Cheap arbitrary high order methods for single integrand SDEs

Debrabant, Kristian; Kværnø, Anne

Published in:
Bit (Lisse)

DOI:
10.1007/s10543-016-0619-8

Publication date:
2017

Document version
Early version, also known as pre-print

Citation for pulished version (APA):
Debrabant, K., & Kværnø, A. (2017). Cheap arbitrary high order methods for single integrand SDEs. Bit (Lisse), 57(1), 153-168. https://doi.org/10.1007/s10543-016-0619-8

To appear in BIT Numerical Mathematics
DOI: 10.1007/s10543-016-0619-8

Cheap arbitrary high order methods for single integrand SDEs

Kristian Debrabant · **Anne Kværnø**

Abstract For a particular class of Stratonovich SDE problems, here denoted as single integrand SDEs, we prove that by applying a deterministic Runge–Kutta method of order p_d we obtain methods converging in the mean-square and weak sense with order $\lfloor p_d/2 \rfloor$. The reason is that the B-series of the exact solution and numerical approximation are, due to the single integrand and the usual rules of calculus holding for Stratonovich integration, similar to the ODE case. The only difference is that integration with respect to time is replaced by integration with respect to the measure induced by the single integrand SDE.

Keywords Stochastic differential equation · Runge–Kutta methods · single integrand SDEs · B-series

Mathematics Subject Classification (2010) MSC 65C30 · MSC 60H35 · MSC 65C20

1 Introduction

In this paper we consider a particular class of Stratonovich stochastic differential equations (SDEs), single integrand SDEs, given by

$$\mathrm{d}X = \lambda f(X)\,\mathrm{d}t + \sigma f(X) \circ \mathrm{d}W, \qquad X(t_0) = x_0 \tag{1.1}$$

where $W(t)$ is a Wiener process, $\lambda \in \{0,1\}$ and $\sigma \in \mathbb{R}$ is a given constant. We assume the coefficient $f : \mathbb{R}^d \to \mathbb{R}^d$ to be differentiable and that f and $f'f$ satisfy a

Kristian Debrabant
Department of Mathematics and Computer Science, University of Southern Denmark, 5230 Odense M, Denmark
E-mail: debrabant@imada.sdu.dk

Anne Kværnø
Department of Mathematical Sciences, Norwegian University of Science and Technology, 7491 Trondheim, Norway
E-mail: anne.kvarno@math.ntnu.no

Lipschitz condition such that there exists a unique solution [12]. The case $\lambda = 0$ covers Stratonovich SDEs without drift term. For $\lambda = 1$, this class of methods arises frequently in applications, especially when modelling phenomena by ordinary differential equations and then introducing multiplicative random fluctuations of uncertainties in time.

All results in this paper hold as well for single integrand SDEs with multidimensional Wiener process,

$$\mathrm{d}X = \lambda f(X)\,\mathrm{d}t + \sum_{i=1}^{m} \sigma_i f(X) \circ \mathrm{d}W_i, \qquad X(t_0) = x_0, \tag{1.2}$$

as this case can be reduced to (1.1) using $\sigma := \sqrt{\sum_{i=1}^m \sigma_i^2}$ and the Wiener process $W := \frac{1}{\sigma}\sum_{i=1}^m \sigma_i W_i$.

Some well known examples for single integrand SDEs are the SDE describing fatigue cracking [12, 20], the Kubo oscillator [6, 17], the stochastic Van der Pol equation [21] and certain stochastic Hamiltonian problems, see [7].

We are interested in solving (1.1) on the interval $I = [t_0, T]$. The equation can also be written in integral form as

$$X(t) = x_0 + \int_{t_0}^{t} f(X(s)) \circ \mathrm{d}\mu(s), \qquad \mu(s) := \lambda s + \sigma W(s). \tag{1.3}$$

Let a discretization $I^{\tilde{h}} := \{t_0, t_1, \ldots, t_N\}$ 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 $n = 0, 1, \ldots, N-1$ and maximal step size $\tilde{h} := \max_{n=1}^{N-1} h_n$ be given.

The SDE is solved by an s-stage stochastic Runge–Kutta method defined by

$$H_i = Y_n + \Delta_{t_n,h_n}\mu \sum_{j=1}^{s} a_{ij} f(H_j), \tag{1.4a}$$

$$Y_{n+1} = Y_n + \Delta_{t_n,h_n}\mu \sum_{i=1}^{s} b_i f(H_i). \tag{1.4b}$$

Here, $\Delta_{t_n,h_n}\mu = \mu_{t_n}(h_n)$ with

$$\mu_t(s) := \mu(t+s) - \mu(t) = \lambda s + \sigma(W(t+s) - W(t)),$$

and typically, the coefficients a_{ij} and b_i will be those of a known Runge–Kutta method for ordinary differential equations. For a general SDE, this simple generalization will result in a strong order 1 method, at the best [1, 2]. But for the single integrand SDEs the situation is far better, as stated in the main result of this paper:

Theorem 1.1 *The Runge–Kutta method* (1.4) *of deterministic order p_d is of mean square as well as weak order $p_\mu = \lfloor p_d/2 \rfloor$, under the conditions on f specified in Assumption 4.1. For weak convergence, it suffices that $\Delta_{t,h}\mu$ is chosen such that at least the first $2p_\mu + 1$ moments coincide with those of $\mu_t(h)$, and all the others are in $\mathcal{O}(h^{p_\mu+1})$.*

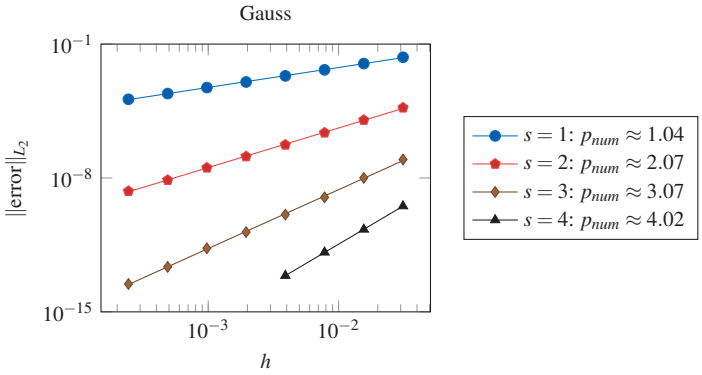

Fig. 1.1 The mean square error of Gauss methods applied to (1.5).

Here and in the following, the $\mathcal{O}$ notation refers to the absolute value and $h \to 0$, and for $x \in \mathbb{R}$, $\lfloor x \rfloor$ denotes the largest integer not larger than x. Theorem 1.1 will be proved in Section 4. Before going into the details, let us justify the result by a simple numerical experiment:

Example 1.1 We apply the s-stage Gauss method of deterministic order $2s$ [5] to the SDE [12]

$$\mathrm{d}X = \sqrt{1+X^2}\,\mathrm{d}t + \sigma\sqrt{1+X^2} \circ \mathrm{d}W, \qquad X(0) = 0 \tag{1.5}$$

with the exact solution $X(t) = \sinh(t + \sigma W(t))$, with $\sigma = 0.8$. The solution is approximated on the interval $[0,1]$ with step sizes $2^{-12} - 2^{-5}$ and the sample average of $M = 10,000$ independent simulated realizations of the absolute error is calculated in order to estimate the expectation. The results at $t = 1$ are presented in Figure 1.1.

The outline of this paper is as follows: In section 2 some results on convergence and consistency are recalled. In section 3, we will show that the B-series of the exact and the numerical solution are exactly as in the ODE case [5], with the exception that integration is now performed with respect to μ instead of h. This is due to the following lemma:

Lemma 1.1 *For all $h \geq 0$ it holds*

$$\int_0^h {\mu_t}^k \circ \mathrm{d}\mu_t = \frac{1}{k+1}\mu_t(h)^{k+1}.$$

To deduce the mean square and the weak order of the approximations, the following lemma is pivotal:

Lemma 1.2 *For $n \in \mathbb{N}$ it holds that* $\mathrm{E}\,\mu_t(h)^n = \begin{cases} \mathcal{O}(h^{\frac{n}{2}}) & : \text{if } n \text{ is even}, \\ \mathcal{O}(h^{\frac{n+1}{2}}) & : \text{if } n \text{ is odd}. \end{cases}$

The proofs of Theorem 1.1 and Lemmas 1.1 and 1.2 are given in section 4. Finally, more numerical experiments justifying the theoretical results are given in section 5.

2 Convergence and consistency

Here we will give the definitions of 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 [12] and $I^{\tilde{h}}$ be the discretized time interval defined above.

Definition 2.1 A time discrete approximation $Y = (Y(t))_{t\in I^{\tilde{h}}}$ converges weakly with order p to X at time $t \in I^{\tilde{h}}$ as the maximum step size $\tilde{h} \to 0$ if for each $g \in C_P^{2(p+1)}(\mathbb{R}^d,\mathbb{R})$ there exist a constant C_g and a finite $\delta_0 > 0$ such that

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

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

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

$$le_g(h;t,x) = \mathrm{E}\left(g(Y(t+h)) - g(X(t+h))|Y(t) = X(t) = x\right).$$

The following theorem due to Milstein [16], which holds also in the case of general one step methods, shows that, as in the deterministic case, consistency implies convergence:

Theorem 2.1 *Suppose the following conditions hold:*

- *The integrand f of* (1.3) *is differentiable, and f and $f'f$ satisfy a Lipschitz condition and belong to $C_P^{2(p+1)}(\mathbb{R}^d,\mathbb{R}^d)$.*
- *For sufficiently large r (see, e.g., [16] for details) the moments $\mathrm{E}(\|Y(t_n)\|^{2r})$ exist for $t_n \in I^{\tilde{h}}$ and are uniformly bounded with respect to N and $n = 0,1,\ldots,N$.*
- *For all $g \in C_P^{2(p+1)}(\mathbb{R}^d,\mathbb{R})$ there exists a $K \in C_P^0(\mathbb{R}^d,\mathbb{R})$ such that*

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

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

Then the method (1.4) *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 2.2 A time discrete approximation $Y = (Y(t))_{t\in I^{\tilde{h}}}$ converges strongly respectively in the mean square with order p to X at time $t \in I^{\tilde{h}}$ as the maximum step size $\tilde{h} \to 0$ if there exist a constant C and a finite $\delta_0 > 0$ such that

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

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

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

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

so 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.

$$\begin{aligned} le^m(h;t,x) &= \mathrm{E}\left(Y(t+h)-X(t+h)|Y(t)=X(t)=x\right),\\ le^{ms}(h;t,x) &= \sqrt{\mathrm{E}\left((Y(t+h)-X(t+h))^2|Y(t)=X(t)=x\right)}. \end{aligned}$$

The following theorem due to Milstein [16], 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 2.2 *Suppose the following conditions hold:*

- *The integrand f of* (1.3) *is differentiable, and f and $f'f$ satisfy a Lipschitz condition.*
- *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 $t, t+h \in I$, 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 (1.4) is convergent of order p in the sense of mean square approximation.

3 B-series and rooted trees

In order to apply Theorems 2.1 and 2.2 we will now use B–series and rooted tree theory to study the order of the local errors of the method (1.4). B-series for deterministic ODEs were introduced by Butcher [4]. Today such series appear as a fundamental tool to do local error analysis on a wide range of problems. B-series for SDEs and their numerical solution by stochastic Runge–Kutta methods have been developed by Burrage and Burrage [2, 3] to study strong convergence in the Stratonovich case, by Komori, Mitsui and Sugiura [14] and Komori [13] to study weak convergence in the Stratonovich case and by Rößler [18, 19] 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. A uniform and self-contained theory for the construction of stochastic B-series for the exact solution of

SDEs and its numerical approximation by stochastic Runge–Kutta methods is given in [8]. Based on the notation used there, we will now derive the B-series for the exact solution and numerical approximation of single-integrand SDEs. Due to the single integrand we will, similar to the ODE case [5], only need non-colored trees in the expansion of the solution.

Definition 3.1 (Trees) The set of rooted trees T is recursively defined as follows:

a) The empty tree $\emptyset$ and the graph $\bullet = [\emptyset]$ with only one vertex belong to T.

Let $\tau = [\tau_1, \tau_2, \ldots, \tau_\kappa]$ be the tree formed by joining the subtrees $\tau_1, \tau_2, \ldots, \tau_\kappa$ each by a single branch to a common root.

b) If $\tau_1, \tau_2, \ldots, \tau_\kappa \in T$ then $\tau = [\tau_1, \tau_2, \ldots, \tau_\kappa] \in T$.

Definition 3.2 (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) $F(\emptyset)(x) = x$,

b) $F(\bullet_l)(x) = f(x)$,

c) If $\tau = [\tau_1, \tau_2, \ldots, \tau_\kappa] \in T \setminus \{\emptyset\}$ then

$$F(\tau)(x) = f^{(\kappa)}(x)\big(F(\tau_1)(x), F(\tau_2)(x), \ldots, F(\tau_\kappa)(x)\big)$$

where $x \in \mathbb{R}^d$.

Definition 3.3 (B-series) Consider a family $\{\phi(\tau)\}_{\tau \in T}$ of random variables satisfying

$$\phi(\emptyset) \equiv 1 \text{ and } \phi(\tau)(0) = 0, \quad \forall \tau \in T \setminus \{\emptyset\}.$$

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

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

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

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

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

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

Lemma 3.1 ([8]) *If $Y(t+h) = B(\phi, x; h)$ is some B-series and $g \in C^\infty(\mathbb{R}^d, \mathbb{R}^{\hat{d}})$ then $g(Y(t+h))$ can be written as a formal series of the form*

$$g(Y(t+h)) = \sum_{u \in U_g} \beta(u) \cdot \psi_\phi(u)(h) \cdot G(u)(x), \tag{3.1}$$

where U_g is a set of trees derived from T, by

a) $[\emptyset]_g \in U_g$, *and if* $\tau_1, \tau_2, \ldots, \tau_\kappa \in T \setminus \{\emptyset\}$ *then* $[\tau_1, \tau_2, \ldots, \tau_\kappa]_g \in U_g$.

b) $G([\emptyset]_g)(x) = g(x)$ *and*
$G(u = [\tau_1, \ldots, \tau_\kappa]_g)(x) = g^{(\kappa)}(x)\big(F(\tau_1)(x), \ldots, F(\tau_\kappa)(x)\big)$.

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

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

We are now able to derive the B-series of the exact solution. Here and in the following, $\rho(\tau)$ denotes the number of nodes in a tree τ.

Theorem 3.1 *Let $\gamma : T \to \mathbb{N}$ be given by*

$$\gamma(\emptyset) = 1, \qquad \gamma(\bullet) = 1,$$
$$\gamma([\tau_1, \ldots, \tau_\kappa]) = \rho([\tau_1, \ldots, \tau_\kappa]) \prod_{j=1}^{\kappa} \gamma(\tau_j).$$

Then the solution $X(t+h)$ of (1.1) *starting at the point (t, x) can be written as a B-series $B(\varphi_t, x; h)$ with*

$$\varphi_t(\tau)(h) = \frac{\mu_t(h)^{\rho(\tau)}}{\gamma(\tau)} \text{ for } \tau \in T. \tag{3.2}$$

Proof Write the exact solution as some B-series $X(t+h) = B(\varphi_t, x; h)$. By applying Lemma 3.1 to f (in which case $U_g = T$) the integral form (1.3) of the SDE can be written as

$$\sum_{\tau \in T} \alpha(\tau) \cdot \varphi_t(\tau)(h) \cdot F(\tau)(x) = x + \sum_{\tau \in T} \alpha(\tau) \cdot \int_0^h \prod_{j=1}^{\kappa} \varphi_t(\tau_j)(s) \circ \mathrm{d}\mu_t(s) \cdot F(\tau)(x).$$

By comparing term by term we get

$$\varphi_t(\emptyset) \equiv 1, \quad \varphi_t(\bullet)(h) = \mu_t(h),$$
$$\varphi_t(\tau)(h) = \int_0^h \prod_{j=1}^{\kappa} \varphi_t(\tau_j)(s) \circ \mathrm{d}\mu_t(s) \quad \text{for} \quad \tau = [\tau_1, \ldots, \tau_\kappa] \in T.$$

This proves the theorem for $\tau = \emptyset$ and $\tau = \bullet$. The rest is proved by induction on the height of τ. If $\tau = [\tau_1, \dots, \tau_\kappa]$ then

$$\varphi_t(\tau)(h) = \int_0^h \prod_{j=1}^{\kappa} \varphi_t(\tau_j)(s) \circ \mathrm{d}\mu_t(s) = \int_0^h \prod_{j=1}^{\kappa} \frac{\mu_t(s)^{\rho(\tau_j)}}{\gamma(\tau_j)} \circ \mathrm{d}\mu_t(s)$$
$$\overset{\text{Lemma 1.1}}{=} \frac{\mu_t(h)^{\rho(\tau)}}{\rho(\tau)} \prod_{j=1}^{\kappa} \frac{1}{\gamma(\tau_j)} = \frac{\mu_t(h)^{\rho(\tau)}}{\gamma(\tau)}.$$

□

For the numerical approximation (1.4) the following result holds:

Theorem 3.2 *The numerical solution $Y(t+h)$ after one step with step size h starting at the point (t,x) as well as the corresponding stage values H_i can be written in terms of B-series*

$$H_i = B(\eta_{ti}, x; h), \qquad Y(t+h) = B(\Phi_t, x; h)$$

with $\eta_{ti}(\tau)(h) = (\Delta_{t,t+h}\mu)^{\rho(\tau)}\hat{\eta}_i(\tau)$, where

$$\hat{\eta}_i(\emptyset) = 1, \quad \hat{\eta}_i(\bullet) = \sum_{j=1}^{s} a_{ij}, \tag{3.3a}$$

$$\hat{\eta}_i(\tau) = \sum_{j=1}^{s} a_{ij} \prod_{k=1}^{\kappa} \hat{\eta}_j(\tau_k) \text{ if } \tau = [\tau_1, \dots, \tau_\kappa] \tag{3.3b}$$

and $\Phi_t(\tau)(h) = (\Delta_{t,t+h}\mu)^{\rho(\tau)}\hat{\Phi}(\tau)$, where

$$\hat{\Phi}(\emptyset) = 1, \quad \hat{\Phi}(\bullet) = \sum_{i=1}^{s} b_i, \tag{3.4a}$$

$$\hat{\Phi}([\tau_1, \dots, \tau_\kappa]) = \sum_{i=1}^{s} b_i \prod_{k=1}^{\kappa} \hat{\eta}_i(\tau_k). \tag{3.4b}$$

Proof Write H_i as a B-series, that is

$$H_i = \sum_{\tau \in T} \alpha(\tau)\eta_{ti}(\tau)(h)F(\tau)(x).$$

Use (1.4a) together with Lemma 3.1 to obtain

$$\begin{aligned}
H_i &= x + \Delta_{t,t+h}\mu \sum_{j=1}^{s} a_{ij} f(H_j) \\
&= x + \Delta_{t,t+h}\mu \sum_{j=1}^{s} a_{ij} f\Big(\sum_{\tau \in T} \alpha(\tau)\eta_{tj}(\tau)(h)F(\tau)(x) \Big) \\
&= x + \Delta_{t,t+h}\mu \sum_{j=1}^{s} a_{ij} \sum_{\tau \in T} \alpha(\tau)\psi_{\eta_{tj}}(\tau)(h)F(\tau)(x) \\
&= x + \sum_{\tau \in T} \alpha(\tau) \left(\Delta_{t,t+h}\mu \sum_{j=1}^{s} a_{ij} \psi_{\eta_{tj}}(\tau)(h) \right) F(\tau)(x).
\end{aligned}$$

Thus $\eta_{ti}(\tau) = \Delta_{t,t+h}\mu \sum_{j=1}^{s} a_{ij}\psi_{\eta_{tj}}(\tau) + \begin{cases} 1 & : \tau = \emptyset \\ 0 & : \tau \neq \emptyset \end{cases}$, proving (3.3). Analogously,

$$\begin{aligned} Y(t+h) &= x + \Delta_{t,t+h}\mu \sum_{i=1}^{s} b_i f(H_i) \\ &= x + \sum_{\tau\in T} \alpha(\tau) \left(\Delta_{t,t+h}\mu \sum_{i=1}^{s} b_i \psi_{\eta_{ti}}(\tau)(h) \right) F(\tau)(x), \end{aligned}$$

proving (3.4). □

To decide the weak order we will also need the B-series of a function g, evaluated at the exact and the numerical solution. From Theorems 3.1 and 3.2 and Lemma 3.1 we obtain

$$g(X(t+h)) = \sum_{u\in U_g} \beta(u) \cdot \psi_{\varphi_t}(u)(h) \cdot G(u)(x),$$
$$g(Y(t+h)) = \sum_{u\in U_g} \beta(u) \cdot \psi_{\Phi_t}(u)(h) \cdot G(u)(x),$$

with

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

and

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

So, for the weak local error it follows

$$le_g(h;t,x) = \sum_{u\in U_g} \beta(u) \cdot \mathrm{E}\left[\psi_{\Phi_t}(u)(h) - \psi_{\varphi_t}(u)(h)\right] \cdot G(u)(x).$$

For the mean respectively mean square local error we obtain from Theorems 3.1 and 3.2

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

4 Proofs of Theorem 1.1 and Lemmas 1.1 and 1.2

With all the B-series in place, we can now present the order conditions for the weak and strong convergence. For convenience, we first summarize the assumptions on f:

Assumption 4.1 *Let $f \in C^{2(p+1)}(\mathbb{R}^d, \mathbb{R}^d)$ and f and $f'f$ fulfill a Lipschitz condition. Further, assume*

- *for mean-square convergence, that either*
  - *all elementary differentials $F(\tau)$ fulfill a linear growth condition, or*

  * *there exists a constant C such that $\|f'(y)\| \le C \quad \forall y \in \mathbb{R}^d$ (which implies the global Lipschitz condition) and all necessary partial derivatives exist [3],*
- *respectively for weak convergence, that $f \in C_P^{2(p+1)}(\mathbb{R}^d, \mathbb{R}^d)$.*

We have weak consistency of order p_μ if and only if

$$\mathrm{E}\,\psi_{\Phi_t}(u)(h) = \mathrm{E}\,\psi_{\varphi_t}(u)(h) + \mathscr{O}(h^{p_\mu+1}) \quad \forall u \in U_g \tag{4.1}$$

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

$$\mathrm{E}\left((\Phi_t(\tau)(h) - \varphi_t(\tau)(h))^2\right) = \mathscr{O}(h^{2p_\mu+1}) \quad \forall \tau \in T, \tag{4.2}$$

$$\mathrm{E}\,\Phi_t(\tau)(h) = \mathrm{E}\,\varphi_t(\tau)(h) + \mathscr{O}(h^{p_\mu+1}) \quad \forall \tau \in T. \tag{4.3}$$

Assume $\Delta_{t,h}\mu = \mu_t(h)$. Due to $\mathrm{E}\,\mu_t(h)^{2\rho(\tau)} = \mathscr{O}(h^{\rho(\tau)})$ by Lemma 1.2, (4.2) is then by Theorems 3.1 and 3.2 automatically fulfilled for all $\tau \in T$ with $\rho(\tau) \ge 2p_\mu + 1$, and satisfied for the remaining trees if and only if

$$\hat{\Phi}(\tau) = \frac{1}{\gamma(\tau)} \quad \forall \tau \in T \text{ with } \rho(\tau) \le 2p_\mu. \tag{4.4}$$

Note that (4.4) is just the condition that for the order p_d of the Runge–Kutta method applied to a deterministic system ($\sigma = 0$) it holds $p_d = 2p_\mu$.

Similarly, (4.3) is automatically fulfilled for all $\tau \in T$ with

$$\rho(\tau) \ge \begin{cases} 2p_\mu + 2 & : \text{for even } \rho(\tau) \\ 2p_\mu + 1 & : \text{for odd } \rho(\tau) \end{cases},$$

and satisfied for the remaining trees if and only if

$$\hat{\Phi}(\tau) = \frac{1}{\gamma(\tau)} \quad \forall \tau \in T \text{ with } \rho(\tau) \le \begin{cases} 2p_\mu + 1 & : \text{for even } \rho(\tau) \\ 2p_\mu & : \text{for odd } \rho(\tau) \end{cases}. \tag{4.5}$$

Thus, the method will be mean-square consistent of order p_μ if its deterministic order is

$$p_d = \begin{cases} 2p_\mu & : \text{if } p_\mu \in \mathbb{N} \\ 2p_\mu + 1 & : \text{if } p_\mu + \frac{1}{2} \in \mathbb{N} \end{cases} \tag{4.6}$$

or, vice versa, a method of deterministic order p_d will converge with mean-square order $\lfloor \frac{p_d}{2} \rfloor$.

Assume now that $\Delta_{t,h}\mu$ is chosen such that at least the first $2p_\mu + 1$ moments coincide with those of $\mu_t(h)$, and all the others are in $\mathscr{O}(h^{p_\mu+1})$. Analogously to the discussion of (4.3), (4.1) is automatically fulfilled for all $u \in U_g$ with

$$\rho(u) \ge \begin{cases} 2p_\mu + 2 & : \text{for even } \rho(u) \\ 2p_\mu + 1 & : \text{for odd } \rho(u) \end{cases},$$

and satisfied for the remaining trees if and only if (4.5) is fulfilled. Thus, we obtain that the weak order of the method equals its mean-square order, which finishes the proof of Theorem 1.1.

Finally we present the proofs of Lemmas 1.1 and 1.2.

Proof (Lemma 1.1) As μ_t is a semimartingale with continuous paths, it holds for $f \in C^2(\mathbb{R},\mathbb{R})$ that

$$f(\mu_t(h)) - f(\mu_t(0)) = \int_0^h f'(\mu_t(s)) \circ \mathrm{d}\mu_t(s),$$

which for $f(x) = x^{k+1}$ immediately gives the assertion. □

Proof (Lemma 1.2) The assertion follows from

$$\mathrm{E}\,\mu_t(h)^n = \sum_{i=0}^{n} \binom{n}{i} \lambda^{n-i} h^{n-i} \sigma^i \,\mathrm{E}\left(\big(W(t+h)-W(t)\big)^i\right)$$

and using that

$$\mathrm{E}\left(\big(W(t+h)-W(t)\big)^i\right) = h^{\frac{i}{2}} \begin{cases} 0 & : \text{ if } i \text{ is odd,} \\ (i-1)\cdot(i-3)\cdot\cdots\cdot 3\cdot 1 & : \text{otherwise.} \end{cases}$$

□

5 Numerical experiments

To verify the theoretical result, we solve three test problems by method (1.4), based on some classes of well known deterministic Runge–Kutta methods. The first group consists of the Gauss methods of deterministic order $p_d = 2s$, in which case the predicted stochastic order $p_\mu = s$, as already demonstrated in Example 1.1. Using the results from [11, 15] it is straightforward to confirm that the Gauss methods preserve quadratic invariants. Two of the test problems below have such invariants, for these problems we also demonstrate this conservation property. The second group of methods consists of the Radau IIA methods, these are of deterministic order $p_d = 2s-1$, thus $p_\mu = s-1$. Finally, we consider three explicit Runge–Kutta (ERK) methods: A third order, three stage method [5, RK32, p. 95], the classical fourth order Runge–Kutta method [10, p. 138] and the fifth order Fehlberg method [10, p. 177]. In the following, they will be denoted by ERK3, ERK4 and ERK5, respectively.

For the calculation of the numerical order p_{num}, errors less than 10^{-14} have been ignored.

Example 5.1 This is a continuation of Example 1.1. The SDE (1.5) is solved by the Radau IIA methods, as well as the three explicit methods. The number of independent simulations is still $M = 10,000$. Convergence plots are given in Figure 5.1, as well as the estimated order p_{num}. The order p_{num} is slightly above p_μ, probably because the error of the deterministic part becomes more dominant for larger step sizes.

Example 5.2 The next example is the well known Kubo oscillator, a prototype problem for solving oscillatory problems, [17, sec.4.4.1]. The SDE is given by

$$\mathrm{d}X = \begin{bmatrix} 0 & -a \\ a & 0 \end{bmatrix} X\,\mathrm{d}t + \begin{bmatrix} 0 & -\sigma \\ \sigma & 0 \end{bmatrix} X \circ \mathrm{d}W, \tag{5.1}$$

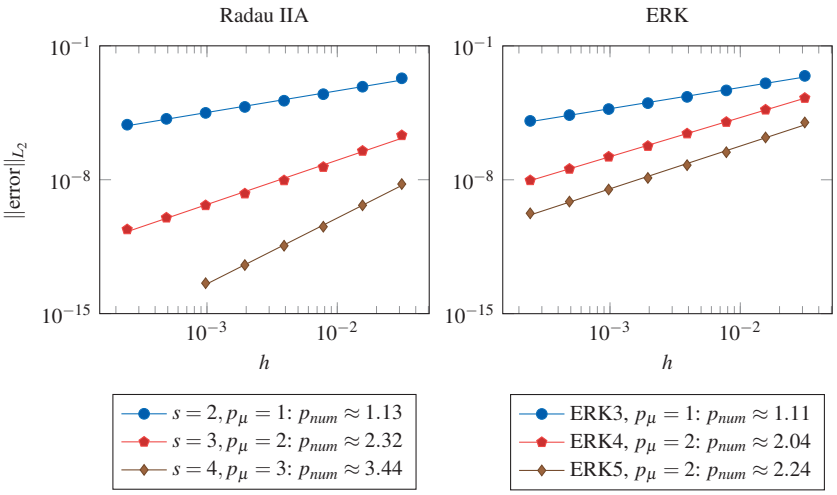

Fig. 5.1 Error plots for the Radau IIA and the ERK methods applied to (1.5).

where a and σ are real parameters. With $X(0)=[1,0]^T$ this problem has as exact solution $X(t)=[\cos(at+\sigma W(t)),\sin(at+\sigma W(t))]^T$. In our experiments, we have used $a=\sigma=1.0$, the mean square error at $t=1.0$ is estimated based on $M=1,000$ simulations.

The Kubo oscillator has the invariant $I(X)=X_1^2+X_2^2$. To see how well this is preserved by the numerical methods, we have computed one solution path by the Gauss ($s=2$) method, one by the Radau IIA ($s=3$) and one by the ERK5 method, all with $p_\mu=2$. The step size was $h=0.5$, and the integration interval [0,1000]. From the picture at the bottom right of Figure 5.2 it is clear that the Gauss solution stays on the circle given by $I(X(t))=I(X(0))$, the others do not.

Example 5.3 This example is based on the deterministic rigid body model from [9]. This model has also been used in [7] for studying energy-preserving integrators. The SDE is given by

$$\mathrm{d}X = A(X)X\,\mathrm{d}t + \sigma A(X)X\circ\mathrm{d}W \tag{5.2}$$

with

$$A(X)=\begin{pmatrix} 0 & X_3/I_3 & -X_2/I_2 \\ -X_3/I_3 & 0 & X_1/I_1 \\ X_2/I_2 & -X_1/I_1 & 0 \end{pmatrix}$$

and parameters $I_1=2$, $I_2=1$ and $I_3=2/3$. As initial value we choose $X(0)=(\cos(1.1),0,\sin(1.1))^T$. This problem conserves the invariants

$$H(X)=\frac{1}{2}\left(\frac{X_1^2}{I_1}+\frac{X_2^2}{I_2}+\frac{X_3^2}{I_3}\right),\qquad C(X)=X_1^2+X_2^2+X_3^2.$$

The equation with $\sigma=0.5$ was solved, and the mean square errors at $t=1.0$ based on $M=1,000$ independent simulations are presented in Figure 5.3. For the Gauss and

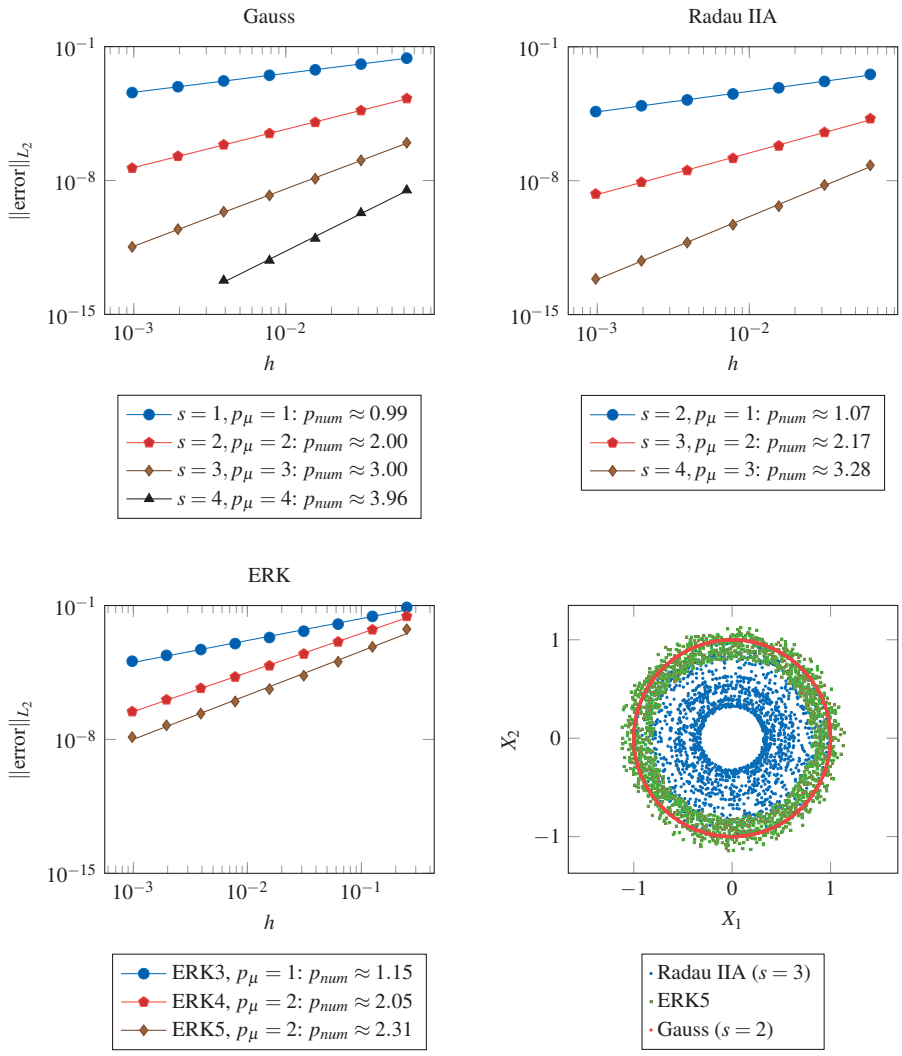

Fig. 5.2 Error plots for the Kubo oscillator (5.1) (*top, and bottom left*). The solution points for one trajectory, using ERK5 (Fehlberg), Radau IIA ($s=3$) and Gauss ($s=2$) with $h=0.5$ for $t\leq 1000$ (*bottom right*).

Radau IIA methods, some Newton iterations fail for the larger step sizes. In order to demonstrate the conservative properties of the Gauss method, a plot of the Casimir $C(X)$ for one trajectory computed by the Gauss ($s=2$) method, the Radau IIA ($s=3$) method and ERK5 has been included. As expected, the Gauss method preserves the Casimir, the others do not.

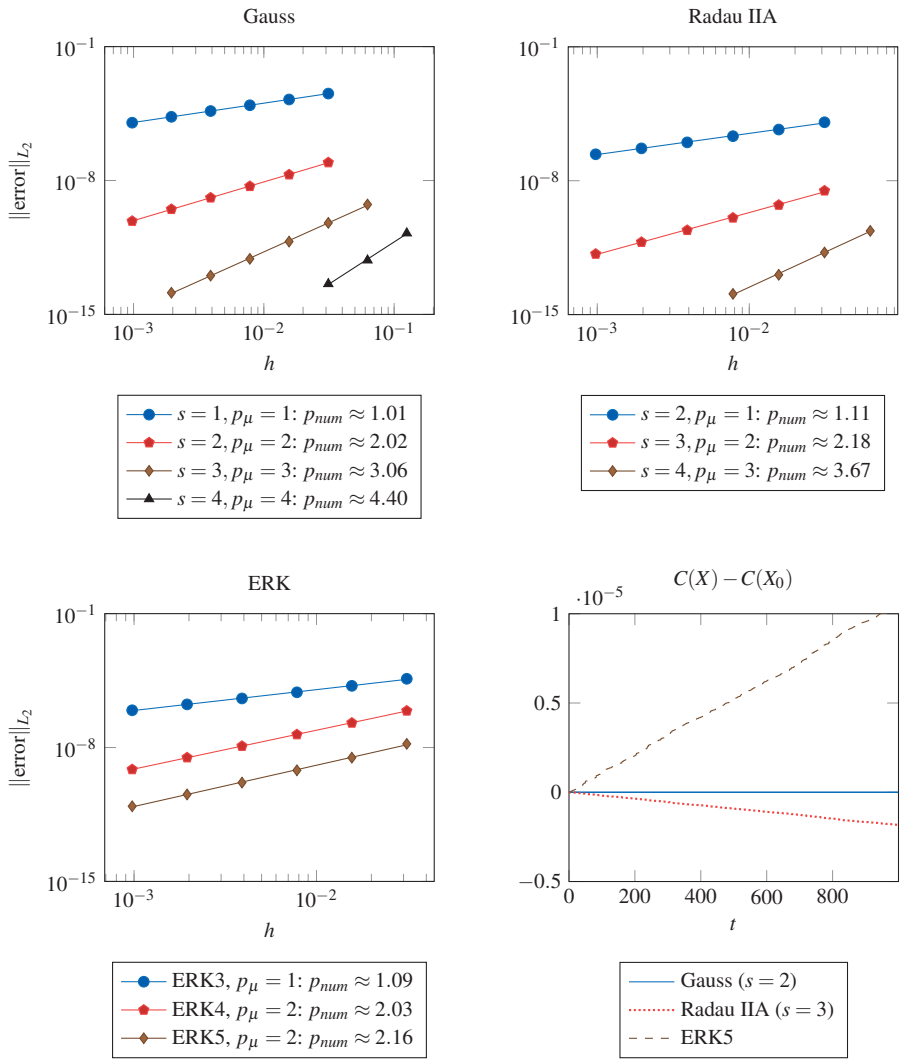

Fig. 5.3 Error plots for the stochastic rigid body problem (5.2) (*top and bottom left*). Drift of the Casimir $C(X)$ for Gauss ($s=2$), Radau IIA ($s=3$) and ERK5 (Fehlberg) using $h=2^{-5}$ for $t\leq 1000$ (*bottom right*).

6 Conclusion

We have proved that a straightforward extension of deterministic Runge–Kutta methods of order p to Stratonovich single integrand SDEs results in methods converging with order $\lfloor p/2\rfloor$ in the mean-square and weak sense. They also inherit certain properties from their deterministic origin, like preservation of quadratic invariants. These methods are cheaply implementable and seem to be preferable for solving single integrand problems.

References

1. Burrage, K., Burrage, P., Higham, D., Kloeden, P., Platen, E.: Comment on "numerical methods for stochastic differential equations". Physical Review E - Statistical, Nonlinear, and Soft Matter Physics **74**(6) (2006). DOI 10.1103/PhysRevE.74.068701
2. Burrage, K., Burrage, P.M.: High strong order explicit Runge–Kutta methods for stochastic ordinary differential equations. Appl. Numer. Math. **22**(1-3), 81–101 (1996). DOI 10.1016/S0168-9274(96)00027-X. Special issue celebrating the centenary of Runge-Kutta methods
3. Burrage, K., Burrage, P.M.: Order conditions of stochastic Runge–Kutta methods by *B*-series. SIAM J. Numer. Anal. **38**(5), 1626–1646 (electronic) (2000). DOI 10.1137/S0036142999363206
4. Butcher, J.C.: Coefficients for the study of Runge–Kutta integration processes. J. Austral. Math. Soc. **3**, 185–201 (1963)
5. Butcher, J.C.: Numerical methods for ordinary differential equations. 2nd revised ed. Hoboken, NJ: John Wiley & Sons. xix, 463 p. (2008)
6. Cohen, D.: On the numerical discretisation of stochastic oscillators. Mathematics and Computers in Simulation **82**(8), 1478 – 1495 (2012). DOI 10.1016/j.matcom.2012.02.004
7. Cohen, D., Dujardin, G.: Energy-preserving integrators for stochastic Poisson systems. Commun. Math. Sci. **12**(8), 1523–1539 (2014). DOI 10.4310/CMS.2014.v12.n8.a7
8. Debrabant, K., Kværnø, A.: 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). DOI 10.1137/070704307
9. Hairer, E., Lubich, C., Wanner, G.: Geometric numerical integration, *Springer Series in Computational Mathematics*, vol. 31, second edn. Springer-Verlag, Berlin (2006). Structure-preserving algorithms for ordinary differential equations
10. Hairer, E., Nørsett, S.P., Wanner, G.: Solving ordinary differential equations. I: Nonstiff problems. 2nd revised ed., 3rd corrected printing. Springer Series in Computational Mathematics 8. Berlin: Springer. xv, 528 p. (2010)
11. Hong, J., Xu, D., Wang, P.: Preservation of quadratic invariants of stochastic differential equations via Runge-Kutta methods. Appl. Numer. Math. **87**, 38–52 (2015). DOI 10.1016/j.apnum.2014.08.003
12. Kloeden, P.E., Platen, E.: Numerical solution of stochastic differential equations, *Applications of Mathematics*, vol. 21, 2 edn. Springer-Verlag, Berlin (1999)
13. Komori, Y.: Multi-colored rooted tree analysis of the weak order conditions of a stochastic Runge-Kutta family. Appl. Numer. Math. **57**(2), 147–165 (2007). DOI 10.1016/j.apnum.2006.02.002
14. Komori, Y., Mitsui, T., Sugiura, H.: Rooted tree analysis of the order conditions of ROW-type scheme for stochastic differential equations. BIT **37**(1), 43–66 (1997). DOI 10.1007/BF02510172
15. Ma, Q., Ding, D., Ding, X.: Symplectic conditions and stochastic generating functions of stochastic Runge-Kutta methods for stochastic Hamiltonian systems with multiplicative noise. Appl. Math. Comput. **219**(2), 635–643 (2012). DOI 10.1016/j.amc.2012.06.053
16. Milstein, G.N.: Numerical integration of stochastic differential equations, *Mathematics and its Applications*, vol. 313. Kluwer Academic Publishers Group, Dordrecht (1995). Translated and revised from the 1988 Russian original
17. Milstein, G.N., Tretyakov, M.V.: Stochastic numerics for mathematical physics. Scientific Computation. Berlin: Springer. ixx, 594 p. (2004)
18. Rößler, A.: Stochastic Taylor expansions for the expectation of functionals of diffusion processes. Stoch. Anal. Appl. **22**(6), 1553–1576 (2004). DOI 10.1081/SAP-200029495
19. Rößler, A.: 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). DOI 10.1080/07362990500397699
20. Sobczyk, K.: Stochastic models for fatigue damage of materials. Adv. in Appl. Probab. **19**(3), 652–673 (1987). DOI 10.2307/1427411
21. Tian, T., Burrage, K.: Implicit Taylor methods for stiff stochastic differential equations. Appl. Numer. Math. **38**(1-2), 167–185 (2001). DOI 10.1016/S0168-9274(01)00034-4