal{B}\text{-measurable } \forall h \geq 0 \big\}$$

and by Ξ_0 the subset of Ξ defined by

$$\Xi_0 := \big\{ \{\varphi(h)\}_{h\geq 0} \in \Xi : \ \varphi(0) \equiv 0 \big\}.$$

Let a discretization $I_h = \{t_0, t_1, \ldots, t_N\}$ with $t_0 < t_1 < \ldots < t_N = T$ of the time interval I with step sizes $h_n = t_{n+1} - t_n$ for $n = 0, 1, \ldots, N-1$ be given. Now, we consider the general class of s-stage SRK methods given by $Y_0 = x_0$ and

$$Y_{n+1} = Y_n + \sum_{l=0}^{m} \sum_{\nu=0}^{M} (z^{(l,\nu)^\top} \otimes I_d) g_l(H^{(l,\nu)}) \tag{1.3a}$$

for $n = 0, 1, \ldots, N-1$ with $Y_n = Y(t_n)$, $t_n \in I_h$, and

$$H^{(l,\nu)} = \mathbb{1}_s \otimes Y_n + \sum_{r=0}^{m} \sum_{\mu=0}^{M} (Z^{(l,\nu)(r,\mu)} \otimes I_d) g_r(H^{(r,\mu)}) \tag{1.3b}$$

for $l = 0, \ldots, m$ and $\nu = 0, \ldots, M$ with

$$g_l(H^{(l,\nu)}) = (g_l(H_1^{(l,\nu)})^\top, \ldots, g_l(H_s^{(l,\nu)})^\top)^\top$$

and

$$z^{(l,\nu)} \in \Xi_0^s, \qquad Z^{(l,\nu)(r,\mu)} \in \Xi_0^{s,s}$$

for $l, r = 0, \ldots, m$, $\mu, \nu = 0, \ldots, M$.

The formulation (1.3) includes the classes of SRK methods considered in [4, 11, 18, 20, 27, 28, 29] as well as the SRK methods considered in [12, 16, 24] and defines a d-dimensional approximation process Y with $Y_n = Y(t_n)$.

Remark 1 (1.3) is equivalent to the class of SRK methods

$$Y_i = y_n + \sum_{l=0}^{m} \sum_{j=1}^{s} Z_{ij}^l g_l(Y_j), \qquad y_{n+1} = y_n + \sum_{l=0}^{m} \sum_{i=1}^{s} z_i^l g_l(Y_i)$$

defined by Burrage and Burrage:
$\Rightarrow$ Put $M = 0$ and choose the coefficients such that $H_i^{(l,0)}$ is independent of l. $\Leftarrow$ Replace the stage index by a multi index, i. e.

$$Y_{(i,l,\nu)} = y_n + \sum_{l=0}^{m} \sum_{(j,r,\mu)} Z_{(i,l,\nu),(j,r,\mu)}^{(k)} g_k(Y_{(j,r,\mu)}), \qquad y_{n+1} = y_n + \sum_{l=0}^{m} \sum_{(i,l,\nu)} z_{(i,l,\nu)}^{(k)} g_k(Y_{(i,l,\nu)})$$

(this means mMs stages) and set $Z_{(i,l,\nu),(j,r,\mu)}^{(r)} = Z_{i,j}^{(l,\nu)(r,\mu)}$, $z_{(i,l,\nu)}^{(l)} = z_i^{(l,\nu)}$ and the remaining coefficients to zero.

Application of an iterative scheme yields

$$H_{k+1}^{(l,\nu)} = \mathbb{1}_s \otimes Y_n + \sum_{r=0}^{m}\sum_{\mu=0}^{M}\left(Z^{(l,\nu)(r,\mu)} \otimes I_d\right) g_r(H_k^{(r,\mu)}) + \sum_{r=0}^{m}\sum_{\mu=0}^{M}\left(Z^{(l,\nu)(r,\mu)} \otimes I_d\right) J_k^{(r,\mu)}\left(H_{k+1}^{(r,\mu)} - H_k^{(r,\mu)}\right), \tag{1.4a}$$

$$Y_{n+1,k} = Y_n + \sum_{l=0}^{m}\sum_{\nu=0}^{M}({z^{(l,\nu)}}^{\top} \otimes I_d) g_l(H_k^{(l,\nu)}) \tag{1.4b}$$

with some approximation $J_k^{(r,\mu)}$ to the Jacobian of $g_r(H_k^{(r,\mu)})$ and a predictor $H_0^{(l,\nu)}$. In the following we assume that (1.4a) can be solved uniquely at least for small enough h_n.

Remark 2 If this iteration is performed for $k = 0, \dots, K$, then in practice we take $y_{n+1} = y_{n+1,K}$ and proceed to the next step, without calculating $y_{n+1,k}$ for $k < K$.

For the approximation $J_k^{(r,\mu)}$ to the Jacobian of $g_r(H_k^{(r,\mu)})$ there exist several common choices. If we choose $J_k^{(r,\mu)}$ to be the exact Jacobian $g_r'(H_k^{(r,\mu)})$, then we obtain the classical Newton iteration method for solving (1.3b), which is locally convergent of order two and will be denoted in the following as full Newton iteration. If we choose instead $J_k^{(r,\mu)} = I_s \otimes g_r'(x_0)$, then we obtain the so called modified Newton iteration method, which is locally convergent of order one. Here, the $J_k^{(r,\mu)}$ are independent of the iteration number k and the stage values $H_k^{(r,\mu)}$. Thus their computation is much cheaper and simpler than in the full Newton iteration case, and in Runge-Kutta implementations this is the method which is usually chosen.
The third and simplest possibility is to choose $J_k^{(r,\mu)}$ equal to zero. In this case we don't even have to solve a linear system for $H_{k+1}^{(l,\nu)}$. This iteration method is called simple iteration method or predictor corrector method. Its disadvantage is that for stiff systems it requires very small step sizes to converge.
In the following, we will mainly consider these three iteration methods.

This paper is organized as follows. In Section 2 we recall the definitions of weak and strong convergence, relate them to consistency and present some preliminary results about stochastic B-series. In Section 3 we develop B-series representations of the iterated solution for the three iteration schemes in consideration and give a description via growth functions. These results are related to the order of convergence of the overall scheme in Sections 4 and 5. Numerical examples in Section 6 finally confirm our theoretical results.

2. Some notation, definitions and preliminary results. In this section we introduce some necessary notation and provide a few definitions and preliminary results that will be used later.

2.1. Convergence and consistency. Here we will give the definitions for both weak and strong convergence and results which relate convergence to consistency.

Let $C_P^l(\mathbb{R}^d, \mathbb{R}^{\hat{d}})$ denote the space of all $g \in C^l(\mathbb{R}^d, \mathbb{R}^{\hat{d}})$ fulfilling a polynomial growth condition [16].

Definition 1 *A time discrete approximation $Y = (Y(t))_{t\in I_h}$ converges weakly with order p to X as $h \to 0$ at time $t \in I_h$ 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*

$$|\,\mathrm{E}(f(Y(t))) - \mathrm{E}(f(X(t)))| \le C_f\, h^p$$

holds for each $h \in]0, \delta_0[$.

Now, let $le_f(h; t, x)$ be the weak local error of the method starting at the point (t, x) with respect to the functional f and step size h, i. e.

$$le_f(h; t, x) = \mathrm{E}\big(f(Y(t+h)) - f(X(t+h))|Y(t) = X(t) = x\big).$$

The following theorem due to Milstein [22], which holds also in the case of general one step methods, shows that, as in the deterministic case, consistency implies convergence:

Theorem 1 *Suppose the following conditions hold:*

- *The coefficients g_l are continuous, satisfy a Lipschitz condition and belong to $C_P^{2(p+1)}(\mathbb{R}^d, \mathbb{R}^d)$ for $l = 0, \ldots, m$.*
- *For sufficiently large r (see, e.g., [22] for details) the moments $\mathrm{E}(\|Y(t_n)\|^{2r})$ exist for $t_n \in I_h$ and are uniformly bounded with respect to N and $n = 0, 1, \ldots, N$.*
- *Assume that for all $f \in C_P^{2(p+1)}(\mathbb{R}^d, \mathbb{R})$ there exists a $K \in C_P^0(\mathbb{R}^d, \mathbb{R})$ such that*

$$|le_f(h; t, x)| \le K(x)\, h^{p+1}$$

is valid for $x \in \mathbb{R}^d$ and $t, t+h \in I_h$, i. e., the approximation is weak consistent of order p.

Then the method (1.3) is convergent of order p in the sense of weak approximation.

Whereas weak approximation methods are used to estimate the expectation of functionals of the solution, strong approximation methods approach the solution pathwise.

Definition 2 *A time discrete approximation $Y = (Y(t))_{t\in I_h}$ converges strongly respectively in the mean square with order p to X as $h \to 0$ at time $t \in I_h$ if there exist a constant C and a finite $\delta_0 > 0$ such that*

$$\mathrm{E}\,\|Y(t) - X(t)\| \le C\, h^p$$

respectively

$$\sqrt{\mathrm{E}(\|Y(t) - X(t)\|^2)} \le C\, h^p$$

holds for each $h \in]0, \delta_0[$.

In this article we will consider convergence in the mean square sense. But as by Jensen's inequality we have

$$(\mathrm{E}\,\|Y(t) - X(t)\|)^2 \le \mathrm{E}(\|Y(t) - X(t)\|^2),$$

mean square convergence implies strong convergence of the same order.

Now, let $le^m(h;t,x)$ respectively $le^{ms}(h;t,x)$ be the mean respectively mean square local error of the method starting at the point (t,x) with respect to the step size h, i. e.

$$le^m(h;t,x) = \mathrm{E}\big(Y(t+h) - X(t+h)|Y(t) = X(t) = x\big),$$
$$le^{ms}(h;t,x) = \sqrt{\mathrm{E}\big((Y(t+h) - X(t+h))^2|Y(t) = X(t) = x\big)}.$$

The following theorem due to Milstein [22] which holds also in the case of general one step methods shows that in the mean square convergence case we obtain order p if the mean local error is consistent of order p and the mean square local error is consistent of order $p - \frac{1}{2}$.

Theorem 2 *Suppose the following conditions hold:*

- *The coefficients g_l are continuous and satisfy a Lipschitz and a linear growth condition for $l = 0, \dots, m$, and $\mathrm{E}(\|X(t_0)\|^2) < \infty$.*
- *There exists a constant K independent of h such that*

$$|le^m(h;t,x)| \le K\sqrt{1+\|x\|^2}\,h^{p_1}, \qquad le^{ms}(h;t,x) \le K\sqrt{1+\|x\|^2}\,h^{p+\frac{1}{2}}$$

  with $p \ge 0$, $p_1 \ge p+1$ is valid for $x \in \mathbb{R}^d$ and $t, t+h \in I_h$, i. e., the approximation is consistent in the mean of order $p_1 - 1 \ge p$ and in the mean square of order $p - \frac{1}{2}$.

Then the SRK method (1.3) is convergent of order p in the sense of mean square approximation.

For Stratonovich SDEs, this result is also obtained by Burrage and Burrage [2].

2.2. Stochastic B-series. In this section we will develop stochastic B-series for the solution of (1.2) as well as for the numerical solution given by (1.3). B-series for deterministic ODEs were introduced by Butcher [6]. Today such series appear as a fundamental tool to do local error analysis on a wide range of problems. B-series for SDEs have been developed by Burrage and Burrage [1, 2] to study strong convergence in the Stratonovich case, by Komori, Mitsui and Sugiura [20] and Komori [18] to study weak convergence in the Stratonovich case and by Rößler [25, 26] to study weak convergence in both the Itô and the Stratonovich case. However, the distinction between the Itô and the Stratonovich integrals only depends on the definition of the integrals, not on how the B-series are constructed. Similarly, the distinction between weak and strong convergence only depends on the definition of the local error. Thus, we find it convenient to present a uniform and self-contained theory for the construction of stochastic B-series. We will present results and proofs in a certain detail, since some intermediate results will be used in later sections.

Following the idea of Burrage and Burrage we introduce the set of colored, rooted trees related to the SDE (1.1), as well as the elementary differentials associated with each of these trees.

Definition 3 (Trees) *The set of $m+1$-colored, rooted trees*

$$T = \{\emptyset\} \cup T_0 \cup T_1 \cup \dots \cup T_m$$

is recursively defined as follows:

a) *The graph $\bullet_l = [\emptyset]_l$ with only one vertex of color l belongs to T_l.*

Let $\tau = [\tau_1, \tau_2, \cdots, \tau_\kappa]_l$ be the tree formed by joining the subtrees $\tau_1, \tau_2, \cdots, \tau_\kappa$ each by a single branch to a common root of color l.
b) *If $\tau_1, \tau_2, \cdots, \tau_\kappa \in T$ then $\tau = [\tau_1, \tau_2, \cdots, \tau_\kappa]_l \in T_l$.*

Thus, T_l is the set of trees with an l-colored root, and T is the union of these sets.

Definition 4 (Elementary differentials) *For a tree $\tau \in T$ the elementary differential is a mapping $F(\tau) : \mathbb{R}^d \to \mathbb{R}^d$ defined recursively by*
a) $F(\emptyset)(x_0) = x_0$,

b) $F(\bullet_l)(x_0) = g_l(x_0)$,

c) *If $\tau = [\tau_1, \tau_2, \cdots, \tau_\kappa]_l \in T_l$ then*

$$F(\tau)(x_0) = g_l^{(\kappa)}(x_0)\big(F(\tau_1)(x_0), F(\tau_2)(x_0), \cdots, F(\tau_\kappa)\big).$$

As will be shown in the following, both the exact and the numerical solutions, including the iterated solutions as we will see later, can formally be written in terms of B-series.

Definition 5 (B-series) *Given a mapping $\phi : T \to \Xi$ satisfying*

$$\phi(\emptyset)(h) \equiv 1 \quad \textit{and} \quad \phi(\tau)(0) \equiv 0, \quad \forall \tau \in T \backslash \{\emptyset\}.$$

A (stochastic) B-series is then a formal series of the form

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

where $\alpha : T \to \mathbb{Q}$ is given by

$$\alpha(\emptyset) = 1, \qquad \alpha(\bullet_l) = 1, \qquad \alpha(\tau = [\tau_1, \cdots, \tau_\kappa]_l) = \frac{1}{r_1! r_2! \cdots r_q!} \prod_{j=1}^{\kappa} \alpha(\tau_j)$$

where $r_1, r_2, \cdots, r_q$ count equal trees among $\tau_1, \tau_2, \cdots, \tau_\kappa$.

If $\phi : T \to \Xi^s$ then $B(\phi, x_0; h) = [B(\phi_1, x_0; h), \cdots, B(\phi_s, x_0; h)]^\top$.

The next lemma proves that if $Y(h)$ can be written as a B-series, then $f(Y(h))$ can be written as a similar series, where the sum is taken over trees with a root of color f and subtrees in T. The lemma is fundamental for deriving B-series for the exact and the numerical solution. It will also be used for deriving weak convergence results.

Lemma 3 *If $Y(h) = B(\phi, x_0; h)$ is some B-series and $f \in C^\infty(\mathbb{R}^d, \mathbb{R}^{\hat{d}})$ then $f(Y(h))$ can be written as a formal series of the form*

$$f(Y(h)) = \sum_{u \in U_f} \beta(u) \cdot \psi_\phi(u)(h) \cdot G(u)(x_0) \tag{2.1}$$

where U_f is a set of trees derived from T, by
a) *$[\emptyset]_f \in U_f$, and if $\tau_1, \tau_2, \cdots, \tau_\kappa \in T$ then $[\tau_1, \tau_2, \cdots, \tau_\kappa]_f \in U_f$.*

b) $G([\emptyset]_f)(x_0) = f(x_0)$ *and*
$G(u = [\tau_1, \cdots, \tau_\kappa]_f)(x_0) = f^{(\kappa)}(x_0)\big(F(\tau_1)(x_0), \cdots, F(\tau_\kappa)(x_0)\big)$.

c) $\beta([\emptyset]_f) = 1$ *and* $\beta(u = [\tau_1, \cdots, \tau_\kappa]_f) = \dfrac{1}{r_1! r_2! \cdots r_q!} \displaystyle\prod_{j=1}^{\kappa} \alpha(\tau_j)$,
where $r_1, r_2, \cdots, r_q$ *count equal trees among* $\tau_1, \tau_2, \cdots, \tau_\kappa$.

d) $\psi_\phi([\emptyset]_f)(h) \equiv 1$ *and* $\psi_\phi(u = [\tau_1, \cdots, \tau_\kappa]_f)(h) = \prod_{j=1}^{\kappa} \phi(\tau_j)(h)$.

Proof. Writing $Y(h)$ as a B-series, we have

$$\begin{aligned} f(Y(h)) &= f\left(\sum_{\tau \in T} \alpha(\tau) \cdot \phi(\tau)(h) \cdot F(\tau)(x_0)\right) \\ &= \sum_{\kappa=0}^{\infty} \frac{1}{\kappa!} f^{(\kappa)}(x_0) \left(\sum_{\tau \in T \setminus \{\emptyset\}} \alpha(\tau) \cdot \phi(\tau)(h) \cdot F(\tau)(x_0)\right)^{\kappa} \\ &= f(x_0) + \sum_{\kappa=1}^{\infty} \frac{1}{\kappa!} \sum_{\{\tau_1, \tau_2, \cdots \tau_\kappa\} \in T \setminus \{\emptyset\}} \frac{\kappa!}{r_1! r_2! \cdots r_q!} \\ &\quad \cdot \left(\prod_{j=1}^{\kappa} \alpha(\tau_j) \cdot \phi(\tau_j)(h)\right) f^{(\kappa)}(x_0)\big(F(\tau_1)(x_0), \cdots, F(\tau_\kappa)(x_0)\big). \end{aligned}$$

where the last sum is taken over all possible unordered combinations of κ trees in T. For each set of trees $\tau_1, \tau_2, \cdots, \tau_\kappa \in T$ we assign a $u = [\tau_1, \tau_2, \cdots, \tau_\kappa]_f \in U_f$. The theorem is now proved by comparing term by term with (2.1). □

Remark 3 E.g. in the definition of weak convergence, just $f \in C_P^{2(p+1)}(\mathbb{R}^d, \mathbb{R})$ is required. Thus $f(Y(h))$ can only be written as a truncated B-series, with a remainder term. However, to simplify the presentation in the following we assume that all derivatives of $f, g_0, \ldots, g_l$ exist.

We will also need the following result:

Lemma 4 *If* $Y(h) = B(\phi_Y, x_0; h)$ *and* $Z(h) = B(\phi_Z, x_0; h)$ *and* $f \in C^\infty(\mathbb{R}^d, \mathbb{R}^{\hat{d}})$ *then*

$$f'(Y(h))B(Z(h)) = \sum_{u \in U_f} \beta(u) \cdot \Upsilon(u)(h) \cdot G(u)(x_0)$$

with

$$\Upsilon([\emptyset]_f)(h) \equiv 0, \qquad \Upsilon([u = [\tau_1, \cdots, \tau_\kappa]_f)(h) = \sum_{i=1}^{\kappa} \left(\prod_{\substack{j=1 \\ j \neq i}}^{\kappa} \phi_Y(\tau_j)(h)\right) \phi_Z(\tau_i)(h)$$

with $\beta(u)$ *and* $G(u)(x_0)$ *given by Lemma 3.*

The proof is similar to the deterministic case, see [23].

When Lemma 3 is applied to the functions g_l on the right hand side of (1.1) we get the following result: If $Y(h) = B(\phi, x_0; h)$ then

$$g_l(Y(h)) = \sum_{\tau \in T_l} \alpha(\tau) \cdot \phi_l'(\tau)(h) \cdot F(\tau)(x_0) \tag{2.2}$$

in which

$$\phi_l'(\tau)(h) = \begin{cases} 1 & \text{if } \tau = \bullet_l, \\ \prod_{j=1}^{\kappa} \phi(\tau_j)(h) & \text{if } \tau = [\tau_1, \cdots, \tau_\kappa]_l \in T_l. \end{cases}$$

By induction on τ, we deduce the following result, giving the B-series of the exact solution of (1.1):

Theorem 5 *The solution $X(t_0 + h)$ of (1.1) can be written as a B-series $B(\varphi, x_0; h)$ with*

$$\varphi(\emptyset)(h) \equiv 1, \quad \varphi(\bullet_l)(h) = W_l(h), \quad \varphi(\tau = [\tau_1, \cdots, \tau_\kappa]_l)(h) = \int_0^h \prod_{j=1}^{\kappa} \varphi(\tau_j)(s) \star dW_l(s).$$

This is proved for the Stratonovich case in [2, 18], but will also apply to the Itô case. The definition of the order of the tree, $\rho(\tau)$, is motivated by the fact that $\mathrm{E}\, W_l(h)^2 = h$ for $l \geq 1$ and $W_0(h) = h$:

Definition 6 (Order) *The order of a tree $\tau \in T$ is defined by*

$$\rho(\emptyset) = 0, \quad \rho([\tau_1, \ldots, \tau_\kappa]_f) = \sum_{i=1}^{\kappa} \rho(\tau_i)$$

and

$$\rho(\tau = [\tau_1, \ldots, \tau_\kappa]_l) = \sum_{i=1}^{\kappa} \rho(\tau_i) + \begin{cases} 1 & \textit{for } l = 0, \\ \frac{1}{2} & \textit{otherwise.} \end{cases}$$

The following result is similar to results given in [1].

Theorem 6 *If the coefficients $Z^{(l,\nu),(r,\mu)} \in \Xi_0^{s,s}$ and $z^{(l,\nu)} \in \Xi_0^s$ then the numerical solution Y_1 as well as the stage values can be written in terms of B-series*

$$H^{(l,\nu)} = B(\Phi^{(l,\nu)}, x_0; h), \qquad Y_1 = B(\Phi, x_0; h)$$

for all l, ν, with

$$\Phi^{(l,\nu)}(\emptyset)(h) \equiv \mathbb{1}_s, \quad \Phi^{(l,\nu)}(\bullet_r)(h) = \sum_{\mu=0}^{M} Z^{(l,\nu)(r,\mu)} \mathbb{1}_s, \tag{2.3a}$$

$$\Phi^{(l,\nu)}(\tau = [\tau_1, \cdots, \tau_\kappa]_r)(h) = \sum_{\mu=0}^{M} Z^{(l,\nu)(r,\mu)} \prod_{j=1}^{\kappa} \Phi^{(r,\mu)}(\tau_j)(h) \tag{2.3b}$$

and

$$\Phi(\emptyset)(h) \equiv 1, \quad \Phi(\bullet_l)(h) = \sum_{\nu=0}^{M} z^{(l,\nu)} \mathbb{1}_s, \tag{2.4a}$$

$$\Phi(\tau = [\tau_1, \cdots, \tau_\kappa]_l)(h) = \sum_{\nu=0}^{M} z^{(l,\nu)} \prod_{j=1}^{\kappa} \Phi^{(l,\nu)}(\tau_j)(h). \tag{2.4b}$$

Proof. Write $H^{(l,\nu)}$ as a B-series, that is

$$H^{(l,\nu)} = \sum_{\tau \in T} \alpha(\tau) \left(\Phi^{(l,\nu)}(h) \otimes I_d \right) (\mathbb{1}_s \otimes F(\tau)(x_0)) .$$

Use the definition of the method (1.3) together with (2.2) to obtain

$$H^{(l,\nu)} = \mathbb{1}_s \otimes x_0 + \sum_{r=0}^{m} \sum_{\mu=0}^{M} \sum_{\tau \in T_r} \alpha(\tau) \left(\left(Z^{(l,\nu)(r,\mu)} \cdot (\Phi^{(r,\mu)})'_r(\tau)(h) \right) \otimes I_d \right) (\mathbb{1}_s \otimes F(\tau)(x_0))$$

with $(\Phi^{(r,\mu)})'_r(\tau)(h) = \left((\Phi_1^{(r,\mu)})'_r(\tau)(h), \ldots, (\Phi_s^{(r,\mu)})'_r(\tau)(h) \right)^\top$. Comparing term by term gives the relations (2.3). The proof of (2.4) is similar. □

To decide the weak order we will also need the B-series of the function f, evaluated at the exact and the numerical solution. From Theorem 5, Theorem 6 and Lemma 3 we obtain

$$f(X(t_0 + h)) = \sum_{u \in U_f} \beta(u) \cdot \psi_\varphi(u)(h) \cdot G(u)(x_0),$$

$$f(Y_1) = \sum_{u \in U_f} \beta(u) \cdot \psi_\Phi(u)(h) \cdot G(u)(x_0),$$

with

$$\psi_\varphi([\emptyset]_f)(h) \equiv 1, \quad \psi_\varphi(u = [\tau_1, \cdots, \tau_\kappa]_f)(h) = \prod_{j=1}^{\kappa} \varphi(\tau_j)(h)$$

and

$$\psi_\Phi([\emptyset]_f)(h) \equiv 1, \quad \psi_\Phi(u = [\tau_1, \cdots, \tau_\kappa]_f)(h) = \prod_{j=1}^{\kappa} \Phi(\tau_j)(h).$$

So, for the weak local error it follows

$$le_f(h; t, x) = \sum_{u \in U_f} \beta(u) \cdot \mathrm{E}\left[\psi_\Phi(u)(h) - \psi_\varphi(u)(h) \right] \cdot G(u)(x).$$

For the mean respectively mean square local error we obtain from Theorem 5 and Theorem 6

$$le^{ms}(h; t, x) = \sqrt{\mathrm{E}\left(\sum_{\tau \in T} \alpha(\tau) \cdot (\Phi(\tau)(h) - \varphi(\tau)(h)) \cdot F(\tau)(x) \right)^2},$$

$$le^{m}(h; t, x) = \sum_{\tau \in T} \alpha(\tau) \cdot \mathrm{E}\left(\Phi(\tau)(h) - \varphi(\tau)(h) \right) \cdot F(\tau)(x).$$

With all the B-series in place, we can now present the order conditions for the weak and strong convergence, for both the Itô and the Stratonovich case. We have weak consistency of order p if and only if

$$\mathrm{E}\,\psi_\Phi(u)(h) = \mathrm{E}\,\psi_\varphi(u)(h) \quad \forall u \in U_f \text{ with } \rho(u) \leq p + \frac{1}{2}, \tag{2.5}$$

where $\rho(u = [\tau_1, \cdots, \tau_\kappa]_f) = \sum_{j=1}^{\kappa} \rho(\tau_j)$, and mean square global order p if

$$\Phi(\tau)(h) = \varphi(\tau)(h) \quad \forall \tau \in T \text{ with } \rho(\tau) \leq p, \tag{2.6}$$

$$\mathrm{E}\,\Phi(\tau)(h) = \mathrm{E}\,\varphi(\tau)(h) \quad \forall \tau \in T \text{ with } \rho(\tau) = p + \frac{1}{2} \tag{2.7}$$

and all elementary differentials $F(\tau)$ fulfill a linear growth condition. Instead of the last requirement it is also enough to claim that there exists a constant C such that $\|g_j'(y)\| \leq C \quad \forall y \in \mathbb{R}^m$, $j = 0, \ldots, M$ (which implies the global Lipschitz condition) and all necessary partial derivatives exist [2].

3. B-series of the iterated solution and growth functions. In this section we will discuss how the iterated solution defined in (1.4) can be written in terms of B-series, that is

$$H_k^{(l,\nu)} = B(\Phi_k^{(l,\nu)}, x_0; h) \quad \text{and} \quad Y_{1,k} = B(\Phi_k, x_0; h).$$

For notational convenience, in the following the h-dependency of the weight functions will be suppressed, so $\Phi(\tau)(h)$ will be written as $\Phi(\tau)$. Further, all results are valid for all $l = 0, \cdots, m$ and $\nu = 0, \cdots, M$. Assume that the predictor can be written as a B-series,

$$H_0^{(l,\nu)} = B(\Phi_0^{(l,\nu)}, x_0; h),$$

satisfying $\Phi_0^{(l,\nu)}(\emptyset) = \mathbb{1}_s$. The most common situation is the use of the trivial predictor $H^{(l,\nu)} = \mathbb{1}_s \otimes x_0$, for which $\Phi_0^{(l,\nu)}(\emptyset) = \mathbb{1}_s$ and $\Phi_0^{(l,\nu)}(\tau) = 0$ otherwise.

The iteration schemes we discuss here only differ in the choice of $J_k^{(r,\mu)}$ in (1.4). For all schemes, the following lemma applies. The proof follows directly from Lemma 3.

Lemma 7 *If $H_k^{(l,\nu)} = B(\Phi_k^{(l,\nu)}, x_0; h)$ then $Y_{1,k} = B(\Phi_k, x_0; h)$ with*

$$\Phi_k(\emptyset) \equiv 1, \quad \Phi_k(\bullet_l) = \sum_{\nu=0}^{M} z^{l,\nu} \mathbb{1}_s, \quad \Phi_k(\tau = [\tau_1, \cdots, \tau_\kappa]_l) = \sum_{\nu=0}^{M} z^{(l,\nu)} \prod_{j=1}^{\kappa} \Phi_k^{(l,\nu)}(\tau_j).$$

Further,

$$f(Y_{1,k}) = \sum_{u \in U_f} \beta(u) \cdot \psi_{\Phi_k}(u) \cdot G(u)(x_0)$$

with

$$\psi_{\Phi_k}([\emptyset]_f) = 1, \quad \psi_{\Phi_k}(u = [\tau_1, \cdots, \tau_\kappa]_f) = \prod_{j=1}^{\kappa} \Phi_k(\tau_j),$$

where $\beta(u)$ and $G(u)(x_0)$ are given in Lemma 3.

We are now ready to study each of the iteration schemes. In each case, we will first find the recurrence formula for $\Phi_k^{(l,\nu)}(\tau)$. From this we define a growth function $\mathfrak{g}(\tau)$:

Definition 7 (Growth function) *A growth function $\mathfrak{g} : T \to \mathbb{N}$ is a function satisfying*

$$\begin{aligned} &\Phi_k^{(l,\nu)}(\tau) = \Phi^{(l,\nu)}(\tau), \quad \forall \tau \in T, \quad \mathfrak{g}(\tau) \le k \\ &\qquad \Rightarrow \qquad \Phi_{k+1}^{(l,\nu)}(\tau) = \Phi^{(l,\nu)}(\tau), \quad \forall \tau \in T, \quad \mathfrak{g}(\tau) \le k+1, \end{aligned} \tag{3.1}$$

for all $k \ge 0$.

This result should be sharp in the sense that in general there exists $\tau \neq \emptyset$ with $\Phi_0^{(l,\nu)}(\tau) \neq \Phi^{(l,\nu)}(\tau)$ and $\Phi_k^{(l,\nu)}(\tau) \neq \Phi^{(l,\nu)}(\tau)$ when $k < \mathfrak{g}(\tau)$. From Lemma 7 we also have

$$\begin{aligned} &\Phi_k(\tau) = \Phi(\tau) && \forall \tau = [\tau_1, \cdots, \tau_\kappa]_l \in T, && \mathfrak{g}'(\tau) = \max_{j=1}^{\kappa} \mathfrak{g}(\tau_i) \le k, \\ &\psi_{\Phi_k}(u) = \psi_{\Phi(\tau)} && \forall u = [\tau_1, \cdots, \tau_\kappa]_f \in U_f, && \mathfrak{g}'(u) = \max_{j=1}^{\kappa} \mathfrak{g}'(\tau_i) \le k. \end{aligned} \tag{3.2}$$

The growth functions give a precise description of the development of the iterations. As we will see, the growth functions are exactly the same as in the deterministic case, see [13, 14]. However, to get applicable results, we will at the end need the relation between the growth functions and the order. Further, we will also take advantage of the fact that $\mathrm{E}\,\Phi(\tau) = 0$ and $\mathrm{E}\,\psi_\Phi(u) = 0$ for some trees. These aspects are discussed in the next sections.

The simple iteration. Simple iterations are described by (1.4a) with $J_k^{(r,\mu)} = 0$, that is

$$H_{k+1}^{(l,\nu)} = \mathbb{1}_s \otimes x_0 + \sum_{r=0}^{m} \sum_{\mu=0}^{M} \left(Z^{(l,\nu)(r,\mu)} \otimes I_d \right) g_r(H_k^{(r,\mu)}). \tag{3.3}$$

By (2.2) and Theorem 6 we easily get the next two results:

Lemma 8 *If $H_0^{(l,\nu)} = B(\Phi_0^{(l,\nu)}, x_0; h)$ then $H_k^{(l,\nu)} = B(\Phi_k^{(l,\nu)}, x_0; h)$, where*

$$\Phi_{k+1}^{(l,\nu)}(\emptyset) \equiv \mathbb{1}_s, \quad \Phi_{k+1}^{(l,\nu)}(\tau = [\tau_1, \cdots, \tau_\kappa]_r) = \sum_{\mu=0}^{M} Z^{(l,\nu)(r,\mu)} \prod_{j=1}^{\kappa} \Phi_k^{(r,\mu)}(\tau_j).$$

The corresponding growth function is given by

$$\mathfrak{h}(\emptyset) = 0, \quad \mathfrak{h}([\tau_1, \ldots, \tau_\kappa]_l) = 1 + \max_{j=1}^{\kappa} \mathfrak{h}(\tau_j).$$

The function $\mathfrak{h}(\tau)$ is the height of τ, that is the maximum number of nodes along one branch. The functions $\mathfrak{h}'(\tau)$ and $\mathfrak{h}'(u)$ are defined by (3.2), with $\mathfrak{g}$ replaced by $\mathfrak{h}$.

The modified Newton iteration. In this subsection, we consider the modified Newton iteration (1.4a) with $J_k^{(r,\mu)} = I_s \otimes g_r'(x_0)$. The B-series for $H_k^{(l,\nu)}$ and the corresponding growth function can now be described by the following lemma:

Lemma 9 *If* $H_0^{(l,\nu)} = B(\Phi_0^{(l,\nu)}, x_0; h)$ *then* $H_k^{(l,\nu)} = B(\Phi_k^{(l,\nu)}, x_0; h)$ *with*

$$\Phi_{k+1}^{(l,\nu)}(\emptyset) \equiv \mathbb{1}_s,$$
$$\Phi_{k+1}^{(l,\nu)}(\tau) = \begin{cases} \displaystyle\sum_{\mu=0}^{M} Z^{(l,\nu)(r,\mu)} \prod_{j=1}^{\kappa} \Phi_k^{(r,\mu)}(\tau_j) & \text{if } \tau = [\tau_1, \cdots, \tau_\kappa]_r \in T \text{ and } \kappa \geq 2, \\ \displaystyle\sum_{\mu=0}^{M} Z^{(l,\nu)(r,\mu)} \Phi_{k+1}^{(r,\mu)}(\tau_1) & \text{if } \tau = [\tau_1]_r \in T. \end{cases} \tag{3.4}$$

The corresponding growth function is given by

$$\mathfrak{r}(\emptyset) = 0, \qquad \mathfrak{r}(\bullet_l) = 1, \qquad \mathfrak{r}(\tau = [\tau_1, \cdots, \tau_\kappa]_l) = \begin{cases} \mathfrak{r}(\tau_1) & \text{if} \quad \kappa = 1, \\ 1 + \max\limits_{j=1}^{\kappa} \mathfrak{r}(\tau_j) & \text{if} \quad \kappa \geq 2. \end{cases}$$

The function $\mathfrak{r}(\tau)$ is one plus the maximum number of ramifications along any branch of the tree.

Proof. The iteration scheme (1.4a) can be rewritten in B-series notation as

$$\begin{aligned} \sum_{\tau \in T} \alpha(\tau) \cdot \Phi_{k+1}^{(l,\nu)}(\tau) \otimes F(\tau)(x_0) &= \mathbb{1} \otimes x_0 + \\ \sum_{r=0}^{m} \sum_{\tau \in T_r} \alpha(\tau) \cdot \left(\sum_{\mu=0}^{M} Z^{(l,\nu)(r,\mu)} (\Phi_k^{(r,\mu)})_r'(\tau) \right) \otimes F(\tau)(x_0) & \\ + \sum_{r=0}^{m} \sum_{\tau_1 \in T_r} \alpha(\tau_1) \cdot \left(\sum_{\mu=0}^{M} Z^{(l,\nu)(r,\mu)} \left(\Phi_{k+1}^{(r,\mu)}(\tau_1) - \Phi_k^{(r,\mu)}(\tau_1) \right) \right) \otimes (g_r'(x_0) F(\tau_1)(x_0)), & \end{aligned} \tag{3.5}$$

where we have used (2.2). Clearly, $\Phi_{k+1}^{(l,\nu)}(\emptyset) \equiv \mathbb{1}_s$ for all $k \geq 0$ and

$$\Phi_{k+1}^{(l,\nu)}(\bullet_r) = \sum_{\mu=0}^{M} Z^{(l,\nu)(r,\mu)} \mathbb{1}_s,$$

proving the lemma for $\tau = \bullet_r = [\emptyset]_r$. Next, let $\tau = [\tau_1]_r$, where $\tau_1 \neq \emptyset$. Then $F(\tau)(x_0) = g_r'(x_0) F(\tau_1)$. Comparing equal terms on both sides of the equation, using $\alpha(\tau) = \alpha(\tau_1)$, we get

$$\Phi_{k+1}^{(l,\nu)}(\tau) = \sum_{\mu=0}^{M} Z^{(l,\nu)(r,\mu)} \left((\Phi_k^{(r,\mu)})_r'(\tau) + \Phi_{k+1}^{(r,\mu)}(\tau_1) - \Phi_k^{(r,\mu)}(\tau_1) \right).$$

Since $(\Phi_k^{(r,\mu)})_r'(\tau) = \Phi_k^{(r,\mu)}(\tau_1)$ the lemma is proved for all $\tau = [\tau_1]_r$. For $\tau = [\tau_1, \cdots, \tau_\kappa]_r$ with $\kappa \geq 2$ the last sum of (3.5) contributes nothing, thus

$$\Phi_{k+1}^{(l,\nu)}(\tau) = \sum_{\mu=0}^{M} Z^{(l,\nu)(r,\mu)} (\Phi_k^{(r,\mu)})_r'(\tau)$$

which concludes the proof of (3.4).

The second statement of the lemma is obviously true for $\tau = \emptyset$. Let τ be any tree satisfying $\mathfrak{r}(\tau) \le k+1$. Then either $\tau = [\tau_1]_l$ with $\mathfrak{r}(\tau_1) \le k+1$ or $\tau = [\tau_1, \cdots, \tau_\kappa]_l$ with $\kappa \ge 2$ and $\mathfrak{r}(\tau_i) \le k$. In the latter case, we have by the hypothesis, by (3.4) and Theorem 6 that

$$\Phi_{k+1}^{(l,\nu)}(\tau) = \sum_{\mu=0}^{M} Z^{(l,\nu)(r,\mu)} \prod_{j=1}^{\kappa} \Phi^{(r,\mu)}(\tau_j) = \Phi^{(l,\nu)}(\tau).$$

In the first case, it follows easily by induction on τ that $\Phi_{k+1}^{(l,\nu)}(\tau) = \Phi^{(l,\nu)}(\tau)$ since $\Phi_{k+1}^{(l,\nu)}(\tau) = \sum_{\mu=0}^{M} Z^{(l,\nu)(r,\mu)} \Phi_{k+1}^{(r,\mu)}(\tau_1)$. □

The full Newton iteration. In this subsection, we consider the full Newton iteration (1.4a) with

$$J_k^{(r,\mu)} = g_r'(H_k^{(r,\mu)}).$$

It follows that the B-series for $H_k^{(l,\nu)}$ and the corresponding growth function satisfy

Lemma 10 *If $H_0^{(l,\nu)} = B(\Phi_0^{(l,\nu)}, x_0; h)$ then $H_k^{(l,\nu)} = B(\Phi_k^{(l,\nu)}, x_0; h)$ with*

$$\begin{aligned}
\Phi_{k+1}^{(l,\nu)}(\emptyset) &\equiv \mathbb{1}_s, \\
\Phi_{k+1}^{(l,\nu)}(\tau) &= \sum_{\mu=0}^{M} Z^{(l,\nu)(r,\mu)} \prod_{j=1}^{\kappa} \Phi_k^{(r,\mu)}(\tau_j) \\
&\quad + \sum_{\mu=0}^{M} Z^{(l,\nu)(r,\mu)} \sum_{i=1}^{\kappa} \left(\prod_{\substack{j=1 \\ j\ne i}}^{\kappa} \Phi_k^{(r,\mu)}(\tau_j) \right) \left(\Phi_{k+1}^{(r,\mu)}(\tau_i) - \Phi_k^{(r,\mu)}(\tau_i) \right)
\end{aligned} \tag{3.6}$$

where $\tau = [\tau_1, \ldots, \tau_\kappa]_r$ and the rightmost $\prod$ is taken to be $\mathbb{1}_s$ if $\kappa = 1$. The corresponding growth function is given by

$$\mathfrak{d}(\emptyset) = 0, \qquad \mathfrak{d}(\bullet_l) = 1,$$
$$\mathfrak{d}(\tau = [\tau_1, \cdots, \tau_\kappa]_l) = \begin{cases} \max_{j=1}^{\kappa} \mathfrak{d}(\tau_j) & \text{if} \quad \gamma = 1, \\ \max_{j=1}^{\kappa} \mathfrak{d}(\tau_j) + 1 & \text{if} \quad \gamma \ge 2, \end{cases}$$

where γ is the number of subtrees in τ satisfying $\mathfrak{d}(\tau_i) = \max_{j=1}^{\kappa} \mathfrak{d}(\tau_j)$.

The function $\mathfrak{d}$ is called the doubling index of τ.

Proof. Using (2.2) and Lemma 4 the scheme (1.4a) can be written as

$$\begin{aligned}
&\sum_{\tau\in T} \alpha(\tau) \cdot \Phi_{k+1}^{(l,\nu)}(\tau) \otimes F(\tau)(x_0) = \mathbb{1}\otimes x_0 + \\
&\quad \sum_{r=0}^{m} \sum_{\tau\in T_r} \alpha(\tau) \cdot \left(\sum_{\mu=0}^{M} Z^{(l,\nu)(r,\mu)} (\Phi_k^{(r,\mu)})_r'(\tau) \right) \otimes F(\tau)(x_0) \\
&\quad + \sum_{r=0}^{m} \sum_{u\in U_{g_r}} \beta(u) \cdot \left(\sum_{\mu=0}^{M} Z^{(l,\nu)(r,\mu)} \Upsilon_k^{(r,\mu)}(u) \right) \otimes G(u)(x_0),
\end{aligned} \tag{3.7}$$

where

$$\Upsilon_k^{(r,\mu)}(u=[\tau_1,\ldots,\tau_\kappa]_{g_r}) = \sum_{i=1}^{\kappa}\left(\prod_{\substack{j=1\\ j\neq i}}^{\kappa}\Phi_k^{(r,\mu)}(\tau_j)\right)\left(\Phi_{k+1}^{(r,\mu)}(\tau_i)-\Phi_k^{(r,\mu)}(\tau_i)\right).$$

From the definition of $F(\tau)$, $G(u=[\tau_1,\cdots,\tau_\kappa]_{g_r})(x_0) = F(\tau=[\tau_1,\cdots,\tau_\kappa]_r)(x_0)$. The sum over all $u\in U_{g_r}$ can be replaced by the sum over all $\tau\in T_r$, and the result is proved. Next, we will prove that $\mathfrak{d}(\tau)$ satisfies the implication (3.1) given in Definition 7. We will do so by induction on $n(\tau)$, the number of nodes in τ. Since $\emptyset$ is the only tree satisfying $n(\tau)=0$, and $\Phi_{k+1}^{(r,\mu)}(\emptyset)=\Phi^{(r,\mu)}(\emptyset)\equiv \mathbb{1}_s$, the conclusion of (3.1) is true for all $\tau\in T$ with $n(\tau)=0$. Let $\bar{n}\geq 1$ and assume by the induction hypothesis that the conclusion of (3.1) holds for any tree satisfying $\mathfrak{d}(\tau)\leq k+1$ and $n(\tau)<\bar{n}$. We will show that $\Phi_{k+1}^{(r,\mu)}(\bar{\tau})=\Phi^{(r,\mu)}(\bar{\tau})$ for all $\bar{\tau}$ satisfying $\mathfrak{d}(\bar{\tau})\leq k+1$ and $n(\bar{\tau})\leq\bar{n}$. Let $\bar{\tau}=[\tau_1,\cdots,\tau_\kappa]_l$ where $n(\tau_j)<\bar{n}$ for $j=1,\cdots,\kappa$. Since $\mathfrak{d}(\bar{\tau})\leq k+1$ there is at most one subtree τ_j satisfying $\mathfrak{d}(\tau_j)=k+1$, let us for simplicity assume this to be the last one. Thus $\mathfrak{d}(\tau_j)\leq k$ for $j=1,\cdots,\kappa-1$ and $\mathfrak{d}(\tau_\kappa)\leq k+1$. Consequently, $\Phi_k^{(r,\mu)}(\tau_j)=\Phi^{(r,\mu)}(\tau_j)$, $j=1,\cdots,\kappa-1$ by the hypothesis of (3.1), and $\Phi_{k+1}^{(r,\mu)}(\tau_j)=\Phi^{(r,\mu)}(\tau_j)$, $j=1,\cdots,\kappa$ by the induction hypothesis. By (3.6) and Theorem 6,

$$\begin{aligned}
\Phi_{k+1}^{(l,\nu)}(\bar{\tau}) &= \sum_{\mu=0}^{M} Z^{(l,\nu)(r,\mu)}\prod_{j=1}^{\kappa}\Phi_k^{(r,\mu)}(\tau_j)\\
&\quad+\sum_{\mu=0}^{M}\sum_{i=1}^{\kappa} Z^{(l,\nu)(r,\mu)}\left(\prod_{\substack{j=1\\ j\neq i}}^{\kappa}\Phi_k^{(r,\mu)}(\tau_j)\right)\left(\Phi_{k+1}^{(r,\mu)}(\tau_i)-\Phi_k^{(r,\mu)}(\tau_i)\right)\\
&= \sum_{\mu=0}^{M} Z^{(l,\nu)(r,\mu)}\left(\prod_{j=1}^{\kappa-1}\Phi_k^{(r,\mu)}(\tau_j)\right)\Phi_k^{(r,\mu)}(\tau_\kappa)\\
&\quad+\sum_{\mu=0}^{M} Z^{(l,\nu)(r,\mu)}\left(\prod_{j=1}^{\kappa-1}\Phi^{(r,\mu)}(\tau_j)\right)\left(\Phi^{(r,\mu)}(\tau_\kappa)-\Phi_k^{(r,\mu)}(\tau_\kappa)\right)\\
&= \Phi^{(l,\nu)}(\bar{\tau}),
\end{aligned}$$

completing the induction proof. □

4. General convergence results for iterated methods. Now we will relate the results of the previous section to the order of the overall scheme. In the following, we assume that the predictors satisfy the conditions

$$\begin{aligned}
&\Phi_0^{(l,\nu)}(\tau)=\Phi^{(l,\nu)}(\tau) && \forall\tau\in T \text{ with } \mathfrak{g}(\tau)\leq\mathcal{G}_0,\\
&\Phi_0^{(l,\nu)}(\tau)\in\{\Phi^{(l,\nu)}(\tau),0\} && \forall\tau\in T \text{ with } \mathfrak{g}(\tau)\leq\hat{\mathcal{G}}_0,
\end{aligned}\tag{4.1}$$

where $\mathcal{G}_0$ and $\hat{\mathcal{G}}_0$ are chosen as large as possible. In particular, the trivial predictor satisfies $\mathcal{G}_0=0$ while $\hat{\mathcal{G}}_0=\infty$. We assume further that in analogy to (3.1) we have

$$\begin{aligned}
&\Phi_k^{(l,\nu)}(\tau)\in\{\Phi^{(l,\nu)}(\tau),0\},\quad \forall\tau\in T,\quad \mathfrak{g}(\tau)\leq k\\
&\quad\Rightarrow\qquad \Phi_{k+1}^{(l,\nu)}(\tau)\in\{\Phi^{(l,\nu)}(\tau),0\},\quad \forall\tau\in T,\quad \mathfrak{g}(\tau)\leq k+1,
\end{aligned}\tag{4.2}$$

for all $k \geq 0$. By Lemmata 8, 9, and 10 this is guaranteed for the iteration schemes considered here.

It follows from (3.1), (3.2) and (4.2) that

$$\begin{aligned} &\Phi_k(\tau) = \Phi(\tau) && \forall \tau \in T \text{ with } \mathfrak{g}'(\tau) \leq \mathcal{G}_0 + k, \\ &\Phi_k(\tau) \in \{\Phi(\tau), 0\} && \forall \tau \in T \text{ with } \mathfrak{g}'(\tau) \leq \hat{\mathcal{G}}_0 + k \end{aligned} \tag{4.3}$$

as well as

$$\begin{aligned} &\psi_{\Phi_k}(u) = \psi_{\Phi}(u) && \forall u \in U_f \text{ with } \mathfrak{g}'(u) \leq \mathcal{G}_0 + k, \\ &\psi_{\Phi_k}(\tau) \in \{\psi_{\Phi}(u), 0\} && \forall u \in U_f \text{ with } \mathfrak{g}'(u) \leq \hat{\mathcal{G}}_0 + k. \end{aligned} \tag{4.4}$$

The next step is to establish the relation between the order and the growth function of a tree. We have chosen to do so by some maximum height functions, given by

$$\begin{aligned} \mathcal{G}_T(q) &= \max_{\tau \in T} \{\mathfrak{g}'(\tau) : \rho(\tau) \leq q\}, \qquad \mathcal{G}_{T,\varphi}(q) = \max_{\tau \in T} \{\mathfrak{g}'(\tau) : \mathrm{E}\,\varphi(\tau) \neq 0, \rho(\tau) \leq q\}, \\ \mathcal{G}_{U_f}(q) &= \max_{u \in U_f} \{\mathfrak{g}'(u) : \rho(u) \leq q\}, \quad \mathcal{G}_{U_f,\psi_\varphi}(q) = \max_{u \in U_f} \{\mathfrak{g}'(u) : \mathrm{E}\,\psi_\varphi(u) \neq 0,\ \rho(u) \leq q\}. \end{aligned} \tag{4.5}$$

Note that the definition relates to the weights of the exact, not the numerical, solution. We are now ready to establish results on weak and strong convergence for the iterated solution.

Weak convergence. Let p be the weak order of the underlying scheme. The weak order of the iterated solution after k iterations is $\min(q_k, p)$ if

$$\mathrm{E}\,\psi_{\Phi_k}(u) = \mathrm{E}\,\psi_{\Phi}(u) \qquad \forall u \in U_f, \quad \rho(u) \leq q_k + \frac{1}{2}.$$

If $q_k \leq p$ we can take advantage of the fact that $\mathrm{E}\,\psi_\Phi(u) = \mathrm{E}\,\psi_\varphi(u) = 0$ for some u, and thereby relax the conditions to

$$\begin{aligned} &\psi_{\Phi_k}(u) = \psi_{\varphi}(u) && \forall u \in U_f \text{ with } \mathrm{E}\,\psi_\varphi(u) \neq 0, \\ &\psi_{\Phi_k}(u) \in \{\psi_{\varphi}(u), 0\} && \forall u \in U_f \text{ with } \mathrm{E}\,\psi_\varphi(u) = 0. \end{aligned} \tag{4.6}$$

By (4.4), this is fulfilled for all u of order $\rho(u) \leq \min(q_k, p)$ if

$$\mathcal{G}_{U_f,\Psi_\varphi}(q_k + \frac{1}{2}) \leq \mathcal{G}_0 + k \qquad \text{and} \qquad \mathcal{G}_{U_f}(q_k + \frac{1}{2}) \leq \hat{\mathcal{G}}_0 + k.$$

The results can then be summarized in the following Theorem:

Theorem 11 *The iterated method is of weak order $q_k \leq p$ after*

$$\max\{\mathcal{G}_{U_f,\psi_\varphi}(q_k + \frac{1}{2}) - \mathcal{G}_0, \mathcal{G}_{U_f}(q_k + \frac{1}{2}) - \hat{\mathcal{G}}_0\}$$

iterations.

Strong convergence. The strong convergence case can be treated similarly. Let p now be the mean square order of the underlying method. The iterated solution is of order $\min(p, q_k)$ if for all $q_k \leq p$

$$\begin{aligned} \Phi_k(\tau) &= \Phi(\tau) \quad \forall \tau \in T \text{ with } \rho(\tau) \leq q_k, \\ \mathrm{E}\,\Phi_k(\tau) &= \mathrm{E}\,\Phi(\tau) \quad \forall \tau \in T \text{ with } \rho(\tau) = q_k + \frac{1}{2}. \end{aligned} \tag{4.7}$$

According to (4.3) these are satisfied if all the conditions

$$\begin{aligned} \mathcal{G}_T(q_k) &\leq \mathcal{G}_0 + k, \\ \mathcal{G}_T(q_k + \frac{1}{2}) &\leq \hat{\mathcal{G}}_0 + k, \\ \mathcal{G}_{T,\varphi}(q_k + \frac{1}{2}) &\leq \mathcal{G}_0 + k \end{aligned}$$

are satisfied. We can summarize this by:

Theorem 12 *The iterated method is of mean square order $q_k \leq p$ after*

$$\max\left\{\max\left\{\mathcal{G}_T(q_k), \mathcal{G}_{T,\varphi}(q_k + \frac{1}{2})\right\} - \mathcal{G}_0, \mathcal{G}_T(q_k + \frac{1}{2}) - \hat{\mathcal{G}}_0\right\}$$

iterations.

5. Growth functions and order. In this section we will discuss the relation between the order of trees and the growth functions defined in section 3. Let us start with the lemma:

Lemma 13 *For $k \geq 1$,*

$$\begin{aligned} \mathfrak{h}'(\tau) = k \quad &\Rightarrow \quad \rho(\tau) \geq \frac{k}{2} + \frac{1}{2}, \\ \mathfrak{r}'(\tau) = k \quad &\Rightarrow \quad \rho(\tau) \geq k, \\ \mathfrak{d}'(\tau) = k \quad &\Rightarrow \quad \rho(\tau) \geq 2^{k-1}. \end{aligned}$$

The same result is valid for $\mathfrak{h}'(u)$, $\mathfrak{r}'(u)$ and $\mathfrak{g}'(u)$.

Proof. Let $\mathcal{T}_{\mathfrak{h},k}$, $\mathcal{T}_{\mathfrak{r},k}$ and $\mathcal{T}_{\mathfrak{d},k}$ be sets of trees of minimal order satisfying $\mathfrak{h}(\tau) = k$ $\forall \tau \in \mathcal{T}_{\mathfrak{h},k}$, $\mathfrak{r}(\tau) = k$ $\forall \tau \in \mathcal{T}_{\mathfrak{r},k}$ and $\mathfrak{d}(\tau) = k$ $\forall \tau \in \mathcal{T}_{\mathfrak{d},k}$, and denote this minimal order by $\rho_{\mathfrak{h},k}$, $\rho_{\mathfrak{r},k}$ and $\rho_{\mathfrak{d},k}$. Minimal order trees are build up only by stochastic nodes. It follows immediately that $\mathcal{T}_{\mathfrak{h},1} = \mathcal{T}_{\mathfrak{r},1} = \mathcal{T}_{\mathfrak{d},1} = \{\bullet_l : \ l \geq 1\}$. Since $\rho(\bullet_l) = 1/2$ for $l \geq 1$, the results are proved for $k = 1$. It is easy to show by induction on k that

$$\begin{aligned} &\mathcal{T}_{\mathfrak{h},k} = \{[\tau]_l : \ \tau \in \mathcal{T}_{\mathfrak{h},k-1}, \ l \geq 1\}, && \rho_{\mathfrak{h},k} = \rho_{\mathfrak{h},k-1} + \frac{1}{2} = \frac{k}{2}, \\ &\mathcal{T}_{\mathfrak{r},k} = \{[\bullet_{l_1}, \tau]_{l_2} : \ \tau \in \mathcal{T}_{\mathfrak{r},k-1}, \ l_1, l_2 \geq 1\}, && \rho_{\mathfrak{r},k} = \rho_{\mathfrak{r},k-1} + 1 = k - \frac{1}{2}, \\ &\mathcal{T}_{\mathfrak{d},k} = \{[\tau_1, \tau_2]_l : \ \tau_1, \tau_2 \in \mathcal{T}_{\mathfrak{d},k-1}, \ l \geq 1\}, && \rho_{\mathfrak{d},k} = 2\rho_{\mathfrak{d},k-1} + \frac{1}{2} = 2^{k-1} - \frac{1}{2}. \end{aligned} \tag{5.1}$$

For each $\mathfrak{g}$ being either $\mathfrak{h}$, $\mathfrak{r}$ or $\mathfrak{d}$, the minimal order trees satisfying $\mathfrak{g}'(\tau'_{\mathfrak{g},k}) = k$, $\mathfrak{g}'(u_{\mathfrak{g},k}) = k$ are $\tau'_{\mathfrak{g},k} = [\tau_{\mathfrak{g},k}]_l$ with $\tau_{\mathfrak{g},k} \in \mathcal{T}_{\mathfrak{g},k}$ and $l \geq 1$, and $u_{\mathfrak{g},k} = [\tau'_{\mathfrak{g},k}]_f$. Both are of order $\rho(\tau_{\mathfrak{g},k}) + 1/2$. □

Let $\mathcal{G}_T(q)$ and $\mathcal{G}_{U_f}(q)$ be defined by (4.5). Then it holds

Corollary 14 *For $q \geq \frac{1}{2}$ we have*

$$\mathcal{G}_T(q) = \mathcal{G}_{U_f}(q) = \begin{cases} 2q-1 & \textit{for simple iterations,} \\ \lfloor q \rfloor & \textit{for modified Newton iterations,} \\ \lfloor \log_2(q) \rfloor + 1 & \textit{for full Newton iterations.} \end{cases}$$

Proof. The minimal order trees are also the maximum height/ramification number/doubling index trees, in the sense that as long as $\rho(\tau'_{\mathfrak{g},k}) \leq q < \rho(\tau'_{\mathfrak{g},k+1})$ there are no trees of order q for which the growth function can exceed k. □

Let $T^S \subset T$ and $U_f^S \subset U_f$ be the set of trees with an even number of each kind of stochastic nodes. Further, let $T^I \subset T_0$ and $U_f^I \subset U_f$ be the set of trees constructed from the root ($\bullet_0$ or $\bullet_f$), by a finite number of steps of the form
i) add one deterministic node, or
ii) add two equal stochastic nodes, neither of them being a father of the other.
Clearly $T^I \subset T^S$ and $U_f^I \subset U_f^S$. From [5, 25] we have

$$\begin{aligned} \mathrm{E}\,\varphi(\tau) = 0 \quad \text{if} \quad \tau \notin \begin{cases} T^S & \text{in the Stratonovich case,} \\ T^I & \text{in the Itô case,} \end{cases} \\ \mathrm{E}\,\psi_\varphi(u) = 0 \quad \text{if} \quad u \notin \begin{cases} U_f^S & \text{in the Stratonovich case,} \\ U_f^I & \text{in the Itô case.} \end{cases} \end{aligned} \tag{5.2}$$

Considering only trees for which $\mathrm{E}\,\varphi$ or $\mathrm{E}\,\psi_\varphi$ are different from zero, we get

Lemma 15 *For $k \geq 1$,*

$$\begin{aligned} \mathfrak{h}'(\tau) = k \quad &\Rightarrow \quad \rho(\tau) \geq \begin{cases} \ulcorner \frac{k+1}{2} \urcorner & \textit{if} \quad \tau \in T^S, \\ k+1 & \textit{if} \quad \tau \in T^I, \end{cases} \\ \mathfrak{r}'(\tau) = k \quad &\Rightarrow \quad \rho(\tau) \geq \begin{cases} k & \textit{if} \quad \tau \in T^S, \\ k+1 & \textit{if} \quad \tau \in T^I, \end{cases} \\ \mathfrak{d}'(\tau) = k \quad &\Rightarrow \quad \rho(\tau) \geq \begin{cases} 2^{k-1} & \textit{if} \quad \tau \in T^S, \\ 2^{k-1}+1 & \textit{if} \quad \tau \in T^I. \end{cases} \end{aligned}$$

This result is also valid for $\mathfrak{h}'(u)$, $\mathfrak{r}'(u)$ and $\mathfrak{g}'(u)$, with $T^{\cdot}$ replaced by $U_f^{\cdot}$.

Proof. In the Stratonovich case, we only consider trees of integer order, which immediately gives the results. In the Itô case, let $\tau_{\mathfrak{g},k}$, $\tau'_{\mathfrak{g},k}$ be the minimal order trees used in the proof of Lemma 13. Unfortunately $\tau'_{\mathfrak{g},k}$ has a stochastic root, so $\tau'_{\mathfrak{g},k} \notin T^I$, and there are no trees $\tau \in T^I$ of order $\rho(\tau_{\mathfrak{g},k}) + 1/2$ satisfying $\mathfrak{g}'(\tau) = k$. When $\mathfrak{g}$ is either $\mathfrak{r}$ or $\mathfrak{d}$ then the tree $[\tau_{\mathfrak{g}}, \bullet_l]_0 \in T^I$ if all the stochastic nodes are of color $l \geq 1$. The order of this tree is $\rho(\tau_{\mathfrak{g}}) + 3/2$, proving the result for $\mathfrak{r}'(\tau)$ and $\mathfrak{d}'(\tau)$. Let $\hat{\tau}'_{\mathfrak{h},k} \in T^I$ be a tree of minimal order satisfying $\mathfrak{h}'(\hat{\tau}'_{\mathfrak{h},k}) = k$. Clearly, $\hat{\tau}'_{\mathfrak{h},1}$ can be either $[\bullet_0]_0$ or $[\bullet_l, \bullet_l]_0$ with $l \geq 1$, both of order 2. From the construction of trees in T^I it is clear that the height of the tree can only be increased by one for each order, thus $\rho(\hat{\tau}'_{\mathfrak{h},k}) = k+1$. The result for U_f^I follows immediately. □

Let $\mathcal{G}_{T,\varphi}(q)$ and $\mathcal{G}_{U_f,\psi_\varphi}(q)$ be given by (4.5). Then the analogue of Corollary 14 is

	Stratonovich			Itô					
	strong/weak appr.			weak appr.			strong appr.		
p	simple	mod.	full	simple	mod.	full	simple	mod.	full
$\frac{1}{2}$	1	1	1	0	0	0	0	0	0
1	1	1	1	0	0	0	1	1	1
$1\frac{1}{2}$	3	2	2	1	1	1	2	1	1
2	3	2	2	1	1	1	3	2	2
$2\frac{1}{2}$	5	3	2	2	2	1	4	2	2
3	5	3	2	2	2	1	5	3	2

TABLE 5.1

Number of iterations needed to achieve order p when using the simple, modified or full Newton iteration scheme in the Itô or Stratonovich case for strong or weak approximation.

Corollary 16 *For $q \geq \frac{1}{2}$ we have in the Stratonovich case*

$$\mathcal{G}_{T,\varphi}(q) = \mathcal{G}_{U_f,\psi_\varphi}(q) = \begin{cases} \max\{0, 2\lfloor q \rfloor - 1\} & \textit{for simple iterations,} \\ \lfloor q \rfloor & \textit{for modified Newton iterations,} \\ \lfloor \log_2(q) \rfloor + 1 & \textit{for full Newton iterations} \end{cases}$$

and in the Itô case

$$\mathcal{G}_{T,\varphi}(q) = \mathcal{G}_{U_f,\psi_\varphi}(q) = \begin{cases} \max\{0, \lfloor q \rfloor - 1\} & \textit{for simple iterations,} \\ \max\{0, \lfloor q \rfloor - 1\} & \textit{for modified Newton iterations,} \\ \max\{0, \lfloor \log_2(q) \rfloor\} & \textit{for full Newton iterations.} \end{cases}$$

For the trivial predictor, Table 5.1 gives the number of iterations needed to achieve a certain order of convergence. The results concerning the Stratonovich case when considering strong approximation and using the simple iteration scheme where already obtained by Burrage and Tian [3] analyzing predictor corrector methods.

6. Numerical examples. As first example, we apply the drift implicit strong order 1.5 scheme due to Kloeden and Platen [16], implemented as stiffly accurate SRK scheme with six stages and denoted by IPS, to a nonlinear Itô SDE in order to analyze its order of convergence in dependence on the kind and number of iterations. Therefore, the solution is approximated with step sizes $2^{-8}, \ldots, 2^{-12}$ and the sample average of $M = 20000$ independent simulated realisations of the absolute error is calculated in order to estimate the expectation.

As test equation, we consider the non-linear SDE [16, 21]

$$dX(t) = \left(\tfrac{1}{2}X(t) + \sqrt{X(t)^2 + 1}\right) dt + \sqrt{X(t)^2 + 1}\, dW(t), \qquad X(0) = 0, \qquad (6.1)$$

on the time interval $I = [0, 1]$ with the solution $X(t) = \sinh(t + W(t))$.

The results at time $t = 1$ without iteration, with one and two simple iterations and one modified Newton iteration step are presented in Figure 6.1, where the orders of convergence correspond to the slope of the regression lines. As predicted by Table 5.1 we observe strong order 0.5 without iteration, strong order one for one simple iteration and strong order 1.5 for two simple or one modified Newton iteration.

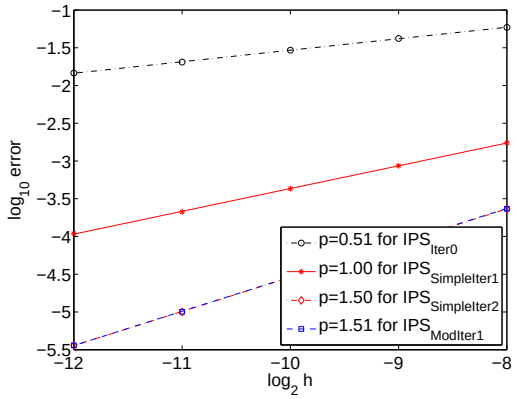


FIGURE 6.1. *Error of IPS applied to (6.1) without iteration, with one or two simple iterations and with one modified Newton iteration (the results for two simple iterations and one modified Newton iteration nearly coincide).*

As second example, we apply the implicit strong order 1.5 method IRK3 given by[1]

$$y_{n+1} = y_n + \sum_{i=1}^{3} \alpha_i g_0(Y_i) + \sum_{i=1}^{3} \left(J_1 \gamma_i^{(1)} + \frac{J_{10}}{h} \gamma_i^{(2)} \right) g_1(Y_i),$$

$$Y_i = y_n + \sum_{j=1}^{3} a_{ij} g_0(Y_j) + \sum_{j=1}^{3} \left(J_1 b_{ij}^{(1)} + \frac{J_{10}}{h} b_{ij}^{(2)} \right) g_1(Y_i)$$

with coefficients

$$\alpha^\top = (3.8987207850703709e{-}01, 2.5303195561124980e{-}01, 3.5709596588171316e{-}01),$$

$$\gamma^{(1)\top} = (-5.1071092434339060e{-}01, 5.8945603509591227e{-}01, 9.2125488924747834e{-}01),$$

$$\gamma^{(2)\top} = (1.4755439991699316, 0, -1.4755439991699316),$$

$$A = \begin{pmatrix} -2.8952337674962747e-1 & 7.4418828501997702e-1 & -2.8516957799526149e-1 \\ 1.0980329196623682e+0 & -5.9356838134753187e-1 & 1.4769092465852457e-2 \\ 4.3056898319396864e-1 & -4.2336418516823161e-2 & 4.5897891676832508e-1 \end{pmatrix},$$

$$B_1 = \begin{pmatrix} -3.7888032376529462e-2 & 3.1212706165530735e-1 & -5.8972962840290177e-1 \\ -5.5669836775129566e-1 & 7.5799500096196037e-1 & -5.4160073101557960e-1 \\ -3.9406687933866946e-1 & 5.4561698120971203e-1 & 4.3403138596161156e-1 \end{pmatrix},$$

$$B_2 = \begin{pmatrix} 5.4900483755770564e-2 & 3.7642537111125235e-1 & 3.2079615732320838e-1 \\ 1.8393033906369787e+0 & -1.4094467152219967e+0 & 1.3018964952626004e+0 \\ 8.0702249594600106e-1 & 3.7642537111125235e-1 & -4.3132585486702213e-1 \end{pmatrix}$$

[1]Note that this method constructed here does not satisfy the strong order 1.5 order conditions of [2], which seem to be flawed.

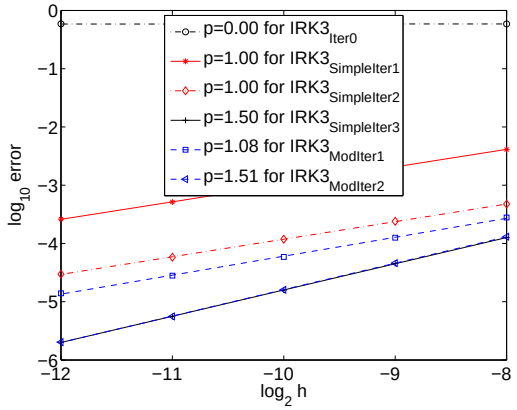


FIGURE 6.2. *Error of IRK3 applied to the Stratonovich version of (6.1) without iteration, with one, two or three simple iterations and with one or two modified Newton iterations (the results for three simple iterations and two modified Newton iterations nearly coincide).*

to the corresponding Stratonovich version of (6.1). Again, the solution is approximated with step sizes $2^{-8}, \dots, 2^{-12}$ and M=20000 simulations are performed. The results at time $t = 1$ without iteration, with one, two and three simple iterations and one and two modified Newton iteration steps are presented in Figure 6.2. As predicted by Table 5.1 we observe no convergence without iteration, strong order one for one or two simple iterations or one modified Newton iteration and strong order 1.5 in the case of three simple iterations or two modified Newton iterations.

Finally, we apply the drift implicit weak order two SRK scheme DDIRDI5 [8] to SDE (6.1). Here, we choose as functional $f(x) = p(\operatorname{arsinh}(x))$, where $p(z) = z^3 - 6z^2 + 8z$ is a polynomial. Then the expectation of the solution can be calculated as

$$\mathrm{E}(f(X(t))) = t^3 - 3t^2 + 2t \ . \tag{6.2}$$

The solution $\mathrm{E}(f(X(t)))$ is approximated with step sizes $2^{-2}, \dots, 2^{-5}$ and $M = 5 \cdot 10^8$ simulations are performed in order to determine the systematic error of DDIRDI5 at time $t = 1$. The results without iteration, with one simple iteration and with one modified Newton iteration steps are presented in Figure 6.3. According to Table 5.1 we expect approximation order one for zero iterations and order two in the other cases, which is approved by Figure 6.3.

7. Conclusion. For stochastic Runge-Kutta methods that use an iterative scheme to compute their internal stage values, we derived convergence results based on the order of the underlying Runge-Kutta method, the choice of the iteration method, the predictor and the number of iterations. This was done by developing a unifying approach for the construction of stochastic B-series, which is valid both for Itô- and Stratonovich-SDEs and can be used both for weak and strong convergence. We expect this to be useful also for the further development and analysis of stochastic Rung-Kutta type methods.

REFERENCES

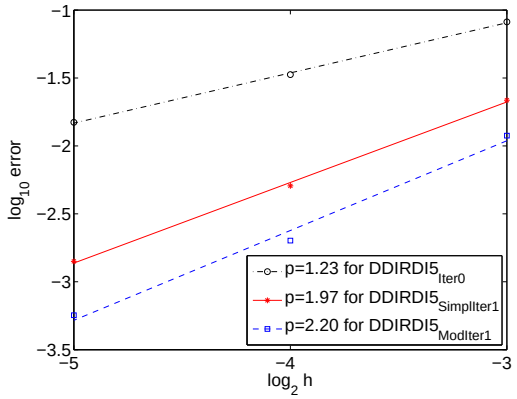


FIGURE 6.3. *Error of DDIRDI5 applied to (6.1) without iteration, with one simple iteration and with one modified Newton iteration.*

[1] K. BURRAGE AND P. M. BURRAGE, *High strong order explicit Runge-Kutta methods for stochastic ordinary differential equations*, Appl. Numer. Math., 22 (1996), pp. 81–101. Special issue celebrating the centenary of Runge-Kutta methods.

[2] ———, *Order conditions of stochastic Runge-Kutta methods by B-series*, SIAM J. Numer. Anal., 38 (2000), pp. 1626–1646 (electronic).

[3] K. BURRAGE AND T. TIAN, *Predictor-corrector methods of Runge-Kutta type for stochastic differential equations*, SIAM J. Numer. Anal., 40 (2002), pp. 1516–1537 (electronic).

[4] ———, *Implicit stochastic Runge-Kutta methods for stochastic differential equations*, BIT, 44 (2004), pp. 21–39.

[5] P. M. BURRAGE, *Runge-Kutta methods for stochastic differential equations*, PhD thesis, The University of Queensland, Brisbane, 1999.

[6] J. C. BUTCHER, *Coefficients for the study of Runge-Kutta integration processes*, J. Austral. Math. Soc., 3 (1963), pp. 185–201.

[7] K. DEBRABANT AND A. RÖSSLER, *Classification of stochastic Runge-Kutta methods for the weak approximation of stochastic differential equations*, Math. Comput. Simulation. in press, doi:10.1016/j.matcom.2007.04.016.

[8] K. DEBRABANT AND A. RÖSSLER, *Diagonally drift-implicit Runge-Kutta methods of weak order one and two for Itô SDEs and stability analysis*, Appl. Numer. Math. in press.

[9] K. DEBRABANT AND A. RÖSSLER, *Families of efficient second order Runge-Kutta methods for the weak approximation of Itô stochastic differential equations*, Appl. Numer. Math. in press.

[10] T. C. GARD, *Introduction to stochastic differential equations*, vol. 114 of Monographs and Textbooks in Pure and Applied Mathematics, Marcel Dekker Inc., New York, 1988.

[11] D. B. HERNÁNDEZ AND R. SPIGLER, *Convergence and stability of implicit Runge-Kutta methods for systems with multiplicative noise*, BIT, 33 (1993), pp. 654–669.

[12] D. J. HIGHAM, *Mean-square and asymptotic stability of the stochastic theta method*, SIAM J. Numer. Anal., 38 (2000), pp. 753–769 (electronic).

[13] K. R. JACKSON, A. KVÆRNØ, AND S. P. NØRSETT, *The use of Butcher series in the analysis of Newton-like iterations in Runge-Kutta formulas*, Appl. Numer. Math., 15 (1994), pp. 341–356. International Conference on Scientific Computation and Differential Equations (Auckland, 1993).

[14] ———, *An analysis of the order of Runge-Kutta methods that use an iterative scheme to compute their internal stage values*, BIT, 36 (1996), pp. 713–765.

[15] I. KARATZAS AND S. E. SHREVE, *Brownian motion and stochastic calculus*, vol. 113 of Graduate Texts in Mathematics, Springer-Verlag, New York, second ed., 1991.

[16] P. E. KLOEDEN AND E. PLATEN, *Numerical solution of stochastic differential equations*, vol. 21 of Applications of Mathematics, Springer-Verlag, Berlin, 2 ed., 1999.

[17] Y. KOMORI, *Weak first- or second-order implicit Runge-Kutta methods for stochastic differential equations with a scalar Wiener process*, J. Comput. Appl. Math., In press.

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

[19] ———, *Weak second-order stochastic Runge-Kutta methods for non-commutative stochastic differential equations*, J. Comput. Appl. Math., 206 (2007), pp. 158–173.
[20] Y. Komori, T. Mitsui, and H. Sugiura, *Rooted tree analysis of the order conditions of ROW-type scheme for stochastic differential equations*, BIT, 37 (1997), pp. 43–66.
[21] V. Mackevičius and J. Navikas, *Second order weak Runge-Kutta type methods of Itô equations*, Math. Comput. Simulation, 57 (2001), pp. 29–34.
[22] G. N. Milstein, *Numerical integration of stochastic differential equations*, vol. 313 of Mathematics and its Applications, Kluwer Academic Publishers Group, Dordrecht, 1995. Translated and revised from the 1988 Russian original.
[23] S. P. Nørsett and A. Wolfbrandt, *Order conditions for Rosenbrock type methods*, Numer. Math., 32 (1979), pp. 1–15.
[24] W. P. Petersen, *A general implicit splitting for stabilizing numerical simulations of Itô stochastic differential equations*, SIAM J. Numer. Anal., 35 (1998), pp. 1439–1451 (electronic).
[25] A. Rössler, *Stochastic Taylor expansions for the expectation of functionals of diffusion processes*, Stochastic Anal. Appl., 22 (2004), pp. 1553–1576.
[26] A. Rössler, *Rooted tree analysis for order conditions of stochastic Runge-Kutta methods for the weak approximation of stochastic differential equations*, Stoch. Anal. Appl., 24 (2006), pp. 97–134.
[27] A. Rössler, *Runge-Kutta methods for Itô stochastic differential equations with scalar noise*, BIT, 46 (2006), pp. 97–110.
[28] A Rössler, *Second order Runge-Kutta methods for Itô stochastic differential equations*, Tech. Report Preprint No. 2479, TU Darmstadt, 2006.
[29] ———, *Second order Runge-Kutta methods for Stratonovich stochastic differential equations*, BIT, 47 (2007), pp. 657–680.
[30] A. Tocino and J. Vigo-Aguiar, *Weak second order conditions for stochastic Runge-Kutta methods*, SIAM J. Sci. Comput., 24 (2002), pp. 507–523 (electronic).