A New Class Of Symplectic Methods For Stochastic Hamiltonian Systems

Cristina Anton

NOTICE: This is the peer reviewed version of the following article: Anton, C. (2025). A new class of symplectic methods for stochastic Hamiltonian systems. *Applied Numerical Mathematics*, 208, 43–59, which has been published in final form at https://doi.org/10.1016/j.apnum.2024.01.021.

A New Class of Symplectic Methods for Stochastic Hamiltonian Systems

Cristina Anton[a,*]

[a]*Department of Mathematics and Statistics Grant MacEwan University, 5-103C, 10700 104 Avenue NW, Edmonton, AB T5J 4S2, Canada*

Abstract

We propose a systematic approach to construct a new family of stochastic symplectic schemes for the strong approximation of the solution of stochastic Hamiltonian systems. Our approach is based both on B-series and generating functions. The proposed schemes are a generalization of the implicit midpoint rule, they require derivatives of the Hamiltonian functions of at most order two, and are constructed by defining a generating function. We construct some schemes with strong convergence order one and a half, and we illustrate numerically their long term performance.

Keywords: stochastic Hamiltonian systems, stochastic Runge-Kutta methods, generating function, symplectic integration.

1. Introduction

For deterministic Hamiltonian systems, symplectic schemes give lower errors for long time simulations [8], and similarly, for stochastic Hamiltonian systems (SHSs), it was illustrated through long time numerical simulations that symplectic methods produce more accurate results than non-symplectic ones (see, e.g., [7, 15, 14, 18]).

In recent years, several derivative free symplectic stochastic Runge-Kutta (SRK) methods for strong approximation for the solutions of SHS were proposed, but, unless a special type of SHS is considered, these methods have strong order 0.5 or 1 [3, 9, 11, 14, 15, 18]. Alternatively, in [7] we present

*corresponding author

Email address: popescuc@macewan.ca (Cristina Anton)

a systematic way, based on generating functions, to construct symplectic schemes, but the schemes with high strong order depend on high order derivatives of the Hamiltonian functions. Here we extend to the stochastic case the approach used for deterministic Hamiltonian systems in [12, 13], and combining ideas used for stochastic Runge-Kutta methods and for methods based on generating functions, we construct a new family of schemes that are symplectic by construction and requires derivatives of Hamiltonian functions of at most order two.

Symplectic stochastic Runge-Kutta (SRK) methods for SHSs are constructed using only the first derivatives of the Hamiltonian functions, but require a large number of conditions, so it is difficult to obtain symplectic schemes of high strong order. In [3, 11, 18] symplectic SRK schemes of strong order 0.5 are constructed for general SHS with multiplicative noise, and symplectic SRK methods of strong order 1 are proposed for SHS with separable Hamiltonians or with commutative noise. For SHS with additive noise, symplectic SRK methods of order 1.5 or 2 are constructed in [9, 15, 18].

On the other hand, following the approach based on generating functions proposed in [7], we could obtain methods of strong order 1.5 or higher that are symplectic by construction, but they require high order derivatives of the Hamiltonian functions (e.g., for general SHS with multiplicative noise, for symplectic schemes of order 1.5 we need 3rd order derivatives of the Hamiltonian functions). Here we propose a new family of symplectic schemes constructed by defining a generating function similar with the ones associated to symplectic SRK methods [14, 11, 9]. We also incorporate ideas used to construct computationally efficient SRK methods in [17]. Since they require derivatives of Hamiltonian functions of at most order two and are symplectic by construction, the proposed schemes are a trade-off between the SRK methods and the symplectic methods based on generating functions.

Let $(\Omega, \mathcal{F}, P)$ be a probability space with a filtration $\{\mathcal{F}_t\}_{t\geq 0}$. A $2d$-dimensional SHS [15] in the sense of Stratonovich has the form:

$$
\begin{aligned}
dP &= -\nabla_Q H_0(P,Q)dt - \sum_{r=1}^{m} \nabla_Q H_r(P,Q) \circ dW_r(t), \quad P(0) = p_0, \\
dQ &= \nabla_P H_0(P,Q)dt + \sum_{r=1}^{m} \nabla_P H_r(P,Q) \circ dW_r(t), \quad Q(0) = q_0,
\end{aligned} \tag{1}
$$

where $t \in [0,T]$, $P(t)$, $Q(t)$, p_0, q_0 are d-dimensional vectors, $H_r : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$, $r = 0, \dots, m$, $\nabla_Q H_r$ and $\nabla_P H_r$ denote the gradients with respect to Q and P respectively, and $W_r, r = 1, \ldots, m$ are independent standard Wiener processes with respect to $\{\mathcal{F}_t\}_{t\geq 0}$. We suppose that H_r are differentiable functions with bounded second order partial derivatives such that $H_r \in \mathcal{C}_P^k(\mathbb{R}^{2d}, \mathbb{R}) := \{g \in C^k(\mathbb{R}^d, \mathbb{R}), |\partial_x^i g(x)| \leq K(1 + |x|^{2n})$ for all $x \in \mathbb{R}^d$ and any partial derivative $\partial_x^i g$ of order $i \leq k$, where $K > 0$ and $n \in \mathbb{N}$ depend on $g\}$. Here, and in the rest of the paper, $|\cdot|$ denotes the Euclidean norm. The order k will be specified later and it depends on the numerical scheme applied to approximate the solutions of the SHS (1). The solution $X(t, x_0) = (P(t, x_0), Q(t, x_0))^\top \in \mathbb{R}^{2d}$, $x_0 = (p_0, q_0)^\top \in \mathbb{R}^{2d}$, of (1) preserves the symplectic structure [14]:

$$\left(\frac{\partial X(t, x_0)}{\partial x_0}\right)^T J \left(\frac{\partial X(t, x_0)}{\partial x_0}\right) = J \quad \text{with} \quad J = \begin{pmatrix} 0 & I_d \\ -I_d & 0 \end{pmatrix},$$

where I_d denotes the d-dimensional identity matrix.

Notice that we can re-write (1) as a $2d-$ dimensional autonomous SDE

$$dX(t) = J^{-1}\nabla H_0(X(t))dt + \sum_{l=1}^{m} J^{-1}\nabla H_l(X(t)) \circ dW_l(t), \quad X(0) = x_0. \quad (2)$$

We denote $g_l := J^{-1}\nabla H_l$, $l = 0, \ldots, m$, and with this notation (2) becomes

$$dX(t) = \sum_{l=0}^{m} g_l(X(t)) \circ dW_l(t), X(0) = x_0 \in \mathbb{R}^{2d}, \quad W_0(t) = t. \quad (3)$$

We consider the generating function

$$S(z, h) := \sum_{i=1}^{s}\sum_{\mu=0}^{M} a_i^{(0,\mu)} H_0(Y_i^{(0,\mu)}) + \sum_{i=1}^{s}\sum_{l=1}^{m}\sum_{\mu=0}^{M} b_i^{(l,\mu)} H_l(Y_i^{(l,\mu)}) \quad (4)$$

where for $i = 1, \ldots, s$, $l = 0, \ldots, m$

$$\begin{aligned} Y_i^{(l,\mu)} &= z + J^{-1}\sum_{j=1}^{s}\sum_{\nu=0}^{M} A_{ij}^{(l,\mu)(0,\nu)}\nabla H_0(Y_j^{(0,\nu)}) \\ &+ J^{-1}\sum_{j=1}^{s}\sum_{k=1}^{m}\sum_{\nu=0}^{M} B_{ij}^{(l,\mu)(k,\nu)}\nabla H_k(Y_j^{(k,\nu)}) \end{aligned} \quad (5)$$

$$= z + \sum_{j=1}^{s}\sum_{\nu=0}^{M} A_{ij}^{(l,\mu)(0,\nu)} g_0(Y_j^{(0,\nu)}) + \sum_{j=1}^{s}\sum_{k=1}^{m}\sum_{\nu=0}^{M} B_{ij}^{(l,\mu)(k,\nu)} g_k(Y_j^{(k,\nu)}). \tag{6}$$

The random variables $a_i^{(0,\mu)}$, $b_i^{(l,\mu)}$, $A_{ij}^{(l,\mu)(0,\nu)}$, and $B_{ij}^{(l,\mu)(k,\nu)}$ are chosen such that we have

$$(a_i^{(0,\mu)})^{2^k} = \mathcal{O}(h^{2^k}), \quad (A_{ij}^{(l,\mu)(0,\nu)})^{2^k} = \mathcal{O}(h^{2^k}), \quad (b_i^{(l,\mu)})^{2^k} = \mathcal{O}(h^{2^{k-1}}), \tag{7}$$

$$(B_{ij}^{(l,\mu)(k,\nu)})^{2^k} = \mathcal{O}(h^{2^{k-1}}), \tag{8}$$

where the $\mathcal{O}(\cdot)$ notation refers to the $L^2(\Omega)$ norm and the time step $h \to 0$, [6]. The one-step approximation $y^* = \psi_h(y)$ of the proposed method is defined as

$$y^* = y + J^{-1}\nabla S\left(\frac{1}{2}(y + y^*), h\right). \tag{9}$$

For any index $\alpha = (j_1, j_2, \ldots, j_l), j_i \in \{0, 1, \ldots, m\}, i = 1, \ldots, l$ we denote by J_α and I_α the multiple Stratonovich and Ito integrals respectively [10]:

$$J_\alpha = \int_0^t \int_0^{s_l} \cdots \int_0^{s_2} \circ dW_{j_1}(s_1) \cdots \circ dW_{j_{l-1}}(s_{l-1}) \circ dW_{j_l}(s_l)$$

$$I_\alpha = \int_0^t \int_0^{s_l} \cdots \int_0^{s_2} dW_{j_1}(s_1) \ldots dW_{j_{l-1}}(s_{l-1}) dW_{j_l}(s_l)$$

Let $\mathcal{A}_k = \{\alpha : l(\alpha) + n(\alpha) \leq 2k\}$ and $\mathcal{B}_k = \{\alpha : l(\alpha) + n(\alpha) \leq 2k$ or $l(\alpha) = n(\alpha) = k + 0.5\}$, where $l(\alpha)$ is the length of the multi-index α and $n(\alpha)$ is the number of zero components (e.g. $l((0,0,1)) = 3$, $n((0,0,1)) = 2$).

The same formula (9) is used in [7] to construct symplectic schemes based on generating function, but the generating function S is defined as a truncation of $\sum_\alpha G_\alpha J_\alpha$. The coefficients G_α are defined by a recurrence [7], and include derivatives of the Hamiltonian functions of maximum order equal with $l(\alpha) - 1$. For example, for a scheme of strong order $k = 1, 2, \ldots$, we have $S = \sum_{\alpha \in \mathcal{A}_k} G_\alpha J_\alpha$. For $k = 1/2, 3/2, \ldots$, we use the relation of the Ito stochastic multiple integrals and the Stratonovich stochastic multiple integrals [10], and we replace in the expansion of S each Stratonovich integral in terms of Ito integrals I_α. We truncate the series by keeping only terms corresponding to Ito integrals I_α with $\alpha \in \mathcal{B}_k$.

In the next section we present the proposed methods. Section 3 include results regarding colored trees and the B-series associated with the proposed schemes. Based on these results in section 4 we study the convergence order and show that for these symplectic schemes only one order condition per unrooted tree remains to be satisfied. In section 5 we construct some schemes of strong order 1.5 and section 6 includes numerical experiments illustrating the performance of the schemes. The last section contains the conclusions.

2. The methods

For any function $f : \mathbb{R}^k \to \mathbb{R}$, $f \in C^2(\mathbb{R}^k)$, we denote by $\nabla f = \left(\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_k}\right)^T$ the gradient of f and by $\nabla^2 f = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}\right)_{i,j=1,\ldots,k} \in \mathcal{M}_{k,k}(\mathbb{R})$ the Hessian matrix of f.

To implement the scheme given by the one-step approximation (9) with the generating function defined by (4) and (5) we need to compute the gradient of $S(z,h)$:

$$\nabla S(z,h) = \sum_{i=1}^{s} \sum_{\mu=0}^{M} a_i^{(0,\mu)} \left(\frac{\partial Y_i^{(0,\mu)}}{\partial z}\right)^\top \nabla H_0(Y_i^{(0,\mu)}) + \sum_{i=1}^{s} \sum_{l=1}^{m} \sum_{\mu=0}^{M} b_i^{(l,\mu)} \left(\frac{\partial Y_i^{(l,\mu)}}{\partial z}\right)^\top \nabla H_l(Y_i^{(l,\mu)}),$$

where for any $i = 1, \ldots, s$, $l = 0, \ldots, m$, $\mu = 0, \ldots, M$, the Jacobian matrices $\frac{\partial Y_i^{(l,\mu)}}{\partial z}$ are given by

$$\frac{\partial Y_i^{(l,\mu)}}{\partial z} = I_{2d} + J^{-1} \sum_{j=1}^{s} \sum_{\nu=0}^{M} A_{ij}^{(l,\mu)(0,\nu)} \frac{\partial K_j^{(0,\nu)}}{\partial z} + J^{-1} \sum_{j=1}^{s} \sum_{k=1}^{m} \sum_{\nu=0}^{M} B_{ij}^{(l,\mu)(k,\nu)} \frac{\partial K_j^{(k,\nu)}}{\partial z},$$

with

$$K_i^{(l,\mu)} := \nabla H_l(Y_i^{(l,\mu)}), \quad \frac{\partial K_i^{(l,\mu)}}{\partial z} = \nabla^2 H_l\left(Y_i^{(l,\mu)}\right) \frac{\partial Y_i^{(l,\mu)}}{\partial z}.$$

To avoid the expensive computation of the Jacobian matrices $\frac{\partial Y_i^{(l,\mu)}}{\partial z}$ and $\frac{\partial K_i^{(l,\mu)}}{\partial z}$ we proceed as in [12]. Firstly, for arbitrary column vectors $\hat{Y}_i^{(l,\mu)}$,

$v_i^{(l,\mu)} \in \mathbb{R}^{2d}$ we have

$$\begin{aligned}
\nabla S(z,h) &= \sum_{i=1}^{s}\sum_{\mu=0}^{M} a_i^{(0,\mu)} \left(\frac{\partial Y_i^{(0,\mu)}}{\partial z}\right)^\top \nabla H_0(Y_i^{(0,\mu)}) \\
&+ \sum_{i=1}^{s}\sum_{l=1}^{m}\sum_{\mu=0}^{M} b_i^{(l,\mu)} \left(\frac{\partial Y_i^{(l,\mu)}}{\partial z}\right)^\top \nabla H_l(Y_i^{(l,\mu)}) \\
&+ \sum_{i=1}^{s}\sum_{l=0}^{m}\sum_{\mu=0}^{M}\Bigg(I_{2d} - \sum_{j=1}^{s}\sum_{\nu=0}^{M} A_{ij}^{(l,\mu)(0,\nu)} \left(\frac{\partial K_j^{(0,\nu)}}{\partial z}\right)^\top J^{-1} \\
&- \sum_{j=1}^{s}\sum_{k=1}^{m}\sum_{\nu=0}^{M} B_{ij}^{(l,\mu)(k,\nu)} \left(\frac{\partial K_j^{(k,\nu)}}{\partial z}\right)^\top J^{-1} - \left(\frac{\partial Y_i^{(l,\mu)}}{\partial z}\right)^\top \Bigg)\hat{Y}_i^{(l,\mu)} \\
&+ \sum_{i=1}^{s}\sum_{l=0}^{m}\sum_{\mu=0}^{M} \left(\left(\frac{\partial Y_i^{(l,\mu)}}{\partial z}\right)^\top \nabla^2 H_l\left(Y_i^{(l,\mu)}\right) - \left(\frac{\partial K_i^{(l,\mu)}}{\partial z}\right)^\top \right) v_i^{(l,\mu)}.
\end{aligned}$$

Here we have used $(J^{-1})^\top = -J^{-1}$. Reordering terms we obtain

$$\begin{aligned}
\nabla S(z,h) = &\sum_{i=1}^{s}\sum_{l=0}^{m}\sum_{\mu=0}^{M} \hat{Y}_i^{(l,\mu)} + \sum_{i=1}^{s}\sum_{\mu=0}^{M} \left(\frac{\partial Y_i^{(0,\mu)}}{\partial z}\right)^\top \Big(a_i^{(0,\mu)} \nabla H_0(Y_i^{(0,\mu)}) \\
&+ \nabla^2 H_0\left(Y_i^{(0,\mu)}\right) v_i^{(0,\mu)} - \hat{Y}_i^{(0,\mu)} \Big) + \sum_{i=1}^{s}\sum_{l=1}^{m}\sum_{\mu=0}^{M} \left(\frac{\partial Y_i^{(l,\mu)}}{\partial z}\right)^\top \Big(b_i^{(l,\mu)} \nabla H_l(Y_i^{(l,\mu)}) \\
&+ \nabla^2 H_l\left(Y_i^{(l,\mu)}\right) v_i^{(l,\mu)} - \hat{Y}_i^{(l,\mu)} \Big) + \sum_{i=1}^{s}\sum_{\mu=0}^{M} \left(\frac{\partial K_i^{(0,\mu)}}{\partial z}\right)^\top \Big(-v_i^{(0,\mu)} \\
&- \sum_{j=1}^{s}\sum_{k=0}^{m}\sum_{\nu=0}^{M} A_{ji}^{(k,\nu)(0,\mu)} J^{-1} \hat{Y}_j^{(k,\nu)} \Big) + \sum_{i=1}^{s}\sum_{l=1}^{m}\sum_{\mu=0}^{M} \left(\frac{\partial K_i^{(l,\mu)}}{\partial z}\right)^\top \Big(-v_i^{(l,\mu)} \\
&- \sum_{j=1}^{s}\sum_{k=0}^{m}\sum_{\nu=0}^{M} B_{ji}^{(k,\nu)(l,\mu)} J^{-1} \hat{Y}_j^{(k,\nu)} \Big).
\end{aligned}$$

Next, for $i = 1, \dots, s$, $l = 1, \dots, m$, $\mu = 0, \dots, M$ we define

$$\hat{Y}_i^{(0,\mu)} := a_i^{(0,\mu)} \nabla H_0(Y_i^{(0,\mu)}) + \nabla^2 H_0\left(Y_i^{(0,\mu)}\right) v_i^{(0,\mu)}, \tag{10}$$

$$\hat{Y}_i^{(l,\mu)} := b_i^{(l,\mu)} \nabla H_l(Y_i^{(l,\mu)}) + \nabla^2 H_l \left(Y_i^{(l,\mu)} \right) v_i^{(l,\mu)}, \tag{11}$$

$$v_i^{(0,\mu)} := -\sum_{j=1}^{s} \sum_{k=0}^{m} \sum_{\nu=0}^{M} A_{ji}^{(k,\nu)(0,\mu)} J^{-1} \hat{Y}_j^{(k,\nu)},$$

$$v_i^{(l,\mu)} := -\sum_{j=1}^{s} \sum_{k=0}^{m} \sum_{\nu=0}^{M} B_{ji}^{(k,\nu)(l,\mu)} J^{-1} \hat{Y}_j^{(k,\nu)},$$

such that we get

$$\nabla S(z,h) = \sum_{i=1}^{s} \sum_{l=0}^{m} \sum_{\mu=0}^{M} \hat{Y}_i^{(l,\mu)}.$$

Let us denote the Jacobian of g_l by $g_l^{'}(z) := J^{-1} \nabla^2 H_l(z)$, $l = 0, \ldots, m$. Replacing from (10) and (11) we get

$$\Theta(z,h) := J^{-1} \nabla S(z,h) = \sum_{i=1}^{s} \sum_{\mu=0}^{M} \left(a_i^{(0,\mu)} g_0(Y_i^{(0,\mu)}) + g_0^{'} \left(Y_i^{(0,\mu)} \right) v_i^{(0,\mu)} \right)$$

$$+ \sum_{i=1}^{s} \sum_{l=1}^{m} \sum_{\mu=0}^{M} \left(b_i^{(l,\mu)} g_l(Y_i^{(l,\mu)}) + g_l^{'} \left(Y_i^{(l,\mu)} \right) v_i^{(l,\mu)} \right), \tag{12}$$

and for $i = 1, \ldots, s$, $l = 1, \ldots, m$, $\mu = 0, \ldots, M$ we have

$$v_i^{(0,\mu)} = -\sum_{j=1}^{s} \sum_{\nu=0}^{M} A_{ji}^{(0,\nu)(0,\mu)} a_j^{(0,\nu)} g_0(Y_j^{(0,\nu)}) - \sum_{j=1}^{s} \sum_{k=1}^{m} \sum_{\nu=0}^{M} A_{ji}^{(k,\nu)(0,\mu)} b_j^{(k,\nu)} g_k(Y_j^{(k,\nu)})$$

$$-\sum_{j=1}^{s} \sum_{k=0}^{m} \sum_{\nu=0}^{M} A_{ji}^{(k,\nu)(0,\mu)} g_k^{'} \left(Y_j^{(k,\nu)} \right) v_j^{(k,\nu)}, \tag{13}$$

$$v_i^{(l,\mu)} = -\sum_{j=1}^{s} \sum_{\nu=0}^{M} B_{ji}^{(0,\nu)(l,\mu)} a_j^{(0,\nu)} g_0(Y_j^{(0,\nu)}) - \sum_{j=1}^{s} \sum_{k=1}^{m} \sum_{\nu=0}^{M} B_{ji}^{(k,\nu)(l,\mu)} b_j^{(k,\nu)} g_k(Y_j^{(k,\nu)})$$

$$-\sum_{j=1}^{s} \sum_{k=0}^{m} \sum_{\nu=0}^{M} B_{ji}^{(k,\nu)(l,\mu)} g_k^{'} \left(Y_j^{(k,\nu)} \right) v_j^{(k,\nu)}. \tag{14}$$

Thus the one-one step approximation (9) can be expressed as

$$y^* = y + \Theta \left(\frac{1}{2}(y + y^*), h \right),$$

where $\Theta(z,h)$ is given by (12) and (6), (13), (14). Notice that if for any $l=0,\dots,m$, $k=1,\dots,m$, $\mu,\nu=0,\dots,M$ the matrices $\left(A_{ij}^{(l,\mu)(0,\mu)}\right)$, $\left(B_{ij}^{(l,\mu)(k,\nu)}\right)$ are strictly lower triangular, the scheme is only implicit in y^*. For arbitrary matrices, the method is fully implicit and we can also express it as

$$
\begin{aligned}
y^* &= y + \sum_{i=1}^{s}\sum_{\mu=0}^{M} a_i^{(0,\mu)} g_0(Y_i^{(0,\mu)}) + \sum_{i=1}^{s}\sum_{l=1}^{m}\sum_{\mu=0}^{M} b_i^{(l,\mu)} g_l(Y_i^{(l,\mu)}) \\
&\quad + \sum_{i=1}^{s}\sum_{l=0}^{m}\sum_{\mu=0}^{M} g_l^{'}\left(Y_i^{(l,\mu)}\right) v_i^{(l,\mu)}, \qquad (15)
\end{aligned}
$$

$$
\begin{aligned}
Y_i^{(l,\mu)} &= y + \sum_{j=1}^{s}\sum_{\nu=0}^{M} \left(A_{ij}^{(l,\mu)(0,\nu)} + \frac{1}{2}a_j^{(0,\nu)}\right) g_0(Y_j^{(0,\nu)}) \\
&\quad + \sum_{j=1}^{s}\sum_{k=1}^{m}\sum_{\nu=0}^{M} \left(B_{ij}^{(l,\mu)(k,\nu)} + \frac{1}{2}b_j^{(k,\nu)}\right) g_k(Y_j^{(k,\nu)}) \\
&\quad + \frac{1}{2}\sum_{j=1}^{s}\sum_{k=0}^{m}\sum_{\nu=0}^{M} g_k^{'}\left(Y_j^{(k,\nu)}\right) v_j^{(k,\nu)}, \qquad (16)
\end{aligned}
$$

and $v_i^{(l,\mu)}$ given in (13), (14), with $i=1,\dots,s$, $l=0,\dots,m$, $\mu=0,\dots,M$.

3. Stochastic B-series

We construct stochastic B-series for the solution of the ODE (3) and the numerical approximation (13), (14), (15), (16).

Following [4, 6, 17] we define the set $\mathcal{T}$ of $m+1-$colored rooted trees recursively as $\mathcal{T}=\{\emptyset\}\cup T_0\cup T_1\dots,T_m$ where

1. The graph $\bullet_l=[\emptyset]_l$ with only one vertex of color l belongs to T_l.
2. If $\tau_1,\dots,\tau_k\in\mathcal{T}$ then $\tau=[\tau_1,\dots,\tau_k]_l\in T_l$, where τ is the tree formed by joining the sub-trees $\tau_1,\dots,\tau_k$ each by a single branch to a common root of color l.

The number of nodes of a tree is denoted $|\tau|$, and the order of a tree is given by

$$
\rho(\emptyset)=0, \rho([\tau_1,\dots,\tau_k]_l) = \sum_{i=1}^{k}\rho(\tau_i) + \begin{cases} 1 \text{ for } l=0, \\ 1/2 \text{ otherwise.} \end{cases}
$$

Similar to [6] we also define $\alpha : \mathcal{T} \to \mathbb{Q}$ given by $\alpha(\emptyset) = 1$, $\alpha(\bullet_l) = 1$,

$$\alpha([\tau_1, \ldots, \tau_k]_l) = \frac{1}{r_1! \cdots r_n!} \prod_{j=1}^{k} \alpha(\tau_j),$$

where $r_1, \ldots, r_n$ count equal trees among $\tau_1, \ldots, \tau_k$.

For the SDE (3) and for every tree $t \in \mathcal{T}$ we define the corresponding elementary differential recursively: $F(\emptyset)(x) = x$, $F(\bullet_l)(x) = g_l(x)$, and for any $x =\in \mathrm{R}^d$

$$F([\tau_1, \ldots, \tau_k]_l)(x) = g_l^{(k)}(x)\left(F(\tau_1)(x), \ldots, F(\tau_k)(x)\right), \quad l = 1, \ldots, m.$$

In [6] a stochastic B-series for (3) is defined as a formal series

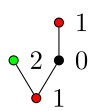
$$B(\phi, x; h) = \sum_{\tau \in T} \alpha(\tau) \cdot \phi(\tau)(h) \cdot F(\tau)(x).$$

From [6, Theorem 5] we know that the exact solution $X(h, x_0)$ can be written as a stochastic B-series: $X(h, x_0) = B(\varphi, x_0; h)$ with

$$\varphi(\emptyset)(h) = 1, \varphi(\bullet_l) = W_l(h), \varphi(\tau)(h) = \int_0^h \left(\prod_{j=1}^{k} \varphi(\tau_k)(s) \right) \circ dW_l(s) \quad (17)$$

where $\tau = [\tau_1, \ldots, \tau_k]_l$.

For example, for the tree $\tau = [\bullet_2, [\bullet_1]_0]_1$, let $\bullet 0 \in T_0$ (deterministic), $\bullet 1 \in T_1$, and $\bullet 2 \in T_2$ (stochastic). The graphical representation of τ and its corresponding terms are given below



$$|\tau| = 4, \rho(\tau) = 5/2, \alpha(\tau) = 1,$$
$$F(\tau) = g_1''(g_2, g_0' g_1),$$
$$\varphi(\tau)(h) = \int_0^h \left(W_2(s) \int_0^s W_1(u) du \right) \circ dW_1(s).$$

Let $Y(h) = B(\phi_Y, x_0; h)$ and $Z(h) = B(\phi_Z, x_0; h)$. From [6] (see Lemma 3 and its consequences on page 187) we know that

$$g_l(Y(h)) = \sum_{\tau \in T_l} \alpha(\tau) \cdot \phi^l_{\phi_Y}(\tau)(h) \cdot F(\tau)(x_0), \quad (18)$$

where $\phi^l_{\phi_Y}(\bullet_l)(h) = \phi^l_{\phi_Y}([\emptyset]_l)(h) \equiv 1$ and

$$\phi^l_{\phi_Y}(\tau)(h) = \prod_{j=1}^{k} \phi_Y(\tau_j)(h) \text{ if } \tau = [\tau_1, \ldots, \tau_k]_l \in T_l. \quad (19)$$

From [6, Lemma 4] we have

$$g_l^{'}\left(Y(h)\right)Z(h) = \sum_{\tau\in T_l} \alpha(\tau)\cdot\Gamma^l(\tau)(h)\cdot F(\tau)(x_0), \tag{20}$$

where $\Gamma^l(\bullet_l)(h) = \Gamma^l([\emptyset]_l)(h) \equiv 0$,

$$\Gamma^l(\tau)(h) = \sum_{i=1}^{k}\left(\prod_{\substack{j=1\\ j\neq i}}^{k}\phi_Y(\tau_j)(h)\right)\phi_Z(\tau_i)(h) \text{ if } \tau = [\tau_1,\ldots,\tau_k]_l \in T_l. \tag{21}$$

Proposition 1. *The numerical solution y^* as well as the stage values can be written as B-series*

$$y^* = B(\theta, y, h), \quad Y_i^{(l,\mu)} = B(\Phi_i^{(l,\mu)}, y, h), \quad v_i^{(l,\mu)} = B(\Psi_i^{(l,\mu)}, y, h),$$

for all $i = 1,\ldots,s$, $l = 0,\ldots,m$, $\mu = 0,\ldots,M$, with $\theta(\emptyset)\equiv 1$, $\Phi_i^{(l,\mu)}(\emptyset)\equiv 1$, $\Psi_i^{(l,\mu)}(\emptyset)\equiv 0$. For $\tau = [\tau_1,\ldots,\tau_{k_0}]_0$

$$\theta(\tau) = \sum_{i=1}^{s}\sum_{\mu=0}^{M} a_i^{(0,\mu)}\prod_{j_1=1}^{k_0}\Phi_i^{(0,\mu)}(\tau_{j_1}) + \sum_{i=1}^{s}\sum_{\mu=0}^{M}\sum_{i_1=1}^{k_0}\left(\prod_{\substack{j_1=1\\ j_1\neq i_1}}^{k_0}\Phi_i^{(0,\mu)}(\tau_{j_1})\right)\Psi_i^{(0,\mu)}(\tau_{i_1}), \tag{22}$$

$$\Phi_i^{(l,\mu)}(\tau) = \sum_{j=1}^{s}\sum_{\nu=0}^{M}\left(\frac{1}{2}a_j^{(0,\nu)} + A_{ij}^{(l,\mu)(0,\nu)}\right)\prod_{j_1=1}^{k_0}\Phi_j^{(0,\nu)}(\tau_{j_1})$$

$$+\frac{1}{2}\sum_{j=1}^{s}\sum_{\nu=0}^{M}\sum_{i_1=1}^{k_0}\left(\prod_{\substack{j_1=1\\ j_1\neq i_1}}^{k_0}\Phi_j^{(0,\nu)}(\tau_{j_1})\right)\Psi_j^{(0,\nu)}(\tau_{i_1}) \textit{ for } l = 0,\ldots,m,$$

$$\Psi_i^{(0,\mu)}(\tau) = -\sum_{j=1}^{s}\sum_{\nu=0}^{M} A_{ji}^{(0,\nu)(0,\mu)} a_j^{(0,\nu)}\prod_{j_1=1}^{k_0}\Phi_j^{(0,\nu)}(\tau_{j_1})$$

$$-\sum_{j=1}^{s}\sum_{\nu=0}^{M} A_{ji}^{(0,\nu)(0,\mu)}\sum_{i_1=1}^{k_0}\left(\prod_{\substack{j_1=1\\ j_1\neq i_1}}^{k_0}\Phi_j^{(0,\nu)}(\tau_{j_1})\right)\Psi_j^{(0,\nu)}(\tau_{i_1}),$$

$$\Psi_i^{(l,\mu)}(\tau) = -\sum_{j=1}^{s}\sum_{\nu=0}^{M} B_{ji}^{(0,\nu)(l,\mu)} a_j^{(0,\nu)} \prod_{j_1=1}^{k_0} \Phi_j^{(0,\nu)}(\tau_{j_1})$$

$$-\sum_{j=1}^{s}\sum_{\nu=0}^{M} B_{ji}^{(0,\nu)(l,\mu)} \sum_{i_1=1}^{k_0} \left(\prod_{\substack{j_1=1\\ j_1\neq i_1}}^{k_0} \Phi_j^{(0,\nu)}(\tau_{j_1}) \right) \Psi_j^{(0,\nu)}(\tau_{i_1}),\ \textit{for } l = 1,\dots,m.$$

For $\tau = [\tau_1, \dots, \tau_{k_0}]_k$, *k=1,…,m, we have*

$$\theta(\tau) = \sum_{i=1}^{s}\sum_{\mu=0}^{M} b_i^{(k,\mu)} \prod_{j_1=1}^{k_0} \Phi_i^{(k,\mu)}(\tau_{j_1}) + \sum_{i=1}^{s}\sum_{\mu=0}^{M}\sum_{i_1=1}^{k_0} \left(\prod_{\substack{j_1=1\\ j_1\neq i_1}}^{k_0} \Phi_i^{(k,\mu)}(\tau_{j_1}) \right) \Psi_i^{(k,\mu)}(\tau_{i_1}), \tag{23}$$

$$\Phi_i^{(l,\mu)}(\tau) = \sum_{j=1}^{s}\sum_{\nu=0}^{M} \left(\frac{1}{2} b_j^{(k,\nu)} + B_{ij}^{(l,\mu)(k,\nu)} \right) \prod_{j_1=1}^{k_0} \Phi_j^{(k,\nu)}(\tau_{j_1})$$

$$+\frac{1}{2}\sum_{j=1}^{s}\sum_{\nu=0}^{M}\sum_{i_1=1}^{k_0} \left(\prod_{\substack{j_1=1\\ j_1\neq i_1}}^{k_0} \Phi_j^{(k,\nu)}(\tau_{j_1}) \right) \Psi_j^{(k,\nu)}(\tau_{i_1})\ \textit{for } l = 0,\dots,m,$$

$$\Psi_i^{(0,\mu)}(\tau) = -\sum_{j=1}^{s}\sum_{\nu=0}^{M} A_{ji}^{(k,\nu)(0,\mu)} b_j^{(k,\nu)} \prod_{j_1=1}^{k_0} \Phi_j^{(k,\nu)}(\tau_{j_1})$$

$$-\sum_{j=1}^{s}\sum_{\nu=0}^{M} A_{ji}^{(k,\nu)(0,\mu)} \sum_{i_1=1}^{k_0} \left(\prod_{\substack{j_1=1\\ j_1\neq i_1}}^{k_0} \Phi_j^{(k,\nu)}(\tau_{j_1}) \right) \Psi_j^{(k,\nu)}(\tau_{i_1}),$$

$$\Psi_i^{(l,\mu)}(\tau) = -\sum_{j=1}^{s}\sum_{\nu=0}^{M} B_{ji}^{(k,\nu)(l,\mu)} b_j^{(k,\nu)} \prod_{j_1=1}^{k_0} \Phi_j^{(k,\nu)}(\tau_{j_1})$$

$$-\sum_{j=1}^{s}\sum_{\nu=0}^{M} B_{ji}^{(k,\nu)(l,\mu)} \sum_{i_1=1}^{k_0} \left(\prod_{\substack{j_1=1\\ j_1\neq i_1}}^{k_0} \Phi_j^{(k,\nu)}(\tau_{j_1}) \right) \Psi_j^{(k,\nu)}(\tau_{i_1}),\ \textit{for } l = 1,\dots,m.$$

Proof. We write y^*, $Y_i^{(l,\mu)}$ and $v_i^{(l,\mu)}$ as B-series:

$$
\begin{aligned}
y^* &= B(\theta, y; h) = \sum_{\tau \in T} \alpha(\tau) \cdot \theta(\tau)(h) \cdot F(\tau)(y), \qquad (24)\\
Y_i^{(l,\mu)} &= B(\Phi_i^{(l,\mu)}, y, h) = \sum_{\tau \in T} \alpha(\tau) \cdot \Phi_i^{(l,\mu)}(\tau)(h) \cdot F(\tau)(y),\\
v_i^{(l,\mu)} &= B(\Psi_i^{(l,\mu)}, y, h) = \sum_{\tau \in T} \alpha(\tau) \cdot \Psi_i^{(l,\mu)}(\tau)(h) \cdot F(\tau)(y).
\end{aligned}
$$

Using the definition of the method (15), (18), and (20) we get

$$
\begin{aligned}
y^* = y &+ \sum_{i=1}^{s} \sum_{\mu=0}^{M} a_i^{(0,\mu)} \sum_{\tau \in T_0} \alpha(\tau) \cdot \phi^0_{\Phi_i^{(0,\mu)}}(\tau)(h) \cdot F(\tau)(y)\\
&+ \sum_{i=1}^{s} \sum_{l=1}^{m} \sum_{\mu=0}^{M} b_i^{(l,\mu)} \sum_{\tau \in T_l} \alpha(\tau) \cdot \phi^l_{\Phi_i^{(l,\mu)}}(\tau)(h) \cdot F(\tau)(y)\\
&+ \sum_{i=1}^{s} \sum_{l=0}^{m} \sum_{\mu=0}^{M} \sum_{\tau \in T_l} \alpha(\tau) \cdot \Gamma_i^{(l,\mu)}(\tau)(h) \cdot F(\tau)(y),
\end{aligned}
$$

where, from (19) and (21), for any $l = 0, \dots, m$, $i = 1, \dots, s$, $\mu = 1, \dots, M$ we have $\phi^l_{\Phi_i^{(l,\mu)}}(\bullet_l)(h) = \phi^l_{\Phi_i^{(0,\mu)}}([\emptyset]_l)(h) \equiv 1$, $\Gamma^l(\bullet_l)(h) = \Gamma^l([\emptyset]_l)(h) \equiv 0$, and if $\tau = [\tau_1, \dots, \tau_{k_0}]_l \in T_l$

$$
\begin{aligned}
\phi^l_{\Phi_i^{(l,\mu)}}(\tau)(h) &= \prod_{j_1=1}^{k_0} \Phi_i^{(l,\mu)}(\tau_{j_1})(h),\\
\Gamma_i^{(l,\mu)}(\tau)(h) &= \sum_{i_1=1}^{k_0} \left(\prod_{\substack{j_1=1\\ j_1 \neq i_1}}^{k_0} \Phi_i^{(l,\mu)}(\tau_{j_1})(h) \right) \Psi_i^{(l,\mu)}(\tau_{i_1})(h).
\end{aligned}
$$

Comparing this expansion term by term with (24) we get (22) and (23). Similarly we can prove the formulas for the coefficients for $Y_i^{(l,\mu)}$ and $v_i^{(l,\mu)}$. $\square$

For example, for the tree $\tau = [\bullet_2, \bullet_0]_1$, the corresponding terms are

$$
|\tau| = 3, \rho(\tau) = 2, \alpha(\tau) = 1, F(\tau) = g_1''(g_2, g_0), \varphi(\tau)(h) = \int_0^h sW_2(s) \circ dW_1(s),
$$

$$\theta(\tau)(h) = \sum_{i,j,j_1=1}^{s} \sum_{\mu,\nu,\nu_1=1}^{M} b_i^{(1,\mu)} \left(\frac{1}{2} b_j^{(2,\nu)} + B_{ij}^{(1,\mu)(2,\nu)}\right) \left(\frac{1}{2} a_{j_1}^{(0,\nu_1)} + A_{ij_1}^{(1,\mu)(0,\nu_1)}\right)$$
$$- \sum_{i,j,j_1=1}^{s} \sum_{\mu,\nu,\nu_1=1}^{M} B_{ji}^{(2,\nu)(1,\mu)} b_j^{(2,\nu)} \left(\frac{1}{2} a_{j_1}^{(0,\nu_1)} + A_{ij_1}^{(1,\mu)(0,\nu_1)}\right)$$
$$- \sum_{i,j,j_1=1}^{s} \sum_{\mu,\nu,\nu_1=1}^{M} B_{j_1 i}^{(0,\nu_1)(1,\mu)} a_{j_1}^{(0,\nu_1)} \left(\frac{1}{2} b_j^{(2,\nu)} + B_{ij}^{(1,\mu)(2,\nu)}\right).$$

4. Convergence order

A time discrete approximation $\bar{X}_{h,x_0}(t_k)$ converges strongly, respectively in mean square, with order n to the solution $X(t_k, x_0)$ of the SHS (3) as $h \to 0$ if there exist $C, \delta_0 > 0$ such that we have $E\left[|\bar{X}_{h,x_0}(t_k) - X(t_k, x_0)|\right] \leq Ch^n$, respectively $\sqrt{E\left[|\bar{X}_{h,x_0}(t_k) - X(t_k, x_0)|^2\right]} \leq Ch^n$, for any fixed $t_k = kh$, $k = 0, \ldots, K$ and for all $0 < h \leq \delta_0$. Notice that by Jensens's inequality mean square convergence implies strong convergence of the same order.

Theorem 1. *We consider the SHS* (1) *written as the* $2d-$ *dimensional autonomous SDE* (3) *and the scheme given by* (15), (16), (13), *and* (14). *Let* $p \in \frac{1}{2}\mathbb{N}_0$ *and assume that* $H_r \in \mathcal{C}_P^{2p+2}(\mathbb{R}^{2d}, \mathbb{R})$ *and* H_r *has bounded second order partial derivatives,* $r = 0, \ldots, m$. *Suppose that the scheme is constructed such that we have* (7)- (8). *If*

$$\theta(\tau)(h) = \varphi(\tau)(h) + \mathcal{O}\left(h^{p+1/2}\right), \text{ for any } \tau \in T, \rho(\tau) \leq p \quad (25)$$
$$E\left[\theta(\tau)(h)\right] = E\left[\varphi(\tau)(h)\right] + \mathcal{O}\left(h^{p+1}\right), \text{ for any } \tau \in T, \rho(\tau) \leq p + 1/2, \quad (26)$$

where the $\mathcal{O}(\cdot)$ *notation refers to the* $L^2(\Omega)$ *norm and* $h \to 0$, *then the scheme is symplectic and has mean square order* p.

Proof. The scheme is symplectic because it is constructed using the generating function (4), [7, 14].

The proof of the strong order is a direct consequence of the similar result included in [6] (see also [5]). The idea is to compare the terms in the B-series expansions for the exact solution $X(h, x_0) = B(\varphi, x_0; h)$ in (17) and for the one-step-approximation $y^* = B(\theta, y; h)$ in (24) (see [17, Theorem 4.1] and [16, Theorem 5.8, Proposition 6.1] for the B-series expansions with remainders). Notice that conditions (7)- (8) imply $\theta(\tau)(h) = \mathcal{O}\left(h^{\rho(\tau)}\right)$ and $\phi^l_{\Phi_i^{(l,\mu)}}(\tau_1)(h) =$

$\mathcal{O}\left(h^{(\rho(\tau)-\rho([\emptyset]_l))}\right)$ for any $i=1,\ldots,s$, $l=0,\ldots,m$, $\mu=0,\ldots,M$, $\tau_1 \in T_l$, and any $\tau \in \mathcal{T}$, [6]. $\square$

Since the scheme given by (15), (16), (13), and (14) is symplectic, we can show that the conditions (25)-(26) should be checked for a substantially smaller number of trees. We first need to establish a preliminary result regarding a property of the scheme related to the Butcher product of two trees, [8].

The Butcher product $u \circ v$ of two trees $u, v \in \mathcal{T} \setminus \{\emptyset\}$, where $u = [u_1, \ldots, u_k]_l$, is defined as the tree obtained by adding v as a branch to the root of u: $u \circ v = [u_1, \ldots, u_k, v]_l$. For example, for $u = [\bullet_2]_0$ and $v = [\bullet_1, \bullet_1]_2$ we have $u \circ v = [\bullet_2, [\bullet_1, \bullet_1]_2]_0$ and $v \circ u = [[\bullet_2]_0, \bullet_1, \bullet_1]_2$.

In terms of the associated colored trees for SDE (3) we know from [1, Lemma 4] that for all $u \in T_l \setminus \{\emptyset\}$ and $v \in T_k \setminus \{\emptyset\}$, $l, k = 0, \ldots, m$ we have

$$\varphi(u)(h)\varphi(v)(h) = \varphi(u \circ v)(h) + \varphi(v \circ u)(h),$$

and a similar result holds for symplectic stochastic Runge Kutta schemes satisfying certain conditions. We extend this result for the symplectic scheme given by (15), (16), (13), and (14).

Lemma 1. *For all $u \in T_l \setminus \{\emptyset\}$ and $v \in T_k \setminus \{\emptyset\}$, $l, k = 0, \ldots, m$, for the symplectic scheme given by* (15)*,* (16)*,* (13) *and* (14) *we have*

$$\theta(u)(h)\theta(v)(h) = \theta(u \circ v)(h) + \theta(v \circ u)(h). \tag{27}$$

Proof. We choose any $u \in T_l \setminus \{\emptyset\}$ and $v \in T_k \setminus \{\emptyset\}$, $l, k = 0, \ldots, m$. We have several cases: $u = [u_1, \ldots, u_{k_1}]_0$ and $v = [v_1, \ldots, v_{k_2}]_0$; $u = [u_1, \ldots, u_{k_1}]_l$, $l = 1, \ldots, m$ and $v = [v_1, \ldots, v_{k_2}]_0$; $u = [u_1, \ldots, u_{k_1}]_l$ and $v = [v_1, \ldots, v_{k_2}]_k$, $l, k = 1, \ldots, m$. We prove (27) for $u = [u_1, \ldots, u_{k_1}]_0$ and $v = [v_1, \ldots, v_{k_2}]_0$. For the other two cases the proof is similar.

Since $u \circ v = [u_1, \ldots, u_{k_1}, v]_0$ and $v \circ u = [v_1, \ldots, v_{k_2}, u]_0$ we get

$$\theta(u \circ v) = \sum_{i=1}^{s}\sum_{\mu=0}^{M} a_i^{(0,\mu)} \Phi_i^{(0,\mu)}(v) \prod_{j_1=1}^{k_1} \Phi_i^{(0,\mu)}(u_{j_1}) + \sum_{i=1}^{s}\sum_{\mu=0}^{M} \left(\prod_{j_1=1}^{k_1} \Phi_i^{(0,\mu)}(u_{j_1})\right) \Psi_i^{(0,\mu)}(v)$$

$$+ \sum_{i=1}^{s}\sum_{\mu=0}^{M}\sum_{i_1=1}^{k_1} \Phi_i^{(0,\mu)}(v) \left(\prod_{\substack{j_1=1 \\ j_1 \neq i_1}}^{k_1} \Phi_i^{(0,\mu)}(u_{j_1})\right) \Psi_i^{(0,\mu)}(u_{i_1})$$

$$\theta(v \circ u) = \sum_{i=1}^{s} \sum_{\mu=0}^{M} a_i^{(0,\mu)} \Phi_i^{(0,\mu)}(u) \prod_{j_1=1}^{k_2} \Phi_i^{(0,\mu)}(v_{j_1}) + \sum_{i=1}^{s} \sum_{\mu=0}^{M} \left(\prod_{j_1=1}^{k_2} \Phi_i^{(0,\mu)}(v_{j_1}) \right) \Psi_i^{(0,\mu)}(u)$$

$$+ \sum_{i=1}^{s} \sum_{\mu=0}^{M} \sum_{i_1=1}^{k_2} \Phi_i^{(0,\mu)}(u) \left(\prod_{\substack{j_1=1 \\ j_1 \neq i_1}}^{k_2} \Phi_i^{(0,\mu)}(v_{j_1}) \right) \Psi_i^{(0,\mu)}(v_{i_1})$$

Replacing the formulas in Proposition 1 for $\Phi_i^{(0,\mu)}(v)$, $\Psi_i^{(0,\mu)}(v)$, $\Phi_i^{(0,\mu)}(u)$, $\Psi_i^{(0,\mu)}(u)$, $\theta(u)$, $\theta(v)$, and adding the expressions for $\theta(u \circ v)$ and $\theta(v \circ u)$, we can easily see that we get the same formula as for $\theta(u)\theta(v)$, so we have (27). □

Let $\tau \in \mathcal{T}$ and $\hat{\tau}$ the corresponding unrooted tree. We denote by $\hat{\mathcal{T}}(\hat{\tau})$ the set of rooted trees obtained from $\hat{\tau}$ by assigning one of its nodes as the root [1].

Theorem 2. *Assume that the scheme given by* (15)*,* (16)*,* (13)*, and* (14) *is constructed such that* $\theta(\tau)(h) = \mathcal{O}\left(h^{\rho(\tau)}\right)$ *for any* $\tau \in \mathcal{T}$*. Let* $p \in \frac{1}{2}\mathbb{N}_0$ *and* $\hat{\tau}$ *be an unrooted tree of order* $q \leq p$*. If* $\varphi(\tau_0)(h) = \theta(\tau_0)(h) + \mathcal{O}(h^{p+1/2})$ *for some rooted tree* $\tau_0 \in \hat{\mathcal{T}}(\hat{\tau})$ *and all rooted trees of order less than* q*, then* $\varphi(\tau)(h) = \theta(\tau)(h) + \mathcal{O}(h^{p+1/2})$ *for all* $\tau \in \hat{\mathcal{T}}(\hat{\tau})$*.*

Proof. From Lemma 1 we know that $\theta(u)(h)\theta(v)(h) = \theta(u \circ v)(h) + \theta(v \circ u)(h)$, for all $u, v \in \mathcal{T} \setminus \{\emptyset\}$. Using this, the proof is identical with the proof of [1, Theorem 5]. □

Application of the previous theorem reduces the order conditions substantially because we have to check them only for one tree from each set $\hat{\mathcal{T}}(\hat{\tau})$. For example, to prove that the scheme given by (15), (16), (13), and (14) applied to the SDE (3) with $m = 1$ has mean square order 2, instead of checking condition (25) for 18 rooted trees with order at most 2, it is enough to check condition (25) for 11 rooted trees that correspond to the unrooted trees with order at most 2. Moreover, the difference between the number of unrooted and rooted trees increases substantially with m [2, Table 1].

5. Examples of schemes for the strong approximation

As in [17] let (p_D, p_S) with $p_D \geq P_S$ denote the order of convergence of the scheme if it is applied to a deterministic or stochastic differential equation, respectively.

5.1. SHS with additive noise

We consider the special case of SHS with additive noise

$$
\begin{aligned}
dP &= -\nabla_Q H_0(P,Q)dt - \sum_{r=1}^{m} \sigma_r \circ dW_r(t), \quad P(0) = p_0, \\
dQ &= \nabla_P H_0(P,Q)dt + \sum_{r=1}^{m} \mu_r \circ dW_r(t), \quad Q(0) = q_0.
\end{aligned} \tag{28}
$$

We write the SHS (28) as the $2d-$ dimensional autonomous SDE (3) with $H_r(P,Q) = P^\top \mu_r + Q^\top \sigma_r$, where $\sigma_r, \mu_r \in \mathbb{R}^d$, $r = 1,\dots,m$, are constant vectors. We propose the order 1.5 method with $M = 0$ and for any $r = 1,\dots,m$, $i,j = 1,\dots,s$

$$
\begin{aligned}
&a_i^{(0,0)} = a_i h, \quad A_{ij}^{(0,0)(0,0)} = A_{ij}^{(0)} h, \quad A_{ij}^{(r,0)(0,0)} = A_{ij}^{(1)} h, \\
&b_i^{(r,0)} = b_i I_{(r)}, \quad B_{ij}^{(0,0)(r,0)} = B_{ij} \frac{I_{(r,0)}}{h}.
\end{aligned}
$$

All the other coefficients are equal with zero. Let $\Sigma_r = \nabla H_r = (\mu_r^\top, \sigma_r^\top)^\top$. Notice that that in this case the scheme becomes

$$
y^* = y + \sum_{i=1}^{s} \left(h a_i g_0(Y_i^{(0,0)}) + g_0^{'} \left(Y_i^{(0,0)} \right) v_i^{(0,0)} \right) + J^{-1} \sum_{i=1}^{s} \sum_{r=1}^{m} \Sigma_r b_i I_{(r)}, \tag{29}
$$

$$
Y_i^{(0,0)} = \frac{y + y*}{2} + h \sum_{j=1}^{s} A_{ij}^{(0)} g_0(Y_j^{(0,0)}) + J^{-1} \sum_{j=1}^{s} \sum_{k=1}^{m} B_{ij} \frac{I_{(k,0)}}{h} \Sigma_k, \tag{30}
$$

$$
\begin{aligned}
v_i^{(0,0)} &= -h^2 \sum_{j=1}^{s} A_{ji}^{(0)} a_j g_0(Y_j^{(0,0)}) - h J^{-1} \sum_{j=1}^{s} \sum_{k=1}^{m} A_{ji}^{(1)} b_j I_{(k)} \Sigma_k \\
&- h \sum_{j=1}^{s} A_{ji}^{(0)} g_0^{'} \left(Y_j^{(0,0)} \right) v_j^{(0,0)}.
\end{aligned} \tag{31}
$$

The method is characterized by the Butcher tableau

$$
\begin{array}{|c|c|}
A^{(0)} & A^{(1)} \\
\hline
 & B \\
\hline
a^\top & b^\top \\
\hline
\end{array}
$$

Theorem 3. *We consider the SHS* (28) *written as the* $2d-$ *dimensional autonomous SDE* (3) *and we assume that* $H_0 \in \mathcal{C}_P^5(\mathbb{R}^{2d}, \mathbb{R})$ *and* H_0 *has bounded second order partial derivatives. If the coefficients of the scheme given by* (29), (30), (31) *fulfill the equations*

$$a^\top e = 1, \quad b^\top e = 1, \quad b^\top A^{(1)} e = \frac{1}{2}, \tag{32}$$

$$a^\top B e = 1, \quad \frac{1}{3} a^\top (Be)^2 - b^\top \left(A^{(1)} (Be) \right) = \frac{1}{4}, \tag{33}$$

then the scheme attains order (2.1.5) for the strong approximation of the solution of the SDE (3)*. If in addition we have*

$$a^\top \left(A^{(0)} \left(A^{(0)} e \right) \right) - a^\top \left(A^{(0)} e \right)^2 = -\frac{1}{3}, \tag{34}$$

then the scheme will attains order (3, 1.5) instead of (2, 1.5).

The proof is included in Appendix A.

Coefficients for two stages strong schemes of order $(2, 1.5)$ and $(3, 1.5)$ are presented in Table 1.

0	0	0	0
0	0	$\frac{1}{2b_2}$	0
		0	0
		$\frac{3}{4}$	0
$-\frac{1}{3}$	$\frac{4}{3}$	$1\text{-}b_2$	b_2

0	0	0	0
$\frac{1}{2}$	0	$\frac{1}{2b_2}$	0
		0	0
		$\frac{3}{4}$	0
$-\frac{1}{3}$	$\frac{4}{3}$	$1\text{-}b_2$	b_2

Table 1: Coefficients of symplectic scheme with $p_D = 2$, $p_S = 1.5$ (left) and with $p_D = 3$, $p_S = 1.5$ (right). $b_2 \in \mathbb{R}$, $b_2 \neq 0$

5.2. SHS with scalar noise

For the SDE (3) with scalar noise ($m = 1$) and $d \geq 1$ we propose the order 1.5 method with $M = 0$ and the coefficients

$$\begin{aligned}
a_i^{(0,0)} &= a_i h, \quad b_i^{(1,0)} = b_i^{(1)} I_{(1)} + b_i^{(2)} \frac{I_{(10)}}{h} + b_i^{(3)} \frac{I_{(1,1,1)}}{h}, \\
A_{ij}^{(0,0)(0,0)} &= A_{ij}^{(0)} h, \quad B_{ij}^{(0,0)(1,0)} = B_{ij}^{(0)} \frac{I_{(1,0)}}{h}, \\
A_{ij}^{(1,0)(0,0)} &= A_{ij}^{(1)} h, \quad B_{ij}^{(1,0)(1,0)} = B_{ij}^{(1)} \sqrt{h}, \quad i, j = 1, \ldots, s.
\end{aligned}$$

The rest of the coefficients are zero. Notice for the scheme constructed with this choice of the coefficients, conditions (7)-(8) hold. The one-step approximation for the method with these coefficients becomes

$$
\begin{aligned}
y^* &= y + h\sum_{i=1}^{s} a_i g_0(Y_i^{(0,0)}) + \sum_{i=1}^{s}\left(b_i^{(1)} I_{(1)} + b_i^{(2)}\frac{I_{(1,0)}}{h} + b_i^{(3)}\frac{I_{(1,1,1)}}{h}\right) g_1\left(Y_i^{(1,0)}\right) \\
&+ \sum_{i=1}^{s}\left(g_0^{'}\left(Y_i^{(0,0)}\right) v_i^{(0,0)} + g_1^{'}\left(Y_i^{(1,0)}\right) v_i^{(1,0)}\right), \text{ where} \\
Y_i^{(0,0)} &= \frac{1}{2}(y+y^*) + h\sum_{j=1}^{s} A_{ij}^{(0)} g_0(Y_j^{(0,0)}) + \frac{I_{(1,0)}}{h}\sum_{j=1}^{s} B_{ij}^{(0)} g_1(Y_j^{(1,0)}), \\
Y_i^{(1,0)} &= \frac{1}{2}(y+y^*) + h\sum_{j=1}^{s} A_{ij}^{(1)} g_0(Y_j^{(0,0)}) + \sqrt{h}\sum_{j=1}^{s} B_{ij}^{(1)} g_1(Y_j^{(1,0)}), \\
v_i^{(0,0)} &= -h^2\sum_{j=1}^{s} A_{ji}^{(0)} a_j g_0(Y_j^{(0,0)}) - h\sum_{j=1}^{s} A_{ji}^{(1}\left(b_j^{(1)} I_{(1)} + b_j^{(2)}\frac{I_{(1,0)}}{h} + b_j^{(3)}\frac{I_{(1,1,1)}}{h}\right) \\
& g_1(Y_j^{(1,0)}) - h\sum_{j=1}^{s}\left(A_{ji}^{(0)} g_0^{'}\left(Y_j^{(0,0)}\right) v_j^{(0,0)} + A_{ji}^{(1)} g_1^{'}\left(Y_j^{(1,0)}\right) v_j^{(1,0)}\right), \\
v_i^{(0,1)} &= -I_{(1,0)}\sum_{j=1}^{s} B_{ji}^{(0)} a_j g_0(Y_j^{(0,0)}) - \sqrt{h}\sum_{j=1}^{s} B_{ji}^{(1}\left(b_j^{(1)} I_{(1)} + b_j^{(2)}\frac{I_{(1,0)}}{h} + b_j^{(3)}\frac{I_{(1,1,1)}}{h}\right) \\
& g_1(Y_j^{(1,0)}) - \frac{I_{(1,0)}}{h}\sum_{j=1}^{s} B_{ji}^{(0)} g_0^{'}\left(Y_j^{(0,0)}\right) v_j^{(0,0)} - \sqrt{h}\sum_{j=1}^{s} B_{ji}^{(1)} g_1^{'}\left(Y_j^{(1,0)}\right) v_j^{(1,0)}.
\end{aligned}
\tag{35}
$$

The method is characterized by the Butcher tableau

$A^{(0)}$	$B^{(0)}$		
$A^{(1)}$	$B^{(1)}$		
$a^\top$	$\left(b^{(1)}\right)^\top$	$\left(b^{(2)}\right)^\top$	$\left(b^{(3)}\right)^\top$

Theorem 4. *We consider the SHS* (1) *written as the* $2d-$ *dimensional autonomous SDE* (3) *and we assume that* $H_r \in \mathcal{C}_P^5(\mathbb{R}^{2d}, \mathbb{R})$ *and* H_r *has bounded second order partial derivatives,* $r = 0, \ldots, 1$. *If the coefficients of the scheme*

given by (35) *fulfill the equations*

$$\begin{aligned}
&a^\top e = 1, \quad \left(b^{(1)}\right)^\top e = 1, \quad \left(b^{(2)}\right)^\top e = 0, \quad \left(b^{(3)}\right)^\top e = 0,\\
&\left(b^{(1)}\right)^\top A^{(1)} e = \frac{1}{2}, \quad \left(b^{(3)}\right)^\top A^{(1)} e = 0, \quad \left(b^{(2)}\right)^\top A^{(1)} e - a^\top B^{(0)} e = -1,\\
&2\left(b^{(1)}\right)^\top \left(B^{(1)}\left(B^{(1)} e\right)\right) - \left(b^{(1)}\right)^\top \left(B^{(1)} e\right)^2 = -\frac{1}{4},\\
&2\left(b^{(2)}\right)^\top \left(B^{(1)}\left(B^{(1)} e\right)\right) - \left(b^{(2)}\right)^\top \left(B^{(1)} e\right)^2 = 0,\\
&2\left(b^{(3)}\right)^\top \left(B^{(1)}\left(B^{(1)} e\right)\right) - \left(b^{(3)}\right)^\top \left(B^{(1)} e\right)^2 = -\frac{1}{2},\\
&\frac{1}{3} a^\top \left(B^{(0)} e\right)^2 - \left(b^{(1)}\right)^\top \left(A^{(1)}\left(B^{(0)} e\right)\right) - \frac{2}{3}\left(b^{(2)}\right)^\top \left(A^{(1)}\left(B^{(0)} e\right)\right) = \frac{1}{4},
\end{aligned}$$

then the scheme attains order (2,1.5) for the strong approximation of the solution of the SDE (3). *If in addition we have*

$$a^\top \left(A^{(0)}\left(A^{(0)} e\right)\right) - a^\top \left(A^{(0)} e\right)^2 = -\frac{1}{3},$$

then the scheme will attains order (3, 1.5) instead of (2, 1.5).

Proof. We apply Theorem 1 with $p = 1.5$. Since the scheme is symplectic, Theorem 2 implies that we need to consider only the trees: $\tau_1 = \emptyset$, $\tau_2 = \bullet_1$, $\tau_3 = \bullet_0$, $\tau_4 = [\bullet_1]_1$, $\tau_5 = [\bullet_1]_0$, $\tau_6 = [[\bullet_1]_1]_1$, $\tau_7 = [\bullet_0]_0$, $\tau_8 = [\bullet_1, \bullet_1]_0$, $\tau_9 = [[\bullet_1]_1]_0$, $\tau_{10} = [[[\bullet_1]_1]_1]_1$, $\tau_{11} = [\bullet_1, \bullet_1, \bullet_1]_1$. The order conditions can be obtained proceeding as in the proof of Theorem 3. □

Coefficients for the order 1.5 strong schemes or order $(2, 1.5)$ and $(3, 1.5)$ are presented in Tables 2 and 3, respectively.

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

Table 2: Coefficients of symplectic scheme with $p_D = 2$, $p_S = 1.5$

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

Table 3: Coefficients of symplectic scheme with $p_D = 3$, $p_S = 1.5$

6. Numerical simulations

We illustrate numerically the performance of the proposed symplectic schemes for SHS with additive and multiplicative noise. As in [14], for the proposed implicit schemes with time step $h < 1$, we replace the random variable $\xi \sim N(0,1)$ with the bounded random variables ξ_h:

$$\xi_h = \begin{cases} -A_h & \text{if } \xi < -A_h \\ \xi & \text{if } |\xi| \leq A_h \\ A_h & \text{if } \xi > A_h, \end{cases} \quad \text{where } A_h = \sqrt{6|\ln h|}. \tag{36}$$

6.1. Hénon-Helles problem with two additive noises

Consider the SHS corresponding to the additive noise Hénon-Helles problem, [8, 9]:

$$\begin{aligned} dP_1(t) &= (\alpha Q_1^2(t) - Q_1(t) - \alpha Q_2^2(t))dt + \sigma_1 dW_1(t), \\ dP_2(t) &= -(Q_2(t) + 2\alpha Q_1(t)Q_2(t)dt + \sigma_2 dW_2(t), \\ dQ_1(t) &= P_1(t)dt, \\ dQ_2(t) &= P_2(t)dt. \end{aligned} \tag{37}$$

The Hamiltonians are

$$\begin{aligned} H_0(P,Q) &= \frac{1}{2}(P_1^2 + P_2^2) + \frac{1}{2}(Q_1^2 + Q_2^2) + \alpha\left(Q_1Q_2^2 - \frac{1}{3}Q_1^3\right) \\ H_1(P,Q) &= -\sigma_1 Q_1, \quad H_2(P,Q) = -\sigma_2 Q_2, P = (P_1,P_2)^\top, Q = (Q_1,Q_2)^\top \end{aligned}$$

We compare the symplectic schemes given by (29)-(31) and with the coefficients in the left Table 1 ($b_2 = 1/2$) with the Runge-Kutta order 1.5

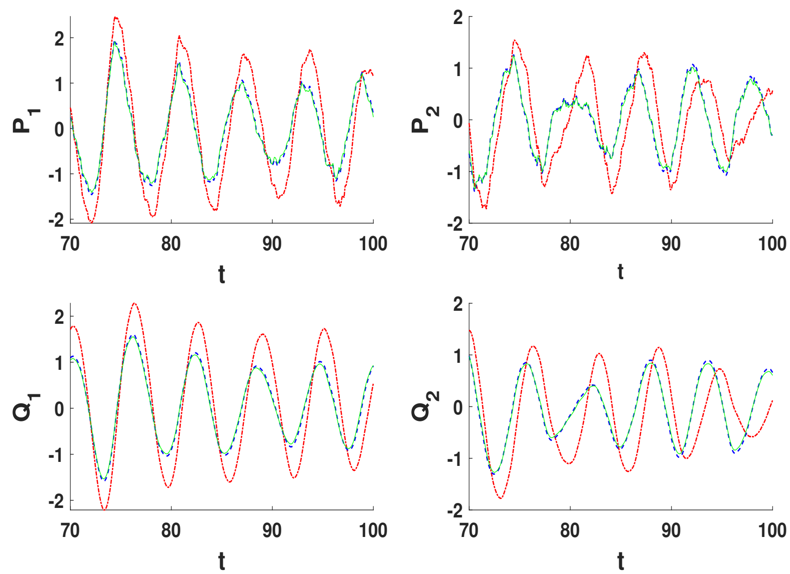


Figure 1: A sample trajectory of (37) with $h = 0.1$, on the time interval $0 \leq t \leq 100$: Runge-Kutta order 1.5 scheme SRA1 in [17] (red dash-dotted lines), symplectic schemes given by (29)-(31) and with the coefficients in the left Table 1 (blue dashed lines), the order 1.5 symplectic scheme (38) based on the truncation of S^3_ω (green solid lines).

scheme SRA1 in [17] and the order 1.5 symplectic midpoint scheme based on the truncation of S^3_ω given in Section 6.1 in [7]:

$$\begin{aligned} Y_{k+1} &= Y_k + J^{-1}\nabla H_0(Y_{k+\frac{1}{2}})h + \sum_{r=1}^{2}\Big(J^{-1}\nabla H_r \xi^{(r)}_{hk}\sqrt{h} \\ &\quad + J^{-1}\nabla G_{(r,0)}(Y_{k+\frac{1}{2}})(\bar{I}^h_{(r,0)} - \bar{I}^h_{(0,r)}) + J^{-1}\nabla G_{(r,r,0)}(Y_{k+\frac{1}{2}})\frac{h^2}{2}\Big), \qquad (38) \\ G_{(r,0)} &= \frac{1}{2}\left(J^{-1}\nabla H_r\right)^\top \nabla H_0, \quad G_{(r,r,0)} = \frac{1}{4}\left(J^{-1}\nabla H_r\right)^\top \nabla^2 H_0 J^{-1}\nabla H_r, \end{aligned}$$

where for each time step k, we have $Y_k = (P_k^\top, Q_k^\top)^\top$ and the arguments are everywhere $Y_{k+\frac{1}{2}} = (Y_{k+1} + Y_k)/2$. Here $\bar{I}^h_{(r,0)} = \frac{h^{\frac{3}{2}}}{2}(\xi^{(r)}_{hk} + \frac{\eta^{(r)}_{hk}}{\sqrt{3}})$ and $\bar{I}^h_{(0,r)} = \xi^{(r)}_{hk} h^{\frac{3}{2}} - \bar{I}^h_{(r,0)}$, where at each time step k, $\xi^{(r)}_{hk}$ and $\eta^{(r)}_{hk}$ are independent bounded random variables as given in (36).

We set the parameters as in [9] $\alpha = 1/16$, $\sigma_1 = \sigma_2 = 0.2$, $P_1(0) = P_2(0) =$

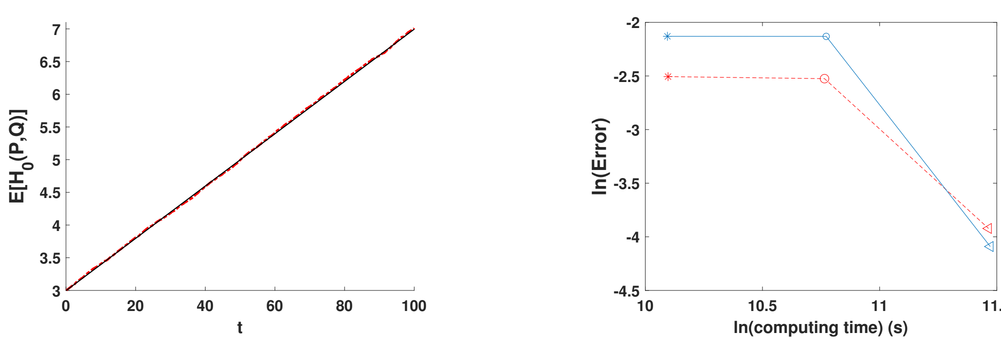
$Q_2(0) = 1$, $Q_1(0) = \sqrt{3}$. From Figure 1 we can see that the proposed symplectic scheme give similar results with the symplectic scheme (38), while the non-symplectic Runge-Kutta order 1.5 scheme SRA1 does not give accurate results for long term simulations for $h = 0.1$.

The moment of the Hamiltonian H_0 has a linear growth [9]:

$$E[H_0(P(t), Q(t))] = E[H_0(P(0), Q(0))] + (\sigma_1^2 + \sigma_2^2)t/2. \tag{39}$$

In Figure 2 (a) we plot the sample averages of Hamiltonian H_0 over 5000 sample trajectories representing the numerical solution given by the symplectic scheme (29)-(31) and with the coefficients in the left Table 1 ($b_2 = 1/2$), and the reference line with the slope $(\sigma_1^2 + \sigma_2^2)/2 = 0.04$. We can notice an excellent agreement with the linear growth of $E[H_0(P(t), Q(t))]$ in (39).

Figure 2 (b) compares the proposed symplectic scheme (29)-(31) to the symplectic scheme (38). The error represents the difference at $T = 100$ between the sample average of the Hamiltonian H_0 over 5000 sample trajectories and the exact value $E[H_0(P(100), Q(100))]$ calculated from (39). The two symplectic schemes are similar in terms of accuracy and CPU time. The computations can be done faster using parallel computing, but these schemes are implicit, so, for the same h, they require substantially larger computational time than the explicit scheme SRA1. However, the symplectic schemes



(a) (b)

Figure 2: (a) Linear growth of $E[H_0(P(t), Q(t))]$ for (37): the sample average of Hamiltonian H_0 with 5000 sample trajectories given by the symplectic scheme (29)-(31) (red dashed lines), h=0.1, and reference line with slope 0.04 (black solid lines). (b) Computing time versus error for the symplectic scheme (29)-(31) (red dashed lines) and the symplectic scheme (38) (blue plain lines) with various time steps for $T = 100$ and 5000 samples, *: $h = 0.02$; $\bigcirc$: $h = 0.01$; $\triangle$: $h = 0.0005$.

give smaller errors than the scheme SRA1. For example, for h=0.005, the error for the scheme SRA1 is 0.3169, while the errors for the scheme (29)-(31) and the scheme (38) are 0.0198 and 0.0167, respectively. Thus, to get a similar accuracy, we need to use a substantially lower h for the explicit scheme SRA1.

6.2. Kubo oscillator

We consider the Kubo oscillator, which is used to demonstrate the advantage of the stochastic symplectic schemes for long time computation [14], [7]:

$$\begin{aligned} dP(t) &= -aQ(t)dt - \sigma Q(t) \circ dW(t), \quad P(0) = p, \\ dQ(t) &= aP(t)dt + \sigma P(t) \circ dW(t), \quad Q(0) = q, \end{aligned} \tag{40}$$

where a and σ are constants. We set $a = 2$, $\sigma = 0.3$, $p = 1$ and $q = 0$. The Hamiltonian functions $H_0(P(t), Q(t)) = a\frac{P(t)^2+Q(t)^2}{2}$ and $H_1(P(t), Q(t)) = \sigma\frac{P(t)^2+Q(t)^2}{2}$ are preserved under the phase flow of the system, so the phase trajectory of (40) lies on the circle with the center at the origin and the radius $\sqrt{p^2+q^2}$, [14]. We can easily find the exact solution of the linear system (40), [7].

The Kubo system has one noise, so we consider the order 1.5 symplectic scheme (35) with the coefficients in Table 2. In Fig. 3 we compare this scheme with the order 1.5 strong Taylor scheme in section 10.4 in [10] and with the order 1.5 symplectic midpoint scheme based on the truncation of S^3_ω [7]:

$$\begin{aligned} Y_{k+1} &= Y_k + J^{-1}\nabla H_0(Y_{k+\frac{1}{2}})h + J^{-1}\nabla H_1 \xi_{hk}\sqrt{h} \\ &+ J^{-1}\nabla G_{(1,1,0)}(Y_{k+\frac{1}{2}})\frac{h^2}{2} + \frac{h\sqrt{h}\xi_{hk}}{2} J^{-1}\nabla G_{(1,1,1)}(Y_{k+\frac{1}{2}}) \\ &+ \frac{h^{3/2}}{2} J^{-1}\nabla G_{(1,1,1)}(Y_{k+\frac{1}{2}}) \left(\frac{\xi_{hk}^3}{3} - \xi_{hk}\right), \\ G_{(1,1,0)} &= \frac{1}{4}\left(J^{-1}\nabla H_1\right)^\top \nabla^2 H_0 \left(J^{-1}\nabla H_1\right), \\ G_{(1,1,1)} &= \frac{1}{4}(J^{-1}\nabla H_1)^T \nabla^2 H_1 (J^{-1}\nabla H_1). \end{aligned} \tag{41}$$

We can notice that for long-term simulation the non-symplectic order 1.5 strong Taylor scheme does not give accurate results for $h = 0.1$, but both

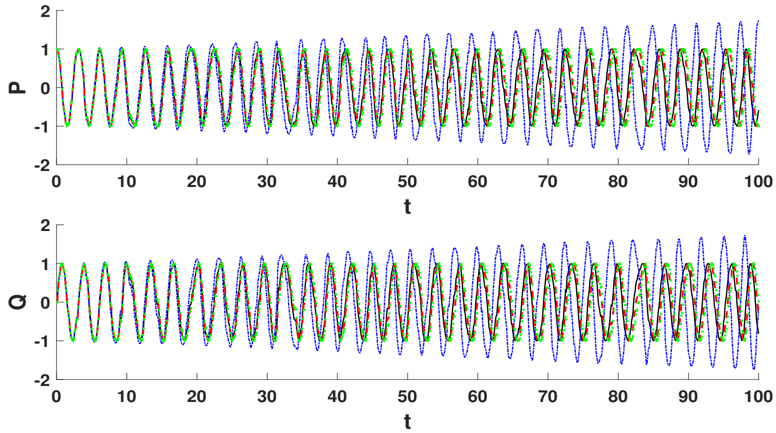


Figure 3: A sample trajectory of (40) with $a = 2$, $\sigma = 0.3$, $p = 1$, $q = 0$, $h = 0.1$, on the time interval $0 \leq t \leq 100$: exact solution (solid black lines), the symplectic scheme (35) (dashed red lines), the symplectic scheme (41) (green dotted lines), the order 1.5 strong Taylor scheme (dash-dotted blue lines).

order 1.5 symplectic schemes give approximations very close to the exact solution. Moreover, the sample phase trajectory of the non-symplectic order 1.5 strong Taylor scheme presented in Fig. 4 a deviates drastically from the exact $P(t)^2 + Q(t)^2 = 1$. However, from Fig. 4 b and Fig. 4 c we can see that the symplectic scheme (35) and the symplectic scheme (41) preserve the quadratic invariant.

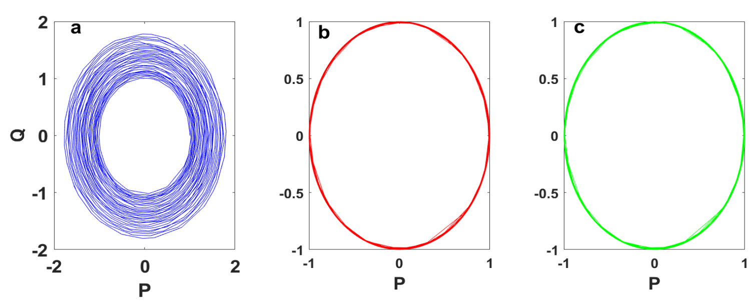


Figure 4: A sample phase trajectory of (40) with $a = 2$, $\sigma = 0.3$, $p = 1$, $q = 0$, $h = 0.1$ obtained by: (a) the order 1.5 strong Taylor scheme, (b) the symplectic scheme (35), (c) the symplectic scheme (41).

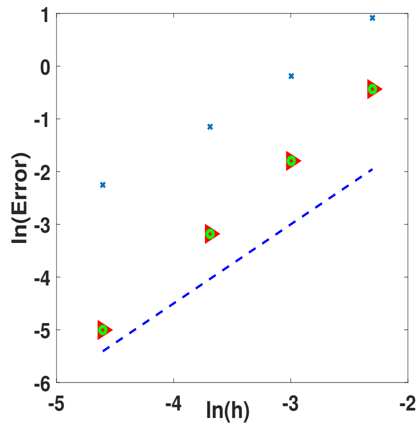


Figure 5: Convergence rate for (40) for the order 1.5 strong Taylor scheme (blue $\times$), the symplectic scheme (35) (red $\triangle$), the symplectic scheme (41) (green $\circ$).

Fig. 5 confirms that the schemes have the expected convergence rate. We have run 10 000 simulations and the mean square errors are calculated at $t = 100$. The order 1.5 reference line is plotted with a dashed line. We can notice that the errors obtained with the symplectic scheme (35) with the coefficients in Table 2 are very close to the one obtained with the symplectic midpoint scheme (41), and are lower than the errors obtained with the non-symplectic order 1.5 strong Taylor scheme.

7. Conclusions

We propose a new family of symplectic methods for SHSs by defining a generating function, so the schemes are symplectic by construction. Similar with symplectic SRK methods and the symplectic methods based on generating functions proposed in [7], the proposed methods are implicit for general SHSs and can be computationally expensive.

As in [12] we avoid computing Jacobian matrices and we evaluate the gradient of the generating function by computing the gradients and the Hessians of the Hamiltonian functions. We show that similar with symplectic SRK methods, to establish the order of convergence for these schemes we have to consider a substantially reduced number of trees. Unlike the symplectic schemes based on generating functions proposed in [7], these new schemes do not require derivatives of the Hamiltonian functions of order greater than 2, which makes them a good alternative for schemes with strong order 1.5 or more. Compared to the symplectic SRK methods, the proposed schemes are symplectic by construction, and they require only the conditions for the order of convergence, so it is easier to construct symplectic schemes with higher strong order.

For SHSs with scalar noise we construct a symplectic scheme with strong order 1.5. Unlike the deterministic case, for stochastic differential equations, order 2 is considered a high order because it is computationally expensive to generate approximations for multiple stochastic integrals of high order.

Acknowledgment

This work has been supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) through the grant DG-2018-04449.

Appendix A. Proof of Theorem 3

We apply Theorem 1 with $p = 1.5$. For additive noise the elementary differentials are zero for all trees that contains a sequence of nodes where at least one node directly succeeds a stochastic node. Thus we need to consider only the trees $\tau_1 = \emptyset$, $\tau_2 = \bullet_r$, $\tau_3 = \bullet_0$, $\tau_4 = [\bullet_r]_0$, $\tau_5 = [\bullet_0]_0$, $\tau_6 = [\bullet_{r_1}, \bullet_{r_2}]_0$. We have the following order conditions.

$$\phi(\tau_1) = \theta(\tau_1) \equiv 1, \quad \phi(\tau_2) = \theta(\tau_2) \Leftrightarrow W_r(h) = b^\top e I_{(r)} \Leftrightarrow b^\top e = 1,$$
$$\phi(\tau_3) = h, \quad \theta(\tau_3) = a^\top e h, \quad \phi(\tau_3) = \theta(\tau_3) \Leftrightarrow a^\top e = 1.$$

Thus we got the first two conditions (32). Replacing we obtain

$$\Phi_j^{(0,0)}(\bullet_0) = \sum_{j_1=1}^{s} \left(\frac{1}{2} a_{j_1}^{(0,0)} + A_{jj_1}^{(0,0)(0,0)} \right) = \frac{h}{2} + h \sum_{j_1=1}^{s} A_{jj_1}^{(0)}, \tag{A.1}$$

$$\Phi_j^{(0,0)}(\bullet_r) = \sum_{j_1=1}^{s} \left(\frac{1}{2} b_{j_1}^{(r,0)} + B_{jj_1}^{(0,0)(r,0)} \right) = \frac{1}{2} I_{(r)} + \sum_{j_1=1}^{s} B_{jj_1} \frac{I_{(r,0)}}{h}, \tag{A.2}$$

$$\Psi_j^{(0,0)}(\bullet_0) = -\sum_{j_1=1}^{s} a_{j1}^{(0,0)} A_{j_1 j}^{(0,0)(0,0)} = -\sum_{j_1=1}^{s} a_{j_1} A_{j_1 j}^{(0)} h^2, \tag{A.3}$$

$$\Psi_j^{(0,0)}(\bullet_r) = -\sum_{j_1=1}^{s} b_{j1}^{(r,0)} A_{j_1 j}^{(r,0)(0,0)} = -\sum_{j_1=1}^{s} b_{j_1} A_{j_1 j}^{(1)} I_{(r)} h. \tag{A.4}$$

Using (A.2), (A.4) we get

$$\phi(\tau_4) = J_{(r,0)}, \quad \theta(\tau_4) = \sum_{j=1}^{s} \left(a_j^{(0,0)} \Phi_j^{(0,0)}(\bullet_r) + \Psi_j^{(0,0)}(\bullet_r) \right)$$
$$= \sum_{j=1}^{s} a_j^{(0,0)} \sum_{j_1=1}^{s} \left(\frac{1}{2} b_{j_1}^{(r,0)} + B_{jj_1}^{(0,0)(r,0)} \right) - \sum_{j=1}^{s} \sum_{j_1=1}^{s} b_{j1}^{(r,0)} A_{j_1 j}^{(r,0)(0,0)}$$
$$= \frac{h}{2} I_{(r)} + a^\top B e I_{(r,0)} - b^\top A^{(1)} e h I_{(r)}.$$

Thus $\phi(\tau_4) = \theta(\tau_4)$ reduces to $b^\top A^{(1)} e = \frac{1}{2}$, $a^\top B^{(0)} e = 1$.

We consider now the trees of order 2. Using (A.1) and (A.3) we get

$$\phi(\tau_5) = J_{(0,0)} = I_{(0,0)} = \frac{h^2}{2}, \quad \theta(\tau_5) = \sum_{i=1}^{s} a_i^{(0,0)} \Phi_i^{(0,0)}(\bullet_0) + \sum_{i=1}^{s} \Psi_i^{(0,0)}(\bullet_0)$$
$$= \frac{h^2}{2} \sum_{j=1}^{s} a_j \sum_{j_1=1}^{s} a_{j_1} + \sum_{j,j_1=1}^{s} a_j A_{jj_1}^{(0)} h^2 - \sum_{j,j_1=1}^{s} a_{j_1} A_{j_1 j}^{(0)} h^2 = \frac{h^2}{2}.$$

If $r_1 \neq r_2$, replacing from (A.2), (A.4), we get

$$E[\phi(\tau_6)] = E[J_{(r_1,r_2,0)} + J_{(r_2,r_1,0)}] = 0,$$
$$\theta(\tau_6) = \sum_{j=1}^{s} a_j^{(0,0)} \Phi_j^{(0,0)}(\bullet_{r_1}) \Phi_j^{(0,0)}(\bullet_{r_2}) + \sum_{j=1}^{s} \Phi_j^{(0,0)}(\bullet_{r_1}) \Psi_j^{(0,0)}(\bullet_{r_2})$$
$$+ \sum_{j=1}^{s} \Phi_j^{(0,0)}(\bullet_{r_2}) \Psi_j^{(0,0)}(\bullet_{r_1})$$
$$= \frac{h}{4} I_{(r_1)} I_{(r_2)} + \frac{1}{2} \left(I_{(r_1)} I_{(r_2,0)} + I_{(r_2)} I_{(r_1,0)} \right) a^\top \left(B^{(0)} e \right)$$
$$+ \frac{1}{h} I_{(r_1,0)} I_{(r_2,0)} a^\top \left(B^{(0)} e \right)^2 - h I_{(r_1)} I_{(r_2)} b^\top \left(A^{(1)} e \right)$$
$$- \left(I_{(r_1)} I_{(r_2,0)} + I_{(r_2)} I_{(r_1,0)} \right) b^\top \left(A^{(1)} \left(B^{(0)} e \right) \right).$$

Since we have $E[I_{(r1)} I_{(r2)}] = E[(I_{(r_1)} I_{(r_2,0)}] = E[I_{(r_1,0)} I_{(r_2,0)}] = 0$, [10], we obtain $E[\theta(\tau_6)] = E[\phi(\tau_6)] = 0$.

If $r_1 = r_2 = r$ we have

$$E[\phi(\tau_6)] = E[2J_{(r,r,0)}] = 2E[I_{(r,r,0)}] + E[I_{(0,0)}] = 0 + \frac{h^2}{2} = \frac{h^2}{2},$$

$$
\begin{aligned}
\theta(\tau_6) &= \sum_{j=1}^{s} a_j^{(r,0)} \left(\Phi_j^{(0,0)}(\bullet_r)\right)^2 + 2\sum_{j=1}^{s} \Phi_j^{(0,0)}(\bullet_r)\Psi_j^{(0,0)}(\bullet_r) \\
&= -\frac{h}{4} I_{(r)}^2 + I_{(r)} I_{(r,0)} \left(1 - 2b^\top \left(A^{(1)}\left(Be\right)\right)\right) + \frac{1}{h} I_{(r,0)}^2 a^\top \left(B^{(0)} e\right)^2 .
\end{aligned}
$$

From $E[I_{(r)}^2] = h$, $E[I_{(r)} I_{(r,0)}] = \frac{h^2}{2}$, $E[I_{(r,0)}^2] = \frac{h^3}{3}$, [10], we obtain

$$
E[\theta(\tau_6)] = h^2 \left(\frac{1}{4} - b^\top \left(A^{(1)}\left(Be\right)\right) + \frac{1}{3} a^\top \left(Be\right)^2 \right),
$$

and we get the second condition (33).

References

[1] S. Anmarkrud, A. Kværnø, Order conditions for stochastic Runge-Kutta methods preserving quadratic invariants of Stratonovich SDEs, J. Comput. Appl. Math. 316 (2017) 40–46.

[2] C. Anton, Explicit pseudo-symplectic Runge-Kutta methods for stochastic Hamiltonian systems, Applied Numerical Mathematics 185 (2023) 18–37.

[3] K. Burrage, P. Burrage, Low rank Runge - Kutta methods, symplecticity and stochastic Hamiltonian problems with additive noise, J. Comp. Appl. Math. 236 (2012) 3920–3930.

[4] K. Burrage, P.M. Burrage, High strong order explicit Runge-Kutta methods for stochastic ordinary differential equations, Appl. Numer. Math. 20 (1996) 1–21.

[5] K. Burrage, T. Tian, Implicit stochastic Runge–Kutta methods for stochastic differential equations, BIT Numerical Mathematics 44 (2004) 21–39.

[6] K. Debrabant, A. Kværnø, B–series analysis of stochastic Runge–Kutta methods that use an iterative scheme to compute their internal stage values, SIAM Journal on Numerical Analysis 47 (2009) 181–203.

[7] J. Deng, C. Anton, Y. Wong, High-order symplectic schemes for stochastic Hamiltonian systems, Commun. Cumput. Phys. 16 (2014) 169–200.

[8] E. Hairer, C. Lubich, G. Wanner, Geometric numerical integration: structure-preserving algorithms for ordinary differential equations, Springer, Berlin, 2006.

[9] M. Han, Q. Ma, X. Ding, High-order stochastic symplectic partitioned Runge-Kutta methods for stochastic Hamiltonian systems with additive noise, Applied Mathematics and Computation 346 (2019) 575–593.

[10] P. Kloeden, E. Platen, Numerical solutions of stochastic differential equations, Springer–Verlag, Berlin, 1992.

[11] Q. Ma, D. Ding, X. Ding, Symplectic conditions and stochastic generating functions of stochastic Runge–Kutta methods for stochastic Hamiltonian systems with multiplicative noise, Appl. Math. Comp. 219 (2012) 635–643.

[12] J. Makazaga, A. Murua, A new class of symplectic integration schemes based on generating functions, Numer. Math. 113 (2009) 631–642.

[13] S. Miesbach, H.J. Pesch, Symplectic phase flow approximation for the numerical integration of canonical systems, Numer. Math. 61 (1992) 501–521.

[14] G.N. Milstein, Y.M. Repin, M.V. Tretyakov, Numerical methods for stochastic systems preserving symplectic structure, SIAM J. Numer. Anal. 40 (2002) 1583–1604.

[15] G.N. Milstein, Y.M. Repin, M.V. Tretyakov, Symplectic integration of Hamiltonian systems with additive noise, SIAM J. Numer. Anal. 39 (2002) 2066–2088.

[16] A. Rößler, Rooted tree analysis for order conditions of stochastic Runge–Kutta methods for the weak approximation of stochastic differential equations, Stochastic Analysis and Applications 24 (2006) 97–134.

[17] A. Rößler, Runge-Kutta methods for the strong approximation of solutions of stochastic differential equations, SIAM J. Numer. Anal. 48 (2010) 922–952.

[18] P. Wang, J. Hong, D. Xu, Construction of symplectic Runge-Kutta methods for stochastic Hamiltonian systems, Communications in Computational Physics 21 (2017) 237–270.

List with figure captions

Fig. 1: A sample trajectory of (37) with $h = 0.1$, on the time interval $0 \leq t \leq 100$: Runge-Kutta order 1.5 scheme SRA1 in [17] (red dash-dotted lines), symplectic schemes given by (29)-(31) and with the coefficients in the left Table 1 (blue dashed lines), the order 1.5 symplectic scheme (38) based on the truncation of S^3_ω (green solid lines).

Fig. 2: (a) Linear growth of $E[H_0(P(t), Q(t))]$ for (37): the sample average of Hamiltonian H_0 with 5000 sample trajectories given by the symplectic scheme (29)-(31) (red dashed lines), h=0.1, and reference line with slope 0.04 (black solid lines). (b) Computing time versus error for the symplectic scheme (29)-(31) (red dashed lines) and the symplectic scheme (38) (blue plain lines) with various time steps for $T = 100$ and 5000 samples, *: $h = 0.02$; $\bigcirc$: $h = 0.01$; $\triangle$: $h = 0.0005$.

Fig. 3: A sample trajectory of (40) with $a = 2$, $\sigma = 0.3$, $p = 1$, $q = 0$, $h = 0.1$, on the time interval $0 \leq t \leq 100$: exact solution (solid black lines), the symplectic scheme (35) (dashed red lines), the symplectic scheme (41) (green dotted lines), the order 1.5 strong Taylor scheme (dash-dotted blue lines).

Fig. 4: A sample phase trajectory of (40) with $a = 2$, $\sigma = 0.3$, $p = 1$, $q = 0$, $h = 0.1$ obtained by: (a) the order 1.5 strong Taylor scheme, (b) the symplectic scheme (35), (c) the symplectic scheme (41).

Fig. 5: Convergence rate for (40) for the order 1.5 strong Taylor scheme (blue $\times$), the symplectic scheme (35) (red $\triangle$), the symplectic scheme (41) (green $\circ$).