

Explicit Pseudo-Symplectic Runge-Kutta Methods for Stochastic Hamiltonian Systems

Cristina Anton

NOTICE: This is the peer reviewed version of the following article: Cristina A. (2023). Explicit pseudo-symplectic Runge-Kutta methods for stochastic Hamiltonian systems. *Applied Numerical Mathematics*, 185, 2023, 18-37, which has been published in final form at <http://dx.doi.org/10.1016/j.apnum.2022.11.013>.

Permanent link to this version <https://hdl.handle.net/20.500.14078/3400>

License © 2023 Elsevier. Licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 user license <http://creativecommons.org/licenses/by-nc-nd/4.0/>

Explicit Pseudo-Symplectic Runge-Kutta 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 give conditions for stochastic Runge-Kutta methods to near preserve quadratic invariants, and we discuss the associated simplified order conditions. For stochastic Hamiltonian systems we propose a systematic approach to construct explicit stochastic Runge-Kutta pseudo-symplectic schemes. Our approach is based on colored trees and B-series. We construct some pseudo-symplectic stochastic Runge-Kutta methods with strong convergence order, and we illustrate numerically the long term performance of the proposed schemes.

Keywords: stochastic Hamiltonian systems, stochastic Runge-Kutta methods, quadratic invariants, symplectic integration, pseudo-symplectic method.

1. Introduction

Recently, several results regarding Runge-Kutta methods preserving invariants of deterministic ordinary differential equations [14] have been extended to stochastic Runge-Kutta (SRK) methods for stochastic differential equations (SDE). Conditions for SRK methods to preserve quadratic invariants are given [16], and in [1] a result obtained for deterministic ordinary differential equations in [25] is extended to SDE, and it is proved that the conditions for preserving quadratic invariants are simplifying assumptions such that only one order condition per unrooted tree remains to be satisfied. Symplectic SRK methods for stochastic Hamiltonian systems (SHSs)

*corresponding author

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

are studied in [7, 26, 15]. However, these SRK methods are implicit and can be computationally expensive.

In the deterministic case it was shown that Runge-Kutta methods with vanishing stability matrix preserve the quadratic invariants of the system [10], and they are symplectic when applied to Hamiltonian systems [24]. Under certain assumptions, for numerical methods being symplectic implies the preservation of quadratic invariants of linear Hamiltonian systems [13], and if a numerical method preserves quadratic invariants generated by those of the original system, then the method is symplectic when applied to a Hamiltonian system [5]. In the stochastic case, the connection between the preservation of quadratic invariants and symplectic structure is studied in [16, Lemma 4.1] for SHSs, and it is proved that, under certain assumptions, a SRK method that preserve quadratic invariants is symplectic. Here for SRK methods we find similar conditions for near preservation of quadratic invariants and pseudo-symplecticity.

Let $(\Omega, \mathcal{F}, P)$ be a probability space with a filtration $\{\mathcal{F}_t\}_{t \geq 0}$. We consider a d -dimensional autonomous stochastic differential equation in the sense of Stratonovich

$$dX(t) = \sum_{l=0}^m g_l(X(t)) \circ dW_l(t), \quad X(0) = x_0 \in \mathbb{R}^d, \quad t \in [0, T], \quad (1)$$

where $W_0(t) = t$ and $W_l(t)$, $l = 1, \dots, m$ are independent Wiener processes with respect to $\{\mathcal{F}_t\}_{t \geq 0}$. We assume that $g_l : \mathbb{R}^d \rightarrow \mathbb{R}^d$, $l = 0, \dots, m$, are differentiable functions and we define the modified drift

$$\tilde{g}_0^i(x) = g_0^i(x) + \frac{1}{2} \sum_{j=1}^d \sum_{k=1}^m g_k^j(x) \frac{\partial g_k^i}{\partial x^j}(x), \quad i = 1, \dots, d. \quad (2)$$

We suppose that $\tilde{g}_0$ and g_l , $l = 1, \dots, m$ satisfy a global Lipschitz condition

$$|\tilde{g}_0(x) - \tilde{g}_0(y)| + \sum_{l=1}^m |g_l(x) - g_l(y)| \leq C|x - y|, \quad (3)$$

for all $x, y \in \mathbb{R}^d$, with some constant $C > 0$. Here, and in the rest of the paper, $|\cdot|$ denotes the Euclidean norm. Under these assumptions a solution of the SDE (1) exists and it is unique [18]. In addition to this, we suppose that for a certain order k , $g_l \in C^k(\mathbb{R}^d, \mathbb{R})$, $l = 0, \dots, m$ are k times continuously

differentiable functions for which all partial derivatives up to order k have polynomial growth: $g_l \in C_P^k(\mathbb{R}^d, \mathbb{R}) := \{g \in C^k(\mathbb{R}^d, \mathbb{R}), |\partial_x^i g(x)| \leq K(1 + |x|^{2r}) \text{ for all } x \in \mathbb{R}^d \text{ and any partial derivative } \partial_x^i g \text{ of order } i \leq k, \text{ where } K > 0 \text{ and } r \in \mathbb{N} \text{ depend on } g\}$. The order k will be specified later and it depends on the numerical scheme applied to approximate the solutions of the SDE (1).

We define the s stages stochastic Runge-Kutta methods

$$X_{n+1} = X_n + \sum_{l=0}^m \sum_{i=1}^s \gamma_i^{(l)} g_l(x_i) \quad X_0 = x_0, \quad (4)$$

$$x_i = X_n + \sum_{l=0}^m \sum_{j=1}^s Z_{ij}^{(l)} g_l(x_j), \quad x_1 = X_n, \quad (5)$$

where the coefficients $\gamma_i^{(l)}$, $Z_{ij}^{(l)}$ include random variables which depend on the step size $h = T/K$, $n = 0, 1, \dots, K-1$.

A differentiable function $I : \mathbb{R}^d \rightarrow \mathbb{R}$ is called an invariant of (1) if it satisfies $\nabla I(x)^\top g_l(x) = 0$, for $l = 0, \dots, m$. We are interested in quadratic invariants $I(x) = x^\top Cx$, where $C \in \mathcal{M}_{d,d}(\mathbb{R})$ is a symmetric constant matrix. In [16, Theorem 2.2] it has been proven that if the method (4)-(5) satisfies

$$\gamma_i^{(l)} Z_{ij}^{(k)} + \gamma_j^{(k)} Z_{ji}^{(l)} = \gamma_i^{(l)} \gamma_j^{(k)}, \quad (6)$$

for all $i, j = 1, \dots, s$ and $k, l = 0, \dots, m$, then it preserves quadratic invariants of (1).

Here we focus on explicit SRK methods and instead of (6) we find conditions for near preservation of quadratic invariants, i.e. such that for some $p \in \frac{1}{2}\mathbb{N}_0$ we have $X_{n+1}^\top C X_{n+1} = X_n^\top C X_n + \mathcal{O}(h^{p+1/2})$. For SHSs we use similar conditions to construct explicit pseudo-symplectic [21] SRK schemes that preserve the symplectic structure with certain accuracy.

A $2d$ -dimensional SHS [20] in the sense of Stratonovich has the form:

$$\begin{aligned} dP &= -\nabla_Q H^{(0)}(P, Q)dt - \sum_{r=1}^m \nabla_Q H^{(r)}(P, Q) \circ dW_r(t), \quad P(0) = p_0, \\ dQ &= \nabla_P H^{(0)}(P, Q)dt + \sum_{r=1}^m \nabla_P H^{(r)}(P, Q) \circ dW_r(t), \quad Q(0) = q_0, \end{aligned} \quad (7)$$

where $P(t)$, $Q(t)$, p_0, q_0 are d -dimensional vectors, $H^{(r)} : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$, $r = 0, \dots, m$, are differentiable functions with bounded second order partial derivatives such that $H^{(r)} \in \mathcal{C}_P^k(\mathbb{R}^{2d}, \mathbb{R})$ for a certain order k , $\nabla_Q H^{(r)}$

and $\nabla_P H^{(r)}$ denote the gradients with respect to Q and P respectively, and $W_r, r = 1, \dots, m$ are independent standard Wiener processes. The solution $X(t, x_0) = (P(t, x_0), Q(t, x_0)) \in \mathbb{R}^{2d}$, $x_0 = (p_0, q_0) \in \mathbb{R}^{2d}$, of (7) preserves the symplectic structure [19]:

$$\left(\frac{\partial X(t, x_0)}{\partial x_0} \right)^T J \left(\frac{\partial X(t, x_0)}{\partial x_0} \right) = J \quad \text{with} \quad J = \begin{pmatrix} 0 & I_d \\ -I_d & 0 \end{pmatrix}, \quad (8)$$

where I_d denotes the d -dimensional identity matrix. In [26, Theorem 2.2] it has been proven that if the method (4)-(5) satisfies (6) then it preserves the symplectic structure.

A one-step numerical method $X_1(h, x_0)$ of mean square order N for the SHS (7) is called a pseudo-symplectic method of mean square order (N, M) , $M > N$, if [21]

$$\begin{aligned} & \left\| \left(\frac{\partial X_1(h, x_0)}{\partial x_0} \right)^T J \left(\frac{\partial X_1(h, x_0)}{\partial x_0} \right) - J \right\|_{L^2(\Omega)} \\ &= \left(E \left\| \left(\frac{\partial X_1(h, x_0)}{\partial x_0} \right)^T J \left(\frac{\partial X_1(h, x_0)}{\partial x_0} \right) - J \right\|^2 \right)^{\frac{1}{2}} = \mathcal{O}(h^{M+1}). \end{aligned} \quad (9)$$

Here h is the time step, E denotes the expectation, and $\| \cdot \|$ denotes the Frobenius norm of a matrix.

In the deterministic case, rooted trees and B-series [9], [14] are used to construct pseudo-symplectic Runge-Kutta schemes [4]. For SHS with additive noise, pseudo symplectic schemes are proposed in [21], including pseudo-symplectic Runge-Kutta methods. In [2] we present a systematic method based on the generating functions [12] to obtain explicit pseudo-symplectic methods of any order for general SHSs with multiplicative or additive noise. These pseudo-symplectic methods depend on higher order derivatives of the Hamiltonian functions $H^{(r)}$, $r = 0, \dots, m$. Here we use a different approach based on colored trees and B-series [8, 11, 23] and we construct pseudo-symplectic Runge-Kutta methods for SHSs.

In the next section we present some preliminary results regarding colored trees and B-series for SDEs. Conditions for near preservation of quadratic invariants are given in section 3. The method to construct pseudo-symplectic schemes and a theoretical study of the pseudo-symplectic order are presented in section 4. In section 5 we construct some explicit pseudo-symplectic

schemes and section 6 includes numerical experiments illustrating the performance of the proposed schemes.

2. Notations and preliminary results

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

For any integer $s > 0$, any function $f : \mathbb{R}^k \rightarrow \mathbb{R}^l$, $f \in C^s(\mathbb{R}^k, \mathbb{R}^l)$, and any $x \in \mathbb{R}^k$, $t_i = (t_{i1}, \dots, t_{ik})^T \in \mathbb{R}^k$, $i = 1, \dots, s$ we have the differential $f^{(s)}(x) \cdot (t_1, \dots, t_s)$ of order s of f at the point x and in the directions t_i , $i = 1, \dots, s$:

$$f^{(s)}(x) \cdot (t_1, \dots, t_s) = \begin{pmatrix} \sum_{\alpha \in \{1,\dots,k\}^s} \frac{\partial^s f_1(x_1, \dots, x_k)}{\partial x_{\alpha_1} \dots \partial x_{\alpha_s}} t_{1\alpha_1} \dots t_{s\alpha_s} \\ \dots \\ \sum_{\alpha \in \{1,\dots,k\}^s} \frac{\partial^s f_l(x_1, \dots, x_k)}{\partial x_{\alpha_1} \dots \partial x_{\alpha_s}} t_{1\alpha_1} \dots t_{s\alpha_s} \end{pmatrix}. \quad (10)$$

For $F : \mathbb{R}^k \rightarrow \mathcal{M}_{l,j}(\mathbb{R})$, $F = (F_1, \dots, F_j)$, with columns $F_1, \dots, F_j \in C^s(\mathbb{R}^k, \mathbb{R}^l)$, and any $x \in \mathbb{R}^k$, $t_i \in \mathbb{R}^k$, $i = 1, \dots, s$ we define

$$D^s F(x) \cdot (t_1, \dots, t_s) = (F_1^{(s)}(x) \cdot (t_1, \dots, t_s), \dots, F_j^{(s)}(x) \cdot (t_1, \dots, t_s)) \in \mathcal{M}_{l,j}(\mathbb{R}). \quad (11)$$

For any index $\alpha = (j_1, j_2, \dots, j_l)$, $j_i \in \{0, 1, \dots, m\}$, $i = 1, \dots, l$ we denote by J_α the multiple Stratonovich integral [17]

$$J_\alpha = \int_0^t \int_0^{s_l} \dots \int_0^{s_2} \circ dW_{j_1}(s_1) \dots \circ dW_{j_{l-1}}(s_{l-1}) \circ dW_{j_l}(s_l) \quad dW_0(s) = ds.$$

2.1. Colored trees and B-series

We study the strong order of convergence of the scheme (4)-(5) using $m+1$ -colored rooted trees and B-series [8, 11, 23]. 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

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

We also define [11] $\alpha : \mathcal{T} \rightarrow \mathbb{Q}$ given by $\alpha(\emptyset) = 1$, $\alpha(\bullet_l) = 1$,

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

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

For the SDE (1) and for every tree $t \in \mathcal{T}$ we define the corresponding elementary differential recursively using (10): $F(\emptyset)(x) = x$, $F(\bullet_l)(x) = g_l(x)$, and for any $x \in \mathbb{R}^d$

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

In [11] a stochastic B-series for (1) is defined as a formal series

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

From [11, Theorem 5, Theorem 6] we know that the exact solution $X(h, x_0)$, the numerical solution X_1 , and the stage values x_i , $i = 1, \dots, s$ can be written as stochastic B-series:

$$X(h, x_0) = B(\varphi, x_0; h), \quad X_1 = B(\Phi, x_0; h), \quad x_i = B(\Phi_i, x_0; h)$$

with

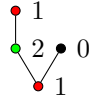
$$\varphi(\emptyset)(h) = 1, \varphi(\bullet_l) = W_l(h), \varphi(\tau)(h) = \int_0^h \left(\prod_{j=1}^k \varphi(\tau_j)(s) \right) \circ dW_l(s) \quad (12)$$

$$\Phi(\emptyset)(h) = 1, \Phi(\bullet_l) = \sum_{i=1}^s \gamma_i^{(l)}, \Phi(\tau)(h) = \sum_{i=1}^s \gamma_i^{(l)} \prod_{j=1}^k \Phi_i(\tau_j)(h) \quad (13)$$

$$\Phi_i(\emptyset)(h) = 1, \Phi_i(\bullet_l) = \sum_{j=1}^s Z_{ij}^{(l)}, \Phi_i(\tau)(h) = \sum_{j=1}^s Z_{ij}^{(l)} \prod_{j_1=1}^k \Phi_j(\tau_{j_1})(h), \quad (14)$$

where $\tau = [\tau_1, \dots, \tau_k]_l$ and $dW_0(t) = dt$.

For example, for the tree $\tau = [\bullet_0, [\bullet_1]_2]_1$, 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



$$|\tau| = 4, \rho(\tau) = 5/2, \alpha(\tau) = 1$$

$$F(\tau) = g_1''(g_0, g_2'g_1)$$

$$\varphi(\tau)(h) = \int_0^h \left(\int_0^s du \int_0^s W_1(u) \circ dW_2(u) \right) dW_1(s)$$

$$\Phi(\tau)(h) = \sum_{i,j,k,l=1}^s \gamma_i^{(1)} Z_{ij}^{(2)} Z_{jk}^{(1)} Z_{il}^{(0)}$$

If $Y(h) = B(\phi, x_0; h)$ and $f \in C^\infty(\mathbb{R}^d, \mathbb{R}^d)$ then from [11, Lemma 3] we have that

$$f(Y(h)) = \sum_{u \in U_f} \beta(u) \Psi_\phi(u)(h) G(u)(x_0) \quad (15)$$

where we define recursively

- U_f is a set of trees such that $[\emptyset]_f \in U_f$, and if $\tau_1, \dots, \tau_k \in \mathcal{T}$ then $[\tau_1, \dots, \tau_k]_f \in U_f$
- $\beta([\emptyset]_f) = 1$ and $\beta([\tau_1, \dots, \tau_k]_f) = \frac{1}{r_1! \dots r_l!} \prod_{j=1}^k \alpha(\tau_j)$, where $r_1, \dots, r_l$ count equal trees among $\tau_1, \dots, \tau_k$
- $\Psi_\phi([\emptyset]_f)(h) = 1$ and $\Psi_\phi([\tau_1, \dots, \tau_k]_f)(h) = \prod_{j=1}^k \phi(\tau_j)(h)$
- $G([\emptyset]_f)(x_0) = f(x_0)$ and $G([\tau_1, \dots, \tau_k]_f)(x_0) = f^{(k)}(x_0)(F(\tau_1)(x_0), \dots, F(\tau_k)(x_0))$.

We define $\rho([\emptyset]_f) = 0$ and $\rho([\tau_1, \dots, \tau_k]_f) = \sum_{j=1}^k \rho(\tau_j)$.

If $f = g_l$, from [11] (see Lemma 3 and its consequences on page 187) we know that

$$g_l(Y(h)) = \sum_{\tau \in T_l} \alpha(\tau) \phi_\phi^l(\tau)(h) F(\tau)(x_0), \quad (16)$$

where

$$\phi_\phi^l(\tau)(h) = \Psi_\phi([\tau_1, \dots, \tau_k]_{g_l})(h) = \prod_{j=1}^k \phi(\tau_j)(h) \text{ if } \tau = [\tau_1, \dots, \tau_k]_l \in T_l, \quad (17)$$

with $\phi^l(\bullet_l)(h) = \phi^l([\emptyset]_l)(h) = \Psi_\phi([\emptyset]_{g_l})(h) = 1$.

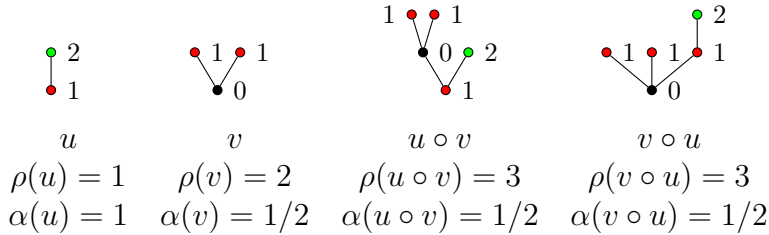
Remark 1. Notice that for any $f, g \in C^\infty(\mathbb{R}^d, \mathbb{R}^d)$ and any $u_1 = [\tau_1, \dots, \tau_k]_f \in U_f$, $u_2 = [\tau_1, \dots, \tau_k]_g \in U_g$, $k \in \mathbb{N}$, $\tau = [\tau_1, \dots, \tau_k]_l \in T_l$, $l = 0, \dots, m$ we have $\rho(u_1) = \rho(u_2) = \rho(\tau) - \rho([\emptyset]_l)$, $\beta(u_1) = \beta(u_2) = \alpha(\tau)$ and $\Psi_\phi(u_1) = \Psi_\phi(u_2) = \phi_\phi^l(\tau)$. This is also true if $u_1 = [\emptyset]_f$, $u_2 = [\emptyset]_g$, and $\tau = [\emptyset]_l$.

In the next sections we will need expansions similar with (15) and (16), but with a remainder. Let us consider the method (4)-(5) and assume that $\tilde{g}_0, g_l$ $l = 1, \dots, m$ satisfy (3), and $f^i, g_l^i \in \mathcal{C}_P^{2(p+1)}(\mathbb{R}^d, \mathbb{R})$, $l = 0, \dots, m$, $i = 1, \dots, d$. For simplicity we denote $\phi_i^l(\tau)(h) := \phi_{\Phi_i}^l(\tau)(h)$, $\tau \in T_l$. Suppose that $\phi_i^l(\tau)(h) = \mathcal{O}(h^{\rho(\tau) - \rho([\emptyset]_l)})$ for any $i = 1, \dots, s$ and any $\tau \in T_l$, $l = 0, \dots, m$, where here $\mathcal{O}(\cdot)$ refers to the $L^2(\Omega)$ norm and $h \rightarrow 0$. From [23, Theorem 4.1] and [22, Theorem 5.8, Proposition 6.1] we get the following expansions

$$f(x_i(h)) = \sum_{\substack{u \in U_f \\ \rho(u) \leq p}} \beta(u) \Psi_{\Phi_i}(u)(h) G(u)(x_0) + \mathcal{O}(h^{p+1/2}), \quad (18)$$

$$g_l(x_i(h)) = \sum_{\substack{\tau \in T_l \\ \rho(\tau) \leq p + \rho([\emptyset]_l)}} \alpha(\tau) \Phi_i^l(\tau)(h) F(\tau)(x_0) + \mathcal{O}(h^{p+1/2}). \quad (19)$$

The Butcher product [14] $u \circ v$ of two trees $u, v \in \mathcal{T} \setminus \{\emptyset\}$ is defined as the tree obtained by adding v as a branch to the root of u . Thus if $u = [u_1, \dots, u_k]_l$ then $u \circ v = [u_1, \dots, u_k, v]_l$. For example, for $u = [\bullet_2]_1$ and $v = [\bullet_1, \bullet_1]_0$ we have $u \circ v = [\bullet_2, [\bullet_1, \bullet_1]_0]_1$ and $v \circ u = [[\bullet_2]_1, \bullet_1, \bullet_1]_0$. The graphical representation is given below.



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 (1) as $h \rightarrow 0$ if there exist $C, \delta_0 > 0$ such that we have $E[|\bar{X}_{h,x_0}(t_k) - X(t_k, x_0)|] \leq Ch^n$, respectively $\sqrt{E[|\bar{X}_{h,x_0}(t_k) - X(t_k, x_0)|^2]} \leq Ch^n$, for any fixed $t_k = kh$, $k = 0, \dots, K$ and for all $0 < h \leq \delta_0$. Notice that by Jensens's inequality mean square convergence implies strong convergence of the same order. In [11] it

is proven that the method (4)-(5) constructed such that $\Phi(\tau)(h) = \mathcal{O}(h^{\rho(\tau)})$ has mean square global order N if

$$\Phi(\tau)(h) = \varphi(\tau)(h) + \mathcal{O}(h^{N+1/2}), \text{ for any } \tau \in T, \rho(\tau) \leq N \quad (20)$$

$$E[\Phi(\tau)(h)] = E[\varphi(\tau)(h)] + \mathcal{O}(h^{N+1}), \text{ for any } \tau \in T, \rho(\tau) \leq N + 1/2, \quad (21)$$

where the $\mathcal{O}(\cdot)$ notation refers to the $L^2(\Omega)$ norm and $h \rightarrow 0$.

3. Near preservation of quadratic invariants

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

$$\varphi(u)(h)\varphi(v)(h) = \varphi(u \circ v)(h) + \varphi(v \circ u)(h), \quad (22)$$

and if the SRK scheme (4)-(5) with time step h preserves the quadratic invariants by satisfying (6) then

$$\Phi(u)(h)\Phi(v)(h) = \Phi(u \circ v)(h) + \Phi(v \circ u)(h), \quad (23)$$

We are interested in SRK schemes that satisfy the following weaker version of (23) for $p \in \frac{1}{2}\mathbb{N}_0$

$$\Phi(u)(h)\Phi(v)(h) = \Phi(u \circ v)(h) + \Phi(v \circ u)(h) \text{ for } u, v \text{ with } \rho(u) + \rho(v) \leq p. \quad (24)$$

Theorem 1. *Suppose that $\tilde{g}_0, g_l$ $l = 1, \dots, m$ satisfy (3), $g_l^i \in \mathcal{C}_P^{2(p+1)}(\mathbb{R}^d, \mathbb{R})$, $l = 0, \dots, m$, $i = 1, \dots, d$, and $I(x) = x^\top Cx$, with $C \in \mathcal{M}_{d,d}(\mathbb{R})$ a symmetric constant matrix, is a quadratic invariant of (1). If the method (4)-(5) is constructed such that $\Phi(\tau)(h) = \mathcal{O}(h^{\rho(\tau)})$, $\Phi_i^l(\tau_1)(h) = \mathcal{O}(h^{\rho(\tau_1) - \rho([\emptyset]_i)})$ for any $i = 1, \dots, s$, $l = 1, \dots, m$, $\tau_1 \in T_l$ and any $\tau \in \mathcal{T}$, and for some $p \in \frac{1}{2}\mathbb{N}_0$ it satisfies (24) then we have*

$$X_{n+1}^\top C X_{n+1} = X_n^\top C X_n + \mathcal{O}(h^{p+1/2}). \quad (25)$$

Proof. Replacing X_{n+1} from (4) we get

$$X_{n+1}^\top C X_{n+1} = \left(X_n^\top + \sum_{l=0}^m \sum_{i=1}^s \gamma_i^{(l)} g_l^\top(x_i) \right) C \left(X_n + \sum_{l=0}^m \sum_{i=1}^s \gamma_i^{(l)} g_l(x_i) \right)$$

$$\begin{aligned}
&= X_n^\top C X_n + \sum_{l=0}^m \sum_{i=1}^s \gamma_i^{(l)} X_n^\top C g_l(x_i) + \sum_{l=0}^m \sum_{i=1}^s \gamma_i^{(l)} g_l^\top(x_i) C X_n \\
&+ \sum_{l_1, l_2=0}^m \sum_{i, j=1}^s \gamma_i^{(l_1)} \gamma_j^{(l_2)} g_{l_1}^\top(x_i) C g_{l_2}(x_j)
\end{aligned}$$

From (4) we obtain

$$\begin{aligned}
X_{n+1}^\top C X_{n+1} &= X_n^\top C X_n + \sum_{l_1=0}^m \sum_{i=1}^s \gamma_i^{(l_1)} \left(x_i^\top - \sum_{l_2=0}^m \sum_{j=1}^s Z_{ij}^{(l_2)} g_{l_2}^\top(x_j) \right) C g_{l_1}(x_i) \\
&+ \sum_{l_1=0}^m \sum_{i=1}^s \gamma_i^{(l_1)} g_{l_1}^\top(x_i) C \left(x_i - \sum_{l_2=0}^m \sum_{j=1}^s Z_{ij}^{(l_2)} g_{l_2}(x_j) \right) \\
&+ \sum_{l_1, l_2=0}^m \sum_{i, j=1}^s \gamma_i^{(l_1)} \gamma_j^{(l_2)} g_{l_1}^\top(x_i) C g_{l_2}(x_j)
\end{aligned}$$

Since I is an invariant for (1), we have $x_i^\top C g_l(x_i) = x_i C g_l^\top(x_i) = 0$, for any $i = 1, \dots, s$ and $l = 1, \dots, m$, so simple calculations yield

$$\begin{aligned}
X_{n+1}^\top C X_{n+1} &= X_n^\top C X_n - \sum_{l_1, l_2=0}^m \sum_{i, j=1}^s \gamma_i^{(l_1)} Z_{ij}^{(l_2)} g_{l_2}^\top(x_j) C g_{l_1}(x_i) \\
&- \sum_{l_1, l_2=0}^m \sum_{i, j=1}^s \gamma_i^{(l_1)} Z_{ij}^{(l_2)} g_{l_1}^\top(x_i) C g_{l_2}(x_j) + \sum_{l_1, l_2=0}^m \sum_{i, j=1}^s \gamma_i^{(l_1)} \gamma_j^{(l_2)} g_{l_1}^\top(x_i) C g_{l_2}(x_j) \\
&= X_n^\top C X_n + \sum_{l_1, l_2=0}^m \sum_{i, j=1}^s \left(\gamma_i^{(l_1)} \gamma_j^{(l_2)} - \gamma_j^{(l_2)} Z_{ji}^{(l_1)} - \gamma_i^{(l_1)} Z_{ij}^{(l_2)} \right) g_{l_1}^\top(x_i) C g_{l_2}(x_j)
\end{aligned} \tag{26}$$

Using (19) we obtain for any $i, j = 1, \dots, s$, $l_1, l_2 = 0, \dots, m$

$$g_{l_1}^\top(x_i) C g_{l_2}(x_j) = \sum_{\substack{\tau_1 \in T_{l_1} \\ \rho(\tau_1) \leq p + \rho([\emptyset]_{l_1})}} \sum_{\substack{\tau_2 \in T_{l_2} \\ \rho(\tau_2) \leq p + \rho([\emptyset]_{l_2})}} \alpha(\tau_1) \alpha(\tau_2) \Phi_i^{l_1}(\tau_1)(h) \Phi_j^{l_2}(\tau_2)(h)$$

$$F(\tau_1)^\top(x_0) C F(\tau_2)(x_0) + \mathcal{O}(h^{p+1/2})$$

Thus from (13) we have

$$\sum_{i, j=1}^s \gamma_i^{(l_1)} \gamma_j^{(l_2)} g_{l_1}^\top(x_i) C g_{l_2}(x_j) = \sum_{\substack{\tau_1 = [u_1, \dots, u_{k_1}]_{l_1} \\ \sum_{i=1}^{k_1} \rho(u_i) \leq p}} \alpha(\tau_1) \left(\sum_{i=1}^s \gamma_i^{(l_1)} \prod_{i_1=1}^{k_1} \Phi_i(u_{i_1})(h) \right)$$

$$\begin{aligned}
& \sum_{\substack{\tau_2=[v_1, \dots, v_{k_2}]_{l_2} \\ \sum_{l=1}^{k_2} \rho(v_l) \leq p}} \alpha(\tau_2) \left(\sum_{j=1}^s \gamma_j^{(l_2)} \prod_{i_2=1}^{k_2} \Phi_j(v_{i_2})(h) \right) F(\tau_1)^\top(x_0) C F(\tau_2)(x_0) + \mathcal{O}(h^{p+1/2}) \\
&= \sum_{\substack{\tau_1=[u_1, \dots, u_{k_1}]_{l_1} \\ \sum_{l=1}^{k_1} \rho(u_l) \leq p \\ \tau_2=[v_1, \dots, v_{k_2}]_{l_2} \\ \sum_{l=1}^{k_2} \rho(v_l) \leq p}} \alpha(\tau_1) \alpha(\tau_2) \Phi(\tau_1)(h) \Phi(\tau_2)(h) F(\tau_1)^\top(x_0) C F(\tau_2)(x_0) + \mathcal{O}(h^{p+1/2})
\end{aligned}$$

Moreover, we can use the Butcher product and (13), (14) to get

$$\begin{aligned}
& \sum_{i,j=1}^s \gamma_j^{(l_2)} Z_{ji}^{(l_1)} g_{l_1}^\top(x_i) C g_{l_2}(x_j) = \sum_{\substack{\tau_1=[u_1, \dots, u_{k_1}]_{l_1} \\ \sum_{l=1}^{k_1} \rho(u_l) \leq p \\ \tau_2=[v_1, \dots, v_{k_2}]_{l_2} \\ \sum_{l=1}^{k_2} \rho(v_l) \leq p}} \alpha(\tau_1) \alpha(\tau_2) \sum_{j=1}^s \gamma_j^{(l_2)} \prod_{i_2=1}^{k_2} \Phi_j(v_{i_2})(h) \\
& \sum_{i=1}^s Z_{ji}^{(l_1)} \prod_{i_1=1}^{k_1} \Phi_i(u_{i_1})(h) F(\tau_1)^\top(x_0) C F(\tau_2)(x_0) + \mathcal{O}(h^{p+1/2}) = \sum_{\substack{\tau_1=[u_1, \dots, u_{k_1}]_{l_1} \\ \sum_{l=1}^{k_1} \rho(u_l) \leq p \\ \tau_2=[v_1, \dots, v_{k_2}]_{l_2} \\ \sum_{l=1}^{k_2} \rho(v_l) \leq p}} \alpha(\tau_1) \\
& \alpha(\tau_2) \sum_{j=1}^s \gamma_j^{(l_2)} \prod_{i_2=1}^{k_2} \Phi_j(v_{i_2})(h) \Phi_j(\tau_1) F(\tau_1)^\top(x_0) C F(\tau_2)(x_0) + \mathcal{O}(h^{p+1/2}) \\
&= \sum_{\substack{\tau_1=[u_1, \dots, u_{k_1}]_{l_1} \\ \sum_{l=1}^{k_1} \rho(u_l) \leq p \\ \tau_2=[v_1, \dots, v_{k_2}]_{l_2} \\ \sum_{l=1}^{k_2} \rho(v_l) \leq p}} \alpha(\tau_1) \alpha(\tau_2) \Phi(\tau_2 \circ \tau_1)(h) F(\tau_1)^\top(x_0) C F(\tau_2)(x_0) + \mathcal{O}(h^{p+1/2})
\end{aligned}$$

Similarly we have

$$\begin{aligned}
& \sum_{i,j=1}^s \gamma_i^{(l_1)} Z_{ij}^{(l_2)} g_{l_1}^\top(x_i) C g_{l_2}(x_j) = \sum_{\substack{\tau_1=[u_1, \dots, u_{k_1}]_{l_1} \\ \sum_{l=1}^{k_1} \rho(u_l) \leq p \\ \tau_2=[v_1, \dots, v_{k_2}]_{l_2} \\ \sum_{l=1}^{k_2} \rho(v_l) \leq p}} \alpha(\tau_1) \alpha(\tau_2) \Phi(\tau_1 \circ \tau_2)(h) \\
& F(\tau_1)^\top(x_0) C F(\tau_2)(x_0) + \mathcal{O}(h^{p+1/2})
\end{aligned}$$

Replacing in (26) we obtain

$$\begin{aligned}
X_{n+1}^\top CX_{n+1} &= X_n^\top CX_n + \sum_{l_1, l_2=0}^m \sum_{\substack{\tau_1=[u_1, \dots, u_{k_1}]_{l_1} \\ \sum_{l=1}^{k_1} \rho(u_l) \leq p \\ \tau_2=[v_1, \dots, v_{k_2}]_{l_2} \\ \sum_{l=1}^{k_2} \rho(v_l) \leq p}} \alpha(\tau_1)\alpha(\tau_2) \left(\Phi(\tau_1)(h)\Phi(\tau_2)(h) \right. \\
&\quad \left. - \Phi(\tau_2 \circ \tau_1)(h) - \Phi(\tau_1 \circ \tau_2)(h) \right) F(\tau_1)^\top(x_0)CF(\tau_2)(x_0) + \mathcal{O}(h^{p+1/2}) \\
&= X_n^\top CX_n + \sum_{l_1, l_2=0}^m \sum_{\substack{\tau_1 \in T_{l_1} \\ \tau_2 \in T_{l_2} \\ \rho(\tau_1) + \rho(\tau_2) \leq p}} \alpha(\tau_1)\alpha(\tau_2) \left(\Phi(\tau_1)(h)\Phi(\tau_2)(h) - \Phi(\tau_2 \circ \tau_1)(h) \right. \\
&\quad \left. - \Phi(\tau_1 \circ \tau_2)(h) \right) F(\tau_1)^\top(x_0)CF(\tau_2)(x_0) + \sum_{l_1, l_2=0}^m \sum_{\substack{\tau_1=[u_1, \dots, u_{k_1}]_{l_1} \\ \sum_{l=1}^{k_1} \rho(u_l) \leq p \\ \tau_2=[v_1, \dots, v_{k_2}]_{l_2} \\ \sum_{l=1}^{k_2} \rho(v_l) \leq p \\ \rho(\tau_1) + \rho(\tau_2) > p}} \alpha(\tau_1)\alpha(\tau_2) \\
&\quad \left(\Phi(\tau_1)(h)\Phi(\tau_2)(h) - \Phi(\tau_2 \circ \tau_1)(h) - \Phi(\tau_1 \circ \tau_2)(h) \right) F(\tau_1)^\top(x_0)CF(\tau_2)(x_0) + \mathcal{O}(h^{p+1/2}) \\
&= X_n^\top CX_n + \sum_{l_1, l_2=0}^m \sum_{\substack{\tau_1=[u_1, \dots, u_{k_1}]_{l_1} \\ \sum_{l=1}^{k_1} \rho(u_l) \leq p \\ \tau_2=[v_1, \dots, v_{k_2}]_{l_2} \\ \sum_{l=1}^{k_2} \rho(v_l) \leq p \\ \rho(\tau_1) + \rho(\tau_2) > p}} \mathcal{O}(h^{\rho(\tau_1) + \rho(\tau_2)}) + \mathcal{O}(h^{p+1/2}) \\
&= X_n^\top CX_n + \mathcal{O}(h^{p+1/2})
\end{aligned}$$

Here we have used the fact that the method (4)-(5) is constructed such that $\Phi(\tau)(h) = \mathcal{O}(h^{\rho(\tau)})$ and it satisfies (24). $\square$

The next result is an extension of [1, Theorem 5], and it reduces substantially the number of trees for which the condition (20) should be checked to establish the order of convergence for the SRK scheme (4)-(5). 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.

Theorem 2. Assume that the method (4)-(5) is constructed such that $\Phi(\tau)(h) = \mathcal{O}(h^{\rho(\tau)})$ for any $\tau \in \mathcal{T}$, and for some $p \in \frac{1}{2}\mathbb{N}_0$ it satisfies (24). Let $r \leq p$ and $\hat{\tau}$ be an unrooted tree of order $q \leq r$. If $\varphi(\tau_0)(h) = \Phi(\tau_0)(h) + \mathcal{O}(h^{r+1/2})$ for some rooted tree $\tau_0 \in \hat{\mathcal{T}}(\hat{\tau})$ and all trees of order less than q , then $\varphi(\tau)(h) = \Phi(\tau)(h) + \mathcal{O}(h^{r+1/2})$ for all $\tau \in \hat{\mathcal{T}}(\hat{\tau})$.

Proof. The proof is similar to the proof of [1, Theorem 5]. The theorem is obviously true for trees with one node. Assume that $\hat{\tau}$ is an unrooted tree with 2 or more nodes and order $q \leq r \leq p$, and there exists $\tau_0 \in \hat{\mathcal{T}}(\hat{\tau})$ such that $\varphi(\tau_0)(h) = \Phi(\tau_0)(h) + \mathcal{O}(h^{r+1/2})$. We pick a branch v from the root of τ_0 and let u be such that $\tau_0 = u \circ v$. Then $\rho(u) + \rho(v) = \rho(\tau_0) = q \leq r \leq p$, so from (24) we get $\Phi(u)(h)\Phi(v)(h) = \Phi(u \circ v)(h) + \Phi(v \circ u)(h) = \Phi(\tau_0)(h) + \Phi(v \circ u)(h)$. Since $\rho(u), \rho(v)$ are less than q we have $\varphi(u)(h) = \Phi(u)(h) + \mathcal{O}(h^{r+1/2})$, and $\varphi(v)(h) = \Phi(v)(h) + \mathcal{O}(h^{r+1/2})$. From these equations and (22) we get $\varphi(v \circ u)(h) = \varphi(u)(h)\varphi(v)(h) - \varphi(u \circ v)(h) = \Phi(u)(h)\Phi(v)(h) - \Phi(\tau_0)(h) + \mathcal{O}(h^{r+1/2}) = \Phi(v \circ u)(h) + \mathcal{O}(h^{r+1/2})$. Since the branch v was arbitrarily chosen, we get the condition for all the trees $\tau \in \hat{\mathcal{T}}(\hat{\tau})$ with the root shifted to one of the nodes adjacent to the root of τ_0 . Repeating this argument we can obtain the condition for all $\tau \in \hat{\mathcal{T}}(\hat{\tau})$. $\square$

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, for one stochastic process ($m = 1$) as mentioned in [1], there are 27 rooted trees of order 2.5 and only 9 unrooted trees. Moreover, the difference between the number of unrooted and rooted trees increases substantially with m . Table 3 compares the number of rooted and unrooted trees of order 1.5 and less for one and two stochastic processes. Thus, to prove that (4)-(5) applied to the SDE (1) with $m = 2$ has mean square order 1.5, we have to check condition (20) for 25 rooted trees with order at most 1.5. However, if the SRK method satisfies (24) with $p = 1.5$, then we only have to check condition (20) for 14 rooted trees that correspond to the unrooted trees with order at most 1.5.

$\rho(\tau)$	$m = 1$ Rooted	$m = 1$ Unrooted	$m = 2$ Rooted	$m = 2$ Unrooted
0.5	1	1	2	2
1	2	2	5	4
1.5	4	2	18	8

Table 1: Number of rooted and unrooted trees for $m = 1$ and $m = 2$

4. Pseudo-Symplectic Stochastic Runge Kutta schemes

We consider the SHS (7) and we notice that we can re-write it as

$$dX(t) = J^{-1}\nabla H^{(0)} + \sum_{l=1}^m J^{-1}\nabla H^{(l)} \circ dW_l(t), \quad X(0) = x_0 = (p_0, q_0) \in \mathbb{R}^{2d}, \quad (27)$$

where $X(t) = (P(t), Q(t)) \in \mathbb{R}^{2d}$, $l = 0, \dots, m$, and J is defined in (8). Notice that (27) is a $2d$ -dimensional autonomous SDE of the form (1) with $g_l = J^{-1}\nabla H^{(l)}$. We also know that $\Sigma(t) = \frac{\partial X(t, x_0)}{\partial x_0}$ is the solution of the stochastic variational equation [16]

$$d\Sigma(t) = \sum_{l=0}^m e_l(X(t))\Sigma(t) \circ dW_l(t), \quad \Sigma(0) = I_{2d}, \quad (28)$$

where $e_l(X(t)) = J^{-1}F_l(X(t))$ and $F_l(\cdot)$ is the Hessian matrix of $H^{(l)}(\cdot)$, $l = 0, \dots, m$.

In [26, Theorem 2.2] it has been proven that if the method (4)-(5) satisfies (6) then it preserves the symplectic structure. In terms of the associated colored trees we know from [1, Lemma 4] that we have (22), and if the SRK preserves the symplectic structure by satisfying (6) then it also satisfies (23). For the general SHS (7) symplectic schemes are implicit, and explicit symplectic schemes can be constructed only if we consider special types of Hamiltonian functions [20, 26]. Here we construct explicit SRK schemes (i.e. schemes of the form (4)-(5) with $Z_{ij}^{(l)} = 0, j = i + 1, \dots, s$) that preserve the symplectic structure with a certain accuracy. We proceed as in the deterministic case [4] and instead of (23) we require

$$\begin{aligned} \Phi(u)(h)\Phi(v)(h) &= \Phi(u \circ v)(h) + \Phi(v \circ u)(h) \text{ for all trees } u, v \in \mathcal{T}, \\ \rho(u) + \rho(v) &\leq M + 1/2. \end{aligned} \quad (29)$$

Theorem 3. *Let us assume that $H^{(r)} \in \mathcal{C}_P^{2M+2}(\mathbb{R}^{2d}, \mathbb{R})$ and $H^{(r)}$ has bounded second order partial derivatives, $r = 0, \dots, m$, $i = 1, \dots, s$. We consider the SHS (7) written as the $2d$ - dimensional autonomous SDE (27) and the explicit SRK scheme given by (4)-(5) with $Z_{ij}^{(l)} = 0, j = i+1, \dots, s, i = 1, \dots, s$. Suppose that the scheme is constructed such that $\Phi(\tau)(h) = \mathcal{O}(h^{\rho(\tau)})$, $\Phi_i^l(\tau_1)(h) = \mathcal{O}(h^{\rho(\tau_1) - \rho([\emptyset]_i)})$ for any $i = 1, \dots, s$, $l = 1, \dots, m$, $\tau_1 \in T_l$, and any $\tau \in \mathcal{T}$. If conditions (20), (21), and (29) are satisfied then it is a pseudo-symplectic explicit SRK method of mean-square order (N, M) .*

The proof is included in appendix Appendix A.

5. Examples of Pseudo-symplectic stochastic Runge Kutta methods

We consider the one-step approximation corresponding to an explicit version of the s -stage stochastic Runge Kutta scheme (4)-(5) with $\gamma_i^{(0)} = \alpha_i h$, $\gamma_i^{(l)} = \beta_i^{(l)} J_l$, $Z_{ij}^{(0)} = A_{ij} h$, $Z_{ij}^{(l)} = B_{ij}^{(l)} J_l$, $i, j = 1, \dots, s$, $l = 1, \dots, m$:

$$P = p - \sum_{i=1}^s \alpha_i^{(0)} h \nabla_Q H^{(0)}(p_i, q_i) - \sum_{l=1}^m \sum_{i=1}^s \beta_i^{(l)} J_l \nabla_Q H^{(l)}(p_i, q_i), \quad (30)$$

$$Q = q + \sum_{i=1}^s \alpha_i^{(0)} h \nabla_P H^{(0)}(p_i, q_i) + \sum_{l=1}^m \sum_{i=1}^s \beta_i^{(l)} J_l \nabla_P H^{(l)}(p_i, q_i), \quad (31)$$

$$p_i = p - \sum_{j=1}^{i-1} A_{ij} h \nabla_Q H^{(0)}(p_j, q_j) - \sum_{l=1}^m \sum_{j=1}^{i-1} B_{ij}^{(l)} J_l \nabla_Q H^{(l)}(p_j, q_j), \quad p_1 = p, \quad (32)$$

$$q_i = q + \sum_{j=1}^{i-1} A_{ij} h \nabla_P H^{(0)}(p_j, q_j) + \sum_{l=1}^m \sum_{j=1}^{i-1} B_{ij}^{(l)} J_l \nabla_P H^{(l)}(p_j, q_j), \quad q_1 = q, \quad (33)$$

Here $J_k = W_k(h)$, $k = 1, \dots, m$ and we simulate them as $\sqrt{h} \xi_k$, where $\xi_k \sim N(0, 1)$ and $\{\xi_k, k = 1, \dots, m\}$ independent. The Butcher tableau associated with this SRK method is

$$\left| \begin{array}{c|c|c|c|c} A & B^{(1)} & B^{(2)} & \dots & B^{(m)} \\ \hline \alpha^\top & (\beta^{(1)})^\top & (\beta^{(2)})^\top & \dots & (\beta^{(m)})^\top \end{array} \right|$$

The condition (6) which ensure that the scheme preserve symplectic structure becomes [26, Theorem 2.2]:

$$\alpha_i \alpha_j - \alpha_j A_{ji} - \alpha_i A_{ij} = 0 \quad (34)$$

$$\alpha_i \beta_j^{(k)} - \beta_j^{(k)} A_{ji} - \alpha_i B_{ij}^{(k)} = 0 \quad (35)$$

$$\beta_i^{(k)} \beta_j^{(l)} - \beta_j^{(l)} B_{ji}^{(k)} - \beta_i^{(k)} B_{ij}^{(l)} = 0, \quad (36)$$

for all $i, j = 1, \dots, s$, $k, l = 1, \dots, m$. These conditions are hard to fulfill and here we replace them with the less restrictive conditions given in (29) to get pseudo-symplectic SRK schemes. Although the application of Theorem 2 reduces the number of order conditions (20), the number of conditions required by (29) increases fast with the scheme order.

We mention that for general SHSs with more than one Wiener process, the maximum strong order for the explicit SRK of the form (4)-(5) with random variables representing all order 1 and order 2 Stratonovich integrals is 0.5, irrespective of the number of stages s [6]. Thus, unless we consider a special type of SHS, from (30)-(33) we can obtain only pseudo-symplectic schemes of order $(0.5, M)$, $M > 0.5$. Let $e = (1, 1, \dots, 1)^\top \in \mathbb{R}^s$.

Corollary 1. *Let us assume that for any $r = 0, \dots, m$, $H^{(r)} \in \mathcal{C}_P^4(\mathbb{R}^{2d}, \mathbb{R})$ and $H^{(r)}$ have bounded second order partial derivatives. The schemes given by (30)-(33) are pseudo-symplectic explicit SRK methods of mean-square order $(0.5, 1)$ if the following conditions are satisfied for any $k, l, l_1 = 1, \dots, m$*

$$\beta^{(k)\top} e = 1, \quad \alpha^\top e = 1, \quad \beta^{(k)\top} B^{(k)} e = 1/2, \quad (37)$$

$$\beta^{(l)\top} B^{(k)} e + \beta^{(k)\top} B^{(l)} e = 1, \quad (38)$$

$$\beta^{(l)\top} A e + \alpha^\top B^{(l)} e = 1, \quad (39)$$

$$\beta^{(k)\top} B^{(l)} (B^{(l_1)} e) + \beta^{(l)\top} ((B^{(k)} e)(B^{(l_1)} e)) = (\beta^{(l)\top} B^{(l_1)} e). \quad (40)$$

Proof. This is a straightforward application of Theorem 3. From (20), (21) the conditions for mean square order 0.5 are (37). As illustrated in Table 2, for $M = 1$, from (29) we get for any $l, k, l_1 = 1, \dots, m$

$$\begin{aligned} (\beta^{(k)\top} e)(\beta^{(l)\top} e) &= \beta^{(l)\top} B^{(k)} e + \beta^{(k)\top} B^{(l)} e \\ (\beta^{(l)\top} e)(\alpha^\top e) &= \beta^{(l)\top} A e + \alpha^\top B^{(l)} e \\ (\beta^{(k)\top} e)(\beta^{(l)\top} B^{(l_1)} e) &= \beta^{(k)\top} B^{(l)} (B^{(l_1)} e) + \beta^{(l)\top} ((B^{(k)} e)(B^{(l_1)} e)) \end{aligned}$$

□

 $\Phi(u) =$ $J_k \sum_{j=1}^s \beta_j^{(k)}$	 $\Phi(v) =$ $J_l \sum_{i=1}^s \beta_i^{(l)}$	 $\Phi(u \circ v) = J_k J_l$ $\sum_{i=1}^s \sum_{j=1}^{i-1} \beta_i^{(k)} B_{ij}^{(l)}$	 $\Phi(v \circ u) = J_l J_k$ $\sum_{i=1}^s \sum_{j=1}^{i-1} \beta_i^{(l)} B_{ij}^{(k)}$
 $\Phi(u) =$ $h \sum_{j=1}^s \alpha_j$	 $\Phi(v) =$ $J_l \sum_{i=1}^s \beta_i^{(l)}$	 $\Phi(u \circ v) = h J_l$ $\sum_{i=1}^s \sum_{j=1}^{i-1} \alpha_i B_{ij}^{(l)}$	 $\Phi(v \circ u) = h J_l$ $\sum_{i=1}^s \sum_{j=1}^{i-1} \beta_i^{(l)} A_{ij}$
 $\Phi(u) =$ $J_k \sum_{j=1}^s \beta_j^{(k)}$	 $\Phi(v) = J_l J_{l_1}$ $\sum_{i=1}^s \sum_{j=1}^{i-1} \beta_i^{(l)} B_{ij}^{(l_1)}$	 $\Phi(u \circ v) = J_k J_l J_{l_1}$ $\sum_{i=1}^s \sum_{j=1}^{i-1} \sum_{i_1=1}^{j-1} \beta_i^{(k)} B_{ij}^{(l)} B_{ji_1}^{(l_1)}$	 $\Phi(v \circ u) = J_l J_k J_{l_1}$ $\sum_{i=1}^s \sum_{j=1}^{i-1} \sum_{i_1=1}^{i-1} \beta_i^{(l)} B_{ij}^{(k)} B_{ii_1}^{(l_1)}$

Table 2: The condition (29) for trees u, v with $\rho(u) + \rho(v) \leq 1.5$

Remark 2. Notice that for a SRK with $s = 2$ stages, condition (40) becomes

$$\sum_{j=1}^s \sum_{i_1=1}^{j-1} \beta_j^{(l)} B_{ji_1}^{l_1} = \sum_{i=1}^s \sum_{j=1}^{i-1} \sum_{i_1=1}^{j-1} \beta_i^{(l)} B_{ij}^{(k)} B_{ii_1}^{(l_1)}. \quad (41)$$

Similarly with the deterministic RK pseudo-symplectic methods, we can easily obtain an explicit pseudo-symplectic SRK by repeating the values for an implicit symplectic method, but the number of stages might get large. For example, starting from the 1-stage symplectic SRK with the tableau 3 (the symplectic midpoint scheme (3.7) in [19]) we can get the a 3-stage explicit

$$\begin{array}{c|c|c|c} 1/2 & 1/2 & \dots & 1/2 \\ \hline 1 & 1 & \dots & 1 \end{array}$$

Table 3: Symplectic scheme of order 0.5

SRK with the tableau 4. Notice that this method satisfies conditions (37)-(40), so its is a pseudo-symplectic method of order (0.5, 1).

0	0	0	0	0	0	...	0	0	0
1/2	0	0	1/2	0	0	...	1/2	0	0
0	1/2	0	0	1/2	0	...	0	1/2	0
0	0	1	0	0	1	...	0	0	1

Table 4: Three-stages pseudo-symplectic SRK scheme of order (0.5,1)

Next we obtain pseudo-symplectic method of order (0.5, 1) by solving the equations (37) -(40) using MAPLE. The explicit SRK with the tableau 5 are two-stages pseudo-symplectic methods of order (0.5, 1), $\alpha_2 \in \mathbb{R}$

0	0	0	0	...	0	0
$-2\alpha_2 + 2$	0	1	0	...	1	0
$1 - \alpha_2$	α_2	$1/2$	$1/2$	...	$1/2$	$1/2$

Table 5: Two-stages pseudo-symplectic SRK scheme of order (0.5,1)

The tableau 6 corresponds to a family of three-stages pseudo-symplectic methods of order (0.5, 1), where $A_{31} = 4b_1\alpha_3\beta_2 - 2\beta_2a_1 - 2\alpha_2\beta_1 - a_3 - 2\alpha_3 + 2$, $B_{31}^{(1)} = \dots = B_{31}^{(m)} = 4b_1\beta_2^2 - 4\beta_2 + 1$, $B_{32}^{(1)} = \dots = B_{32}^{(m)} = -4b_1\beta_2^2 - 2\beta_2b_1 + 4\beta_2$, $b_1, a_1, \alpha_2, \alpha_3, \beta_3 \in \mathbb{R}$.

0	0	0	0	0	0	...	0	0	0
a_1	0	0	b_1	0	0	...	b_1	0	0
A_{31}	a_3	0	$B_{31}^{(1)}$	$B_{32}^{(1)}$	0	...	$B_{31}^{(m)}$	$B_{32}^{(m)}$	0
$1 - \alpha_2 - \alpha_3$	α_2	α_3	$1/2 - \beta_2$	β_2	1/2	...	$1/2 - \beta_2$	β_2	1/2

Table 6: Family of three-stages pseudo-symplectic SRK schemes of order (0.5,1)

5.1. Pseudo-symplectic SRK for SHS with commutative noise

The SDE (1) has commutative noise if for any $i, j = 1, \dots, m$, $i \neq j$, and any x we have

$$[g_i, g_j] := g'_i(x) \cdot g_j(x) - g'_j(x) \cdot g_i(x) = 0.$$

From [6, Theorem 4.2] we know that for the SHS (7) with commutative noise the scheme (30)-(33) attains strong order 1 if for any $k, l = 1, \dots, m$

$$\beta^{(k)\top} e = 1, \quad \alpha^\top e = 1, \quad \beta^{(k)\top} B^{(l)} e + \beta^{(l)\top} B^{(k)} e = 1. \quad (42)$$

From Theorem 3 applied for $M = 1.5$ we obtain the following conditions to construct a pseudo-symplectic SRK scheme of order $(1, 1.5)$.

Corollary 2. *Let us assume that the SHS (7) has commuting noise and for any $r = 0, \dots, m$, $H^{(r)} \in \mathcal{C}_P^5(\mathbb{R}^{2d}, \mathbb{R})$ and $H^{(r)}$ have bounded second order partial derivatives. The schemes given by (30)-(33) are pseudo-symplectic explicit SRK methods of mean-square order $(1, 1.5)$ if the following conditions are satisfied for any $k, l, l_1, l_2, l_3, l_4 = 1, \dots, m$*

$$\beta^{(k)\top} e = 1, \quad \alpha^\top e = 1, \quad \beta^{(k)\top} B^{(l)} e + \beta^{(l)\top} B^{(k)} e = 1, \quad (43)$$

$$\beta^{(l)\top} A e + \alpha^\top B^{(l)} e = 1, \quad (44)$$

$$\beta^{(k)\top} B^{(l)} (B^{(l_1)} e) + \beta^{(l)\top} ((B^{(k)} e) (B^{(l_1)} e)) = (\beta^{(l)\top} B^{(l_1)} e), \quad (45)$$

$$\begin{aligned} & \beta^{(l_1)\top} (B^{(l_2)} ((B^{(l_3)} e) (B^{(l_4)} e))) + \beta^{(l_2)\top} ((B^{(l_1)} e) (B^{(l_3)} e) (B^{(l_4)} e)) \\ &= \left(\beta^{(l_2)\top} ((B^{(l_3)} e) (B^{(l_4)} e)) \right) \end{aligned} \quad (46)$$

$$\begin{aligned} & \beta^{(l_1)\top} (B^{(l_2)} (B^{(l_3)} (B^{(l_4)} e))) + \beta^{(l_2)\top} ((B^{(l_1)} e) (B^{(l_3)} (B^{(l_4)} e))) \\ &= \left(\beta^{(l_2)\top} (B^{(l_3)} (B^{(l_4)} e)) \right) \end{aligned} \quad (47)$$

$$\beta^{(l_1)\top} (B^{(l_2)} (A e)) + \beta^{(l_2)\top} ((B^{(l_1)} e) (A e)) = \left(\beta^{(l_2)\top} (A e) \right) \quad (48)$$

$$\beta^{(l_1)\top} (A (B^{(l_2)} e)) + \alpha^\top ((B^{(l_1)} e) (B^{(l_2)} e)) = (\alpha^\top B^{(l_2)} e) \quad (49)$$

$$\alpha^\top A e = \frac{1}{2} \quad (50)$$

$$\alpha^\top (B^{(l_1)} (B^{(l_2)} e)) + \beta^{(l_1)\top} ((B^{(l_2)} e) (A e)) = \left(\beta^{(l_1)\top} B^{(l_2)} e \right) \quad (51)$$

$$\begin{aligned} & \beta^{(l_1)\top} ((B^{(l_2)} e) (B^{(l_3)} (B^{(l_4)} e))) + \beta^{(l_3)\top} ((B^{(l_4)} e) (B^{(l_1)} (B^{(l_2)} e))) \\ &= \left(\beta^{(l_1)\top} B^{(l_2)} e \right) \left(\beta^{(l_3)\top} B^{(l_4)} e \right) \end{aligned} \quad (52)$$

A family of explicit three-stages pseudo-symplectic SRK schemes of order $(1, 1.5)$ can be constructed from the tableau 7 with $a_1 \in \mathbb{R}$, $b_1 \in \mathbb{R} \setminus \{0, 3/4, 1\}$ and

$$\begin{aligned} A_{31} &= \frac{a_1(2b_1 - 1)^2}{b_1^2(4b_1 - 3)}, \quad A_{32} = \frac{(b_1 - 1)(2b_1^2 + a_1 - 2a_1b_1)}{b_1^2(4b_1 - 3)} \\ B_{31}^{(1)} &= \dots = B_{31}^{(m)} = \frac{(2b_1 - 1)^2}{b_1(4b_1 - 3)}, \quad B_{32}^{(1)} = \dots = B_{32}^{(m)} = \frac{b_1 - 1}{b_1(4b_1 - 3)} \\ \alpha_1 &= \frac{a_1(4b_1 - 1)}{8b_1^2}, \quad \alpha_2 = \frac{2a_1b_1 - 2b_1^2 - a_1}{8b_1^2(b_1 - 1)}, \quad \alpha_3 = \frac{(3 - 4b_1)(a_1 - 2b_1)}{8b_1(b_1 - 1)} \end{aligned}$$

$$\beta_1^{(1)} = \dots = \beta_1^{(m)} = \frac{4b_1 - 1}{8b_1}, \beta_2^{(1)} = \dots = \beta_2^{(m)} = \frac{-1}{8b_1(b_1 - 1)}$$

$$\beta_3^{(1)} = \dots = \beta_3^{(m)} = \frac{4b_1 - 3}{8(b_1 - 1)}$$

0	0	0	0	0	0	...	0	0	0
a_1	0	0	b_1	0	0	...	b_1	0	0
A_{31}	A_{32}	0	$B_{31}^{(1)}$	$B_{32}^{(1)}$	0	...	$B_{31}^{(m)}$	$B_{32}^{(m)}$	0
α_1	α_2	α_3	$\beta_1^{(1)}$	$\beta_2^{(1)}$	$\beta_3^{(1)}$	...	$\beta_1^{(m)}$	$\beta_2^{(m)}$	$\beta_3^{(m)}$

Table 7: Family of 3-stages pseudo-symplectic methods of order (1,1.5) for SHS with commutative noise

6. Numerical simulations

For the deterministic case, [4] includes a theoretical proof and numerical illustrations that pseudo-symplectic Runge-Kutta schemes have a good long-term behavior. In the stochastic case, for SHSs with additive noise numerical simulations show that pseudo-symplectic SRK schemes give accurate long-term approximations [21]. Here we illustrate numerically the performance of the proposed pseudo-symplectic schemes for SHS with multiplicative noise.

6.1. Kubo oscillator

The Kubo oscillator is used to demonstrate the advantage of the stochastic symplectic schemes for long time computation [19], [12]:

$$\begin{aligned} dP &= -aQdt - \sigma Q \circ dW(t), & P(0) &= p, \\ dQ &= aPdt + \sigma P \circ dW(t), & Q(0) &= q, \end{aligned} \tag{53}$$

where a and σ are constants. We can easily find the exact solution of the linear system (53) [3].

As illustrated in [19], 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. This means that the phase trajectory of (53) lies on the circle with the center at the origin and the radius $\sqrt{p^2 + q^2}$. The Hamiltonian functions are also quadratic invariants for the SHS (53).

We consider the SRK scheme corresponding to the tableau 5 with $\alpha_2 = 1/2$. Since the Kubo system has one noise and the conditions (42) are satisfied, the scheme has mean square order 1. From Theorem 3 we know that

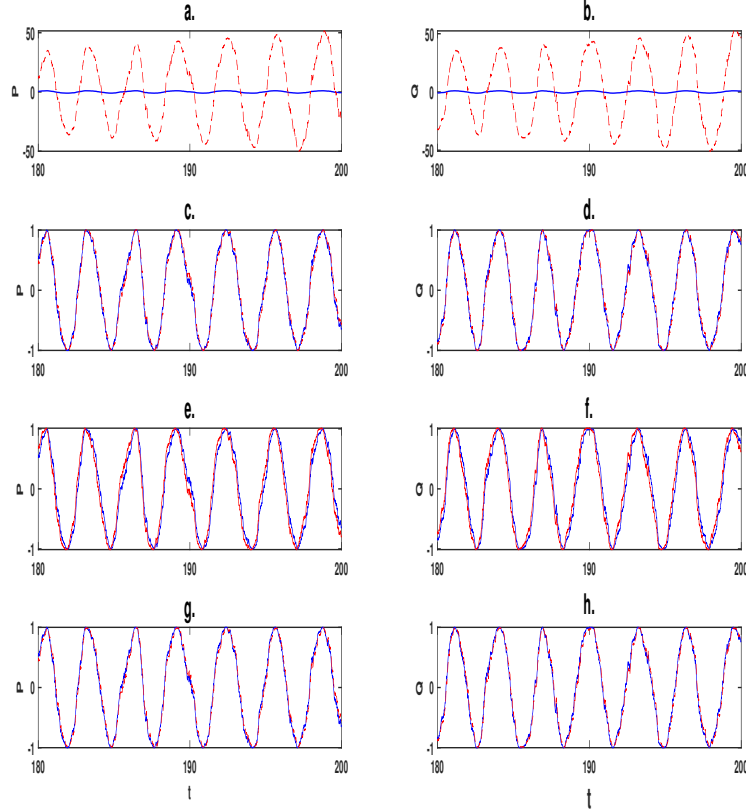


Figure 1: Sample trajectories of (53) with $a = 2$, $\sigma = 0.3$, $p = 1$ and $q = 0$ obtained from the exact solution (blue, solid line) and from numerical methods with time step $h = 0.01$ (red, dashed line): (a), (b) the explicit Milstein first order scheme; (c), (d) the symplectic midpoint scheme; (e), (f) the SRK scheme of order 1 with the tableau 5 with $\alpha_2 = 1/2$; (g) (h) the pseudo-symplectic SRK scheme of order (1,1.5) with the tableau 7 with $a_1 = b_1 = 1/2$.

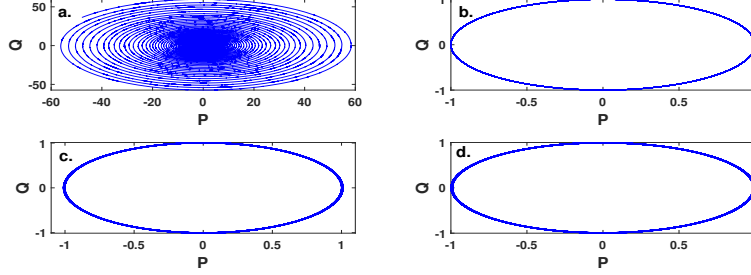


Figure 2: A sample phase trajectory of (53) with $a = 2$, $\sigma = 0.3$, $p = 1$, $q = 0$, $h = 0.01$, on the time interval $0 \leq t \leq 200$: (a) the explicit Milstein first order scheme; (b) the symplectic midpoint scheme; (c) the SRK scheme of order 1 with the tableau 5 with $\alpha_2 = 1/2$; (d) the pseudo-symplectic SRK scheme of order (1,1.5) with the tableau 7 with $a_1 = b_1 = 1/2$.

for the corresponding one-step approximation $X_1 = (P_1, Q_1) \in \mathbb{R}^2$ we have

$$\left\| \left(\frac{\partial X_1}{\partial x} \right)^T J \left(\frac{\partial X_1}{\partial x} \right) - J \right\|_{L^2(\Omega)} = \mathcal{O}(h^2), \quad x = (p, q) \in \mathbb{R}^2. \quad (54)$$

From Theorem 1 we also get near preservation of quadratic invariants, so we have

$$P_{n+1}^2 + Q_{n+1}^2 = P_n^2 + Q_n^2 + \mathcal{O}(h^2), \quad n = 0, 1, \dots, K-1. \quad (55)$$

We also consider a pseudo-symplectic scheme of order (1,1.5) with the tableau 7 with $a_1 = b_1 = 1/2$. For this scheme we have near preservation of quadratic invariants (55), too. Instead of (54) we have the symplectic error (9) of order $\mathcal{O}(h^{2.5})$.

In Fig. 1 we compare the exact solution with the approximations obtained with these schemes, the explicit Milstein first order scheme [17, chapter 10.3] and the symplectic midpoint scheme with the tableau 3 (which has mean square order 1 for SHSs with commutative noise or with one noise). Here we include comparisons only with the symplectic midpoint scheme, but we have obtained similar results for the comparisons with the symplectic schemes constructed using generating functions in [3]. We can notice that while the Milstein scheme fails for $h = 0.01$, the pseudo-symplectic and the symplectic schemes give accurate long term results. The sample phase trajectories of (53) presented in Fig.2 illustrate the preservation of quadratic invariants

by the symplectic midpoint scheme, and the near preservation of quadratic invariants by the schemes corresponding to tableaux 5 and 7. The phase trajectory of the Milstein scheme deviates drastically from the exact $P(t)^2 + Q(t)^2 = 1$, but the order 1 SRK scheme corresponding to the tableau 5 and the pseudo-symplectic SRK scheme of order (1,1.5) with the tableau 7 produce accurate numerical solutions that are similar with the ones given by the symplectic midpoint scheme.

Figure 3 illustrates the convergence curves for the symplectic error (9) for these SRK schemes applied to the SHS (53) with $a = 2$, $\sigma = 0.3$, $p = 1$ and $q = 0$. The error is averaged over 100 000 trajectories. There is a good match between the symplectic error (54) for the SRK scheme corresponding to the tableau 5 and the order 2 reference line plotted with a blue dashed line (- -). We have a similar conclusion about the order 2.5 reference line, plotted with a red dotted line (-.), and the convergence curve for the symplectic error for the pseudo-symplectic scheme of order (1, 1.5) with the tableau 7.

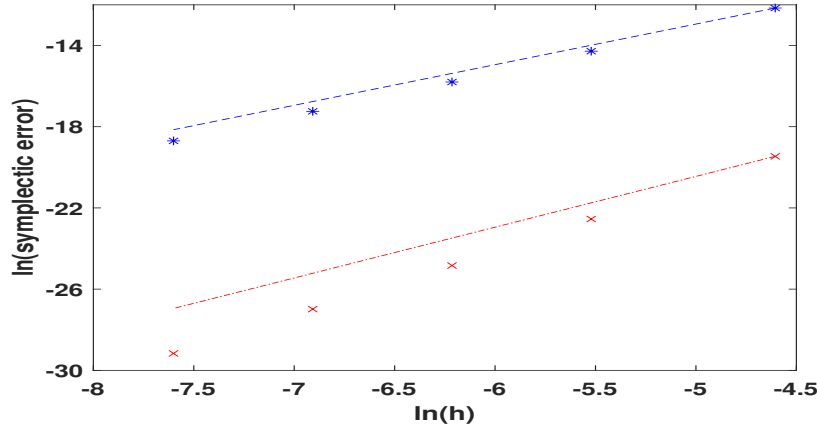


Figure 3: Convergence curves for the symplectic error for SRK of order 1 with the tableau 5 with $\alpha_2 = 1/2$ (blue *); pseudo-symplectic SRK scheme of order (1,1.5) with the tableau 7 with $a_1 = b_1 = 1/2$ (red x).

Fig. 4 confirms that the pseudo-symplectic schemes have the expected convergence rate. The order 1 reference line is plotted with a dashed line (- -). We can notice that the pseudo-symplectic SRK scheme with the tableau 7 has mean square order 1 and gives accurate results, very close to the one obtained with the symplectic midpoint scheme. Fig. 4 also shows that the

SRK scheme with the tableau 5 has mean square order 1 and produce very accurate results.

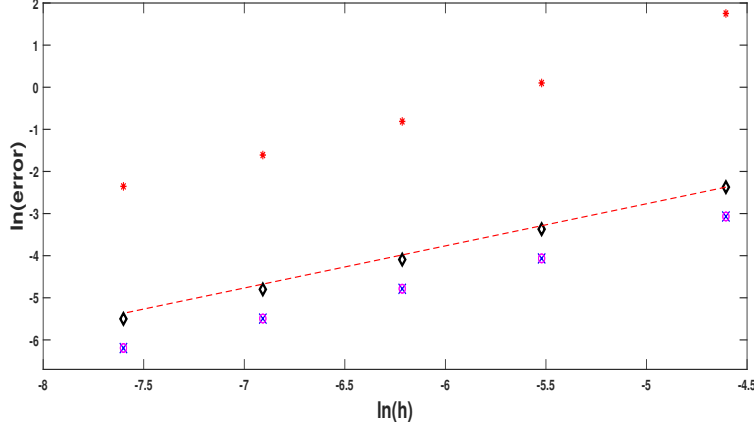


Figure 4: Convergence rate for (53) with $a = 2$, $\sigma = 0.3$, $p = 1$ and $q = 0$, where error is the maximum error of (P, Q) at $T = 100$ averaged over 100 000 trajectories for: red *: Milstein scheme of order 1; blue $\times$: symplectic midpoint scheme; black $\diamond$: SRK scheme of order 1 with the tableau 5 with $\alpha_2 = 1/2$; pink $\circ$: Pseudo-symplectic SRK scheme of order (1,1.5) with the tableau 7 with $a_1 = b_1 = 1/2$.

6.2. Nonlinear stochastic oscillator

We consider the more complex SHS given by the following double well problem [7] with two multiplicative noises:

$$dP = (Q - Q^3)dt - \sigma_1 P dW_1(t) - \sigma_2 Q dW_2(t) \quad (56)$$

$$dQ = P dt + \sigma_1 Q dW_1(t). \quad (57)$$

The accuracy of a pseudo symplectic scheme of order (0.5,1) based on generating functions was illustrated in [2] and a version of this SHS with one additive noise was considered in [21] to show the performance of a pseudo symplectic scheme of order (1,2).

For this nonlinear system with non-commutative noise we cannot find explicitly its exact solution. In Fig. 5 we compare the midpoint symplectic scheme with the tableau 3 of mean square order 0.5, the pseudo-symplectic SRK method of mean square order (0.5,1) with the tableau 5 with $\alpha_2 = 1/2$, and the fully implicit backward Euler scheme [17]. We can clearly

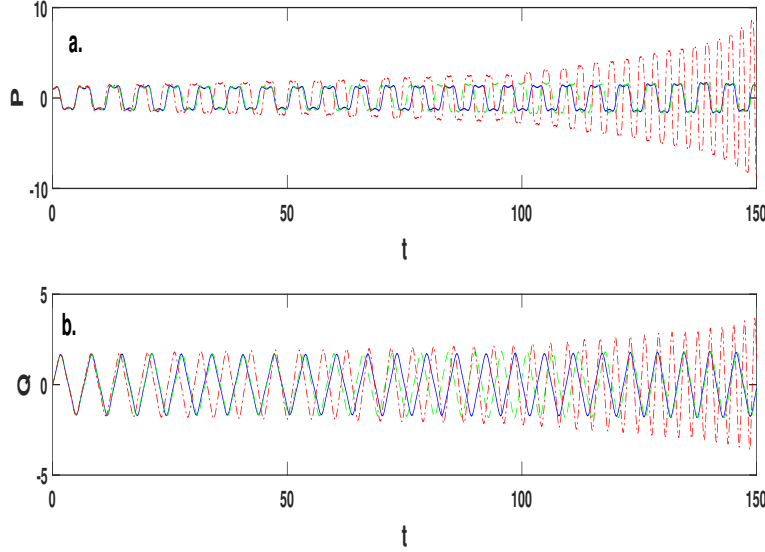


Figure 5: Sample trajectories for the solution of (56)-(57) with $h = 0.005$, $\sigma_1 = \sigma_2 = 0.01$: the symplectic method of strong order 0.5 (blue plain line), the pseudo-symplectic SRK method of mean square order (0.5,1) with the tableau 5 with $\alpha_2 = 1/2$ (green, dashed line), and the backward Euler method (red, dot-dashed line)

conclude that the pseudo-symplectic method gives similar results with the symplectic method, while the backward Euler scheme fails to approximate the oscillations of the nonlinear SHS (56)-(57) when $t > 80$. We mention that in this case the forward and the backward Euler methods give almost identical results.

7. Conclusions

We present a condition, in terms of colored trees, to have near preservation of quadratic invariants by a SRK scheme, and for stochastic Hamiltonian systems we use this condition to find a systematic way of constructing explicit SRK pseudo-symplectic schemes. We show that the same condition can be used to reduce substantially the number of trees that have to be considered to establish the order of convergence of a SRK method.

Similarly with the explicit pseudo-symplectic schemes based on generating functions proposed in [2], overall the explicit pseudo-symplectic SRK schemes offer a good compromise between computational time and accuracy,

because they are more accurate than the explicit strong Taylor approximation methods and less computationally expensive than the implicit symplectic schemes. However, the explicit pseudo-symplectic schemes based on generating functions require the evaluations of many derivatives of the Hamiltonian functions, and the number of these derivatives increases with the pseudo-symplectic order M . Thus the explicit pseudo-symplectic SRK scheme are a good derivative-free alternative.

A limitation of the method proposed here is that although we can construct explicit pseudo-symplectic SRK schemes of any order (N, M) , the number of conditions increases rapidly when we increase the order. However, unlike the deterministic case, for stochastic differential equations order 2 is considered a high order. Schemes with mean square order higher than 2 are not computationally attractive because they require generating 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

Proof. Conditions (20) and (21) implies that the SRK scheme given by (4)-(5) has mean-square order N . It remains to prove that we also have (9).

We consider the first step of the explicit version of scheme (4)-(5) applied to the stochastic system formed by the SHS (7) written as the $2d$ -dimensional autonomous SDE (27) and the stochastic variational equation (28):

$$X_1 = x_0 + \sum_{l=0}^m \sum_{i=1}^s \gamma_i^{(l)} g_l(x_i), \quad (\text{A.1})$$

$$\Sigma_1 = I_{2d} + \sum_{l=0}^m \sum_{i=1}^s \gamma_i^{(l)} e_l(x_i(t)) M_i, \quad (\text{A.2})$$

$$x_i = x_0 + \sum_{l=0}^m \sum_{j=1}^{i-1} Z_{ij}^{(l)} g_l(x_j), \quad (\text{A.3})$$

$$M_i = I_{2d} + \sum_{l=0}^m \sum_{j=1}^{i-1} Z_{ij}^{(l)} e_l(x_j)) M_j, \quad i = 1, \dots, s. \quad (\text{A.4})$$

From [16, Theorem 4.2] we know that $\frac{\partial X_1(h, x_0)}{\partial x_0} = \Sigma_1$. Using (A.2), (A.4) we get

$$\begin{aligned} \Sigma_1^\top J \Sigma_1 &= J + \sum_{l=0}^m \sum_{i=1}^s \gamma_i^{(l)} I_{2d}^\top J e_l(x_i(t)) M_i + \sum_{l=0}^m \sum_{i=1}^s \gamma_i^{(l)} M_i^\top (e_l(x_i(t)))^\top J I_{2d} \\ &+ \sum_{l_1, l_2=0}^m \sum_{i, j=1}^s \gamma_i^{(l_1)} \gamma_j^{(l_2)} M_i^\top (e_{l_1}(x_i(t)))^\top J e_{l_2}(x_j(t)) M_j \\ &= J + \sum_{l=0}^m \sum_{i=1}^s \gamma_i^{(l)} M_i^\top J e_l(x_i(t)) M_i + \sum_{l=0}^m \sum_{i=1}^s \gamma_i^{(l)} M_i^\top (e_l(x_i(t)))^\top J M_i \\ &- \sum_{l_1, l_2=0}^m \sum_{i=1}^s \sum_{j=1}^{i-1} \gamma_i^{(l_1)} Z_{ij}^{(l_2)} M_j^\top (e_{l_2}(x_j(t)))^\top J e_{l_1}(x_i(t)) M_i \\ &- \sum_{l_1, l_2=0}^m \sum_{i=1}^s \sum_{j=1}^{i-1} \gamma_i^{(l_1)} Z_{ij}^{(l_2)} M_i^\top (e_{l_1}(x_i(t)))^\top J e_{l_2}(x_j(t)) M_j \\ &+ \sum_{l_1, l_2=0}^m \sum_{i, j=1}^s \gamma_i^{(l_1)} \gamma_j^{(l_2)} M_i^\top (e_{l_1}(x_i(t)))^\top J e_{l_2}(x_j(t)) M_j \\ &= J - \sum_{l_1, l_2=0}^m \sum_{i=1}^s \sum_{j=1}^{i-1} \gamma_i^{(l_2)} Z_{ij}^{(l_1)} M_j^\top (e_{l_1}(x_j(t)))^\top J e_{l_2}(x_i(t)) M_i \\ &- \sum_{l_1, l_2=0}^m \sum_{i=1}^s \sum_{j=1}^{i-1} \gamma_i^{(l_1)} Z_{ij}^{(l_2)} M_i^\top (e_{l_1}(x_i(t)))^\top J e_{l_2}(x_j(t)) M_j \\ &+ \sum_{l_1, l_2=0}^m \sum_{i, j=1}^s \gamma_i^{(l_1)} \gamma_j^{(l_2)} M_i^\top (e_{l_1}(x_i(t)))^\top J e_{l_2}(x_j(t)) M_j \end{aligned} \quad (\text{A.5})$$

Here we have also used $F_k(x_i)^\top = F_k(x_i)$, $J^{-1} = -J = J^\top$, $J^2 = -I_{2d}$, so

$$\begin{aligned} M_i^\top J e_l(x_i(t)) M_i + M_i^\top (e_l(x_i(t)))^\top J M_i &= M_i^\top J J^{-1} F_l(x_i) M_i + M_i^\top F_l(x_i) (J^{-1})^\top J M_i \\ &= M_i^\top F_l(x_i) M_i + M_i^\top F_l(x_i) J^2 M_i = M_i^\top F_l(x_i) M_i - M_i^\top F_l(x_i) M_i = 0 \end{aligned}$$

Recursively replacing in (A.4) we get

$$\begin{aligned}
M_i &= I_{2d} + \sum_{l=0}^m \sum_{j_1=1}^{i-1} Z_{ij_1}^{(l)} e_l(x_{j_1}) + \sum_{l_1, l_2=0}^m \sum_{1 \leq j_2 < j_1 \leq i-1} Z_{ij_1}^{(l_1)} Z_{j_1 j_2}^{(l_2)} \\
&\quad e_{l_1}(x_{j_1}) e_{l_2}(x_{j_2}) + \dots + \sum_{l_1, \dots, l_{i-1}=0}^m \sum_{1 \leq j_{i-1} < \dots < j_1 \leq i-1} Z_{ij_1}^{(l_1)} Z_{j_1 j_2}^{(l_2)} \dots Z_{j_{i-2} j_{i-1}}^{(l_{i-1})} \\
&\quad e_{l_1}(x_{j_1}) \dots e_{l_{i-1}}(x_{j_{i-1}})
\end{aligned} \tag{A.6}$$

For any $l = 0, \dots, m$ let us define recursively for any $u \in T_l$: $G_l([\emptyset]_l)(x_0) = F_l(x_0)$, and $G_l([\tau_1, \dots, \tau_k]_l)(x_0) = F_l^{(k)}(x_0)(F(\tau_1)(x_0), \dots, F(\tau_k)(x_0))$, where $F_l^{(k)}(x_0) \in \mathcal{M}_{2d}(\mathbb{R})$ is defined by columns according to (11). We apply the expansion (18) to each column of $F_l(x_i)$ and using also Remark 1 we get

$$e_l(x_i(t)) = J^{-1} F_l(x_i) = \sum_{\substack{u \in T_l \\ \rho(u) \leq M+1/2 + \rho([\emptyset]_l)}} \alpha(u) \Phi_i^l(u)(h) J^{-1} G_l(u)(x_0) + \mathcal{O}(h^{M+1}). \tag{A.7}$$

Replacing in (A.6) and using $\Phi_i^l(u)(h) = \mathcal{O}(h^{\rho(u) - \rho([\emptyset]_l)})$, $i = 1, \dots, s$, $l = 0, \dots, m$ we have

$$\begin{aligned}
M_i &= I_{2d} + \sum_{l_1=0}^m \sum_{\substack{u \in T_{l_1} \\ \rho(u) \leq M+1/2 + \rho([\emptyset]_{l_1})}} \alpha(u) \sum_{j_1=1}^{i-1} Z_{ij_1}^{(l_1)} \Phi_{j_1}^{l_1}(u)(h) J^{-1} G_{l_1}(u)(x_0) \\
&\quad + \sum_{l_1, l_2=0}^m \sum_{\substack{u_1 \in T_{l_1}, u_2 \in T_{l_2} \\ \rho(u_1) + \rho(u_2) \leq M+1/2 + \rho([\emptyset]_{l_1}) + \rho([\emptyset]_{l_2})}} \alpha(u_1) \alpha(u_2) \sum_{1 \leq j_2 < j_1 \leq i-1} Z_{ij_1}^{(l_1)} Z_{j_1 j_2}^{(l_2)} \\
&\quad \Phi_{j_1}^{l_1}(u_1)(h) \Phi_{j_2}^{l_2}(u_2)(h) J^{-1} G_{l_1}(u_1)(x_0) J^{-1} G_{l_2}(u_2)(x_0) + \dots + \\
&\quad \sum_{l_1, \dots, l_{i-1}=0}^m \sum_{\substack{u_k \in T_{l_k} \\ \sum_{k=1}^{i-1} \rho(u_k) \leq M+1/2 + \sum_{k=1}^{i-1} \rho([\emptyset]_{l_k})}} \alpha(u_1) \dots \alpha(u_{i-1}) \\
&\quad \sum_{1 \leq j_{i-1} < \dots < j_1 \leq i-1} Z_{ij_1}^{(l_1)} Z_{j_1 j_2}^{(l_2)} \dots Z_{j_{i-2} j_{i-1}}^{(l_{i-1})} \Phi_{j_1}^{l_1}(u_1)(h) \dots \Phi_{j_{i-1}}^{l_{i-1}}(u_{i-1})(h) \\
&\quad J^{-1} G_{l_1}(u_1)(x_0) \dots J^{-1} G_{l_{i-1}}(u_{i-1})(x_0) + \mathcal{O}(h^{M+1})
\end{aligned}$$

This implies

$$\begin{aligned}
e_{k_2}(x_i)M_i &= \sum_{\substack{\theta \in T_{k_2} \\ \rho(\theta) \leq M+1/2+\rho([\emptyset]_{k_2})}} \alpha(\theta)\Phi_i^{k_2}(\theta)(h)J^{-1}G_{k_2}(\theta)(x_0)+ \\
&+ \sum_{l_1=0}^m \sum_{\substack{\theta \in T_{k_2}, u \in T_{l_1} \\ \rho(\theta)+\rho(u) \leq M+1/2+\rho([\emptyset]_{l_1})+\rho([\emptyset]_{k_2})}} \alpha(\theta)\alpha(u)\Phi_i^{k_2}(\theta)(h)\Phi_i(u)(h) \\
&J^{-1}G_{k_2}(\theta)(x_0)J^{-1}G_{l_1}(u)(x_0)+ \\
&+ \sum_{l_1, l_2=0}^m \sum_{\substack{\theta \in T_{k_2}, u_1 \in T_{l_1}, u_2 \in T_{l_2} \\ \rho(\theta)+\rho(u_1 \circ u_2) \leq M+1/2+\rho([\emptyset]_{l_1})+\rho([\emptyset]_{l_2})+\rho([\emptyset]_{k_2})}} \alpha(\theta)\alpha(u_1)\alpha(u_2) \\
&\Phi_i^{k_2}(\theta)(h)\Phi_i(u_1 \circ u_2)(h)J^{-1}G_{k_2}(\theta)(x_0)J^{-1}G_{l_1}(u_1)(x_0)J^{-1}G_{l_2}(u_2)(x_0) \\
&+ \dots \\
&+ \sum_{l_1, \dots, l_{i-1}=0}^m \sum_{\substack{\theta \in T_{k_2}, u_k \in T_{l_k}, k=1, \dots, i-1 \\ \rho(\theta)+\rho(u_1 \circ (u_2 \circ \dots (u_{i-2} \circ u_{i-1}) \dots)) \leq M+ \\ +1/2+\rho([\emptyset]_{k_2})+\sum_{k=1}^{i-1} \rho([\emptyset]_{l_k}), \\ k=1, \dots, i-1}} \alpha(\theta)\alpha(u_1) \cdots \alpha(u_{i-1}) \\
&\Phi_i^{k_2}(\theta)(h)\Phi_i(u_1 \circ (u_2 \circ \dots (u_{i-2} \circ u_{i-1}) \dots))(h)J^{-1}G_{k_2}(\theta)(x_0) \\
&J^{-1}G_{l_1}(u_1)(x_0) \cdots J^{-1}G_{l_{i-1}}(u_{i-1})(x_0) + \mathcal{O}(h^{M+1})
\end{aligned}$$

From this expansion and $\Phi(\tau)(h) = \mathcal{O}(h^{\rho(\tau)})$ we obtain

$$\begin{aligned}
&\sum_{k_1, k_2=0}^m \sum_{i=1}^s \sum_{j=1}^{i-1} \gamma_i^{(k_2)} Z_{ij}^{(k_1)} M_j^\top (e_{k_1}(x_j))^\top J e_{k_2}(x_i) M_i = \\
&= - \sum_{k_1, k_2=0}^m \sum_{\substack{\theta_1 \in T_{k_1}, \theta_2 \in T_{k_2} \\ \rho(\theta_1)+\rho(\theta_2) \leq M+1/2+\rho([\emptyset]_{k_1})+\rho([\emptyset]_{k_2})}} \alpha(\theta_1)\alpha(\theta_2) \\
&\sum_{i=1}^s \sum_{j=1}^{i-1} \gamma_i^{(k_2)} \Phi_i^{k_2}(\theta_2)(h) Z_{ij}^{(k_1)} \Phi_j^{k_1}(\theta_1)(h) (G_{k_1}(\theta_1)(x_0))^\top J^{-1}G_{k_2}(\theta_2)(x_0) \\
&- \sum_{k_1, k_2, l_1=0}^m \sum_{\substack{\theta_1 \in T_{k_1}, \theta_2 \in T_{k_2}, u \in T_{l_1} \\ \rho(\theta_1)+\rho(\theta_2)+\rho(u) \leq M+1/2+\rho([\emptyset]_{k_1})+\rho([\emptyset]_{k_2})+\rho([\emptyset]_{l_1})}} \alpha(\theta_1)\alpha(\theta_2)\alpha(u)
\end{aligned}$$

$$\begin{aligned}
& \sum_{i=1}^s \sum_{j=1}^{i-1} \gamma_i^{(k_2)} \Phi_i^{k_2}(\theta_2)(h) \Phi_i(u)(h) Z_{ij}^{(k_1)} \Phi_j^{k_1}(\theta_1)(h) \\
& (G_{k_1}(\theta_1)(x_0))^\top J^{-1} G_{k_2}(\theta_2)(x_0) J^{-1} G_{l_1}(u)(x_0) - \dots - \\
& - \sum_{\substack{k_1, k_2, \\ l_1, \dots, l_{N_1} \\ n_1, \dots, n_{N_2}=0}}^m \sum_{\substack{\theta_1 \in T_{k_1}, \theta_2 \in T_{k_2} \\ u_k \in T_{l_k}, k=1, \dots, N_1 \\ v_k \in T_{n_k}, k=1, \dots, N_2 \\ \rho(\theta_1) + \rho(\theta_1) + \\ + \rho(u_1 \circ \dots \circ (u_{N_1-1} \circ u_{N_1}) \dots) + \\ + \rho(v_1 \circ \dots \circ (v_{N_2-1} \circ v_{N_2}) \dots) \\ \leq M+1/2 + \rho([\emptyset]_{k_2}) + \sum_{k=1}^{N_1} \rho([\emptyset]_{l_k}) \\ + \rho([\emptyset]_{k_1}) + \sum_{k=1}^{N_2} \rho([\emptyset]_{n_k})}} \alpha(\theta_1) \alpha(\theta_2) \alpha(u_1) \cdots \alpha(u_{N_1}) \\
& \alpha(v_1) \cdots \alpha(v_{N_2}) \sum_{i=1}^s \sum_{j=1}^{i-1} \gamma_i^{(k_2)} \Phi_i^{k_2}(\theta_2)(h) \Phi_i(u_1 \circ (\dots \circ (u_{N_1-1} \circ u_{N_1}) \dots))(h) \\
& Z_{ij}^{(k_1)} \Phi_j^{k_1}(\theta_1)(h) \Phi_j(v_1 \circ (\dots \circ (v_{N_2-1} \circ v_{N_2}) \dots))(h) \\
& (G_{n_{N_2}}(v_{N_2})(x_0))^\top (J^{-1})^\top \cdots (G_{n_1}(v_1)(x_0))^\top (J^{-1})^\top (G_{k_1}(\theta_1)(x_0))^\top \\
& J^{-1} G_{k_2}(\theta_2)(x_0) J^{-1} G_{l_1}(u_1)(x_0) \cdots J^{-1} G_{l_{N_1}}(u_{N_1})(x_0) - \dots + \mathcal{O}(h^{M+1}) \\
& = - \sum_{k_1, k_2=0}^m \sum_{\substack{\theta_1 \in T_{k_1}, \theta_2 \in T_{k_2} \\ \rho(\theta_1) + \rho(\theta_2) \leq M+1/2}} \alpha(\theta_1) \alpha(\theta_2) \Phi(\theta_2 \circ \theta_1)(h) \\
& (G_{k_1}(\theta_1)(x_0))^\top J^{-1} G_{k_2}(\theta_2)(x_0) \\
& - \sum_{k_1, k_2, l_1=0}^m \sum_{\substack{\theta_1 \in T_{k_1}, \theta_2 \in T_{k_2}, u_1 \in T_{l_1} \\ \rho(\theta_1) + \rho(\theta_2) + \rho(u) \leq M+1/2}} \alpha(\theta_1) \alpha(\theta_2) \alpha(u) \Phi((\theta_2 \circ u_1) \circ \theta_1)(h) \\
& (G_{k_1}(\theta_1)(x_0))^\top J^{-1} G_{k_2}(\theta_2)(x_0) J^{-1} G_{l_1}(u_1)(x_0) - \dots - \\
& - \sum_{\substack{k_1, k_2, \\ l_1, \dots, l_{N_1} \\ n_1, \dots, n_{N_2}=0}}^m \sum_{\substack{\theta_1 \in T_{k_1}, \theta_2 \in T_{k_2} \\ u_k \in T_{l_k}, k=1, \dots, N_1 \\ v_k \in T_{n_k}, k=1, \dots, N_2 \\ \rho(\theta_1) + \rho(\theta_1) + \\ + \rho(u_1 \circ \dots \circ (u_{N_1-1} \circ u_{N_1}) \dots) + \\ + \rho(v_1 \circ \dots \circ (v_{N_2-1} \circ v_{N_2}) \dots) \\ \leq M+1/2}} \alpha(\theta_1) \alpha(\theta_2) \alpha(u_1) \cdots \alpha(u_{N_1}) \\
& \alpha(v_1) \cdots \alpha(v_{N_2})
\end{aligned}$$

$$\begin{aligned} & \Phi((\theta_2 \circ (u_1 \circ (\dots \circ (u_{N_1-1} \circ u_{N_1}) \dots))) \circ (\theta_1 \circ (v_1 \circ (\dots \circ (v_{N_2-1} \circ v_{N_2}) \dots))))(h) \\ & (G_{n_{N_2}}(v_{N_2})(x_0))^\top (J^{-1})^\top \dots (G_{n_1}(v_1)(x_0))^\top (J^{-1})^\top (G_{k_1}(\theta_1)(x_0))^\top \\ & J^{-1}G_{k_2}(\theta_2)(x_0)J^{-1}G_{l_1}(u_1)(x_0) \dots J^{-1}G_{l_{N_1}}(u_{N_1})(x_0) - \dots + \mathcal{O}(h^{M+1}) \end{aligned}$$

Here we have a finite number of terms of the form above because the number of trees N_1 and N_2 in the Butcher products cannot be larger than $2M + 1$. Similarly we get

$$\begin{aligned} & \sum_{k_1, k_2=0}^m \sum_{i=1}^s \sum_{j=1}^{i-1} \gamma_i^{(k_1)} Z_{ij}^{(k_2)} M_i^\top (e_{k_1}(x_i))^\top J e_{k_2}(x_j) M_j = \\ & = - \sum_{k_1, k_2=0}^m \sum_{\substack{\theta_1 \in T_{k_1}, \theta_2 \in T_{k_2} \\ \rho(\theta_1) + \rho(\theta_2) \leq M+1/2}} \alpha(\theta_1) \alpha(\theta_2) \Phi(\theta_1 \circ \theta_2)(h) \\ & (G_{k_1}(\theta_1)(x_0))^\top J^{-1}G_{k_2}(\theta_2)(x_0) \\ & - \sum_{k_1, k_2, l_1=0}^m \sum_{\substack{\theta_1 \in T_{k_1}, \theta_2 \in T_{k_2}, u_1 \in T_{l_1} \\ \rho(\theta_1) + \rho(\theta_2) + \rho(u_1) \leq M+1/2}} \alpha(\theta_1) \alpha(\theta_2) \alpha(u_1) \Phi(\theta_1 \circ (\theta_2 \circ u_1))(h) \\ & (G_{k_1}(\theta_1)(x_0))^\top J^{-1}G_{k_2}(\theta_2)(x_0)J^{-1}G_{l_1}(u_1)(x_0) - \dots - \\ & - \sum_{\substack{k_1, k_2, \\ l_1, \dots, l_{N_1} \\ n_1, \dots, n_{N_2}=0}}^m \sum_{\substack{\theta_1 \in T_{k_1}, \theta_2 \in T_{k_2} \\ u_k \in T_{l_k}, k=1, \dots, N_1 \\ v_k \in T_{n_k}, k=1, \dots, N_2 \\ \rho(\theta_1) + \rho(\theta_2) + \\ + \rho(u_1 \circ (\dots \circ (u_{N_1-1} \circ u_{N_1}) \dots)) + \\ + \rho(v_1 \circ (\dots \circ (v_{N_2-1} \circ v_{N_2}) \dots)) \\ \leq M+1/2}} \alpha(\theta_1) \alpha(\theta_2) \alpha(u_1) \dots \alpha(u_{N_1}) \\ & \alpha(v_1) \dots \alpha(v_{N_2}) \\ & \Phi((\theta_1 \circ (v_1 \circ (\dots \circ (v_{N_2-1} \circ v_{N_2}) \dots))) \circ (\theta_2 \circ (u_1 \circ (\dots \circ (u_{N_1-1} \circ u_{N_1}) \dots))))(h) \\ & (G_{n_{N_2}}(v_{N_2})(x_0))^\top (J^{-1})^\top \dots (G_{n_1}(v_1)(x_0))^\top (J^{-1})^\top (G_{k_1}(\theta_1)(x_0))^\top \\ & J^{-1}G_{k_2}(\theta_2)(x_0)J^{-1}G_{l_1}(u_1)(x_0) \dots J^{-1}G_{l_{N_1}}(u_{N_1})(x_0) - \dots + \mathcal{O}(h^{M+1}), \end{aligned}$$

and

$$\sum_{k_1, k_2=0}^m \sum_{i, j=1}^s \gamma_i^{(k_1)} \gamma_j^{(k_2)} M_i^\top (e_{k_1}(x_i))^\top J e_{k_2}(x_j) M_j =$$

$$\begin{aligned}
&= - \sum_{k_1, k_2=0}^m \sum_{\substack{\theta_1 \in T_{k_1}, \theta_2 \in T_{k_2} \\ \rho(\theta_1) + \rho(\theta_2) \leq M+1/2}} \alpha(\theta_1) \alpha(\theta_2) \Phi(\theta_1)(h) \Phi(\theta_2)(h) \\
&\quad (G_{k_1}(\theta_1)(x_0))^\top J^{-1} G_{k_2}(\theta_2)(x_0) \\
&- \sum_{k_1, k_2, l_1=0}^m \sum_{\substack{\theta_1 \in T_{k_1}, \theta_2 \in T_{k_2}, u_1 \in T_{l_1} \\ \rho(\theta_1) + \rho(\theta_2) + \rho(u_1) \leq M+1/2}} \alpha(\theta_1) \alpha(\theta_2) \alpha(u_1) \Phi(\theta_1)(h) \Phi((\theta_2 \circ u_1))(h) \\
&\quad (G_{k_1}(\theta_1)(x_0))^\top J^{-1} G_{k_2}(\theta_2)(x_0) J^{-1} G_{l_1}(u_1)(x_0) - \dots - \\
&- \sum_{\substack{k_1, k_2, \\ l_1, \dots, l_{N_1} \\ n_1, \dots, n_{N_2}=0}}^m \sum_{\substack{\theta_1 \in T_{k_1}, \theta_2 \in T_{k_2} \\ u_k \in T_{l_k}, k=1, \dots, N_1 \\ v_k \in T_{n_k}, k=1, \dots, N_2 \\ \rho(\theta_1) + \rho(\theta_1) + \\ + \rho(u_1 \circ (\dots \circ (u_{N_1-1} \circ u_{N_1}) \dots)) + \\ + \rho(v_1 \circ (\dots \circ (v_{N_2-1} \circ v_{N_2}) \dots)) \\ \leq M+1/2}} \alpha(\theta_1) \alpha(\theta_2) \alpha(u_1) \cdots \alpha(u_{N_1}) \\
&\quad \alpha(v_1) \cdots \alpha(v_{N_2}) \Phi(\theta_1 \circ (v_1 \circ (\dots \circ (v_{N_2-1} \circ v_{N_2}) \dots)))(h) \\
&\quad \Phi(\theta_2 \circ (u_1 \circ (\dots \circ (u_{N_1-1} \circ u_{N_1}) \dots)))(h) \\
&\quad (G_{n_{N_2}}(v_{N_2})(x_0))^\top (J^{-1})^\top \cdots (G_{n_1}(v_1)(x_0))^\top (J^{-1})^\top (G_{k_1}(\theta_1)(x_0))^\top \\
&\quad J^{-1} G_{k_2}(\theta_2)(x_0) J^{-1} G_{l_1}(u_1)(x_0) \cdots J^{-1} G_{l_{N_1}}(u_{N_1})(x_0) - \dots + \mathcal{O}(h^{M+1})
\end{aligned}$$

We replace the previous three expressions in (A.5), group the similar terms, and use (29) and we get $\Sigma_1^\top J \Sigma_1 = J + \mathcal{O}(h^{M+1})$. Thus the scheme is a pseudo-symplectic explicit SRK method of mean-square order (N, M) . $\square$

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 methods based on generating functions for stochastic Hamiltonian systems, J. Comput. Appl. Math. 373 (2020) 112433. Numerical Analysis and Scientific Computation with Applications.

- [3] C. Anton, Y. Wong, J. Deng, Symplectic schemes for stochastic Hamiltonian systems preserving Hamiltonian functions, *Int. J. Numer. Anal. Model.* 11 (2014) 427–451.
- [4] A. Aubry, P. Chartier, Pseudo-symplectic Runge-Kutta methods, *BIT Numer Math* 38 (1998) 439–461.
- [5] P. Bochev, C. Scovel, On quadratic invariants and symplectic structure, *BIT Numerical Mathematics* 34 (1994) 337–345.
- [6] K. Burrage, P. Burrage, General order conditions for stochastic Runge-Kutta methods for both commuting and non-commuting stochastic ordinary differential equation systems, *Appl. Numer. Math.* 28 (1998) 161–177.
- [7] 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.
- [8] K. Burrage, P.M. Burrage, High strong order explicit Runge-Kutta methods for stochastic ordinary differential equations, *Appl. Numer. Math.* 20 (1996) 1–21.
- [9] J.C. Butcher, *Numerical Methods for Ordinary Differential Equations*, John Wiley, 2016. 3rd ed.
- [10] G.J. Cooper, Stability of Runge-Kutta Methods for Trajectory Problems, *IMA Journal of Numerical Analysis* 7 (1987) 1–13.
- [11] 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.
- [12] J. Deng, C. Anton, Y. Wong, High-order symplectic schemes for stochastic Hamiltonian systems, *Commun. Comput. Phys.* 16 (2014) 169–200.
- [13] Z. Ge, K. Feng, On the approximation of linear H-systems, *J. Comput. Math* 6 (1988) 88–97.
- [14] E. Hairer, C. Lubich, G. Wanner, *Geometric numerical integration: structure-preserving algorithms for ordinary differential equations*, Springer, Berlin, 2006.

- [15] 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.
- [16] J. Hong, D. Xu, P. Wang, Preservation of quadratic invariants of stochastic differential equations via Runge-Kutta methods, *Appl. Numer. Math.* 87 (2015) 38–52.
- [17] P. Kloeden, E. Platen, Numerical solutions of stochastic differential equations, Springer-Verlag, Berlin, 1992.
- [18] X. Mao, Stochastic differential equations and applications, Woodhead Publishing, Philadelphia, 2011. 2nd ed.
- [19] 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.
- [20] 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.
- [21] X. Niu, J. Cui, J. Hong, Z. Liu, Explicit pseudo-symplectic methods for stochastic Hamiltonian systems, *BIT Numer Math* 58 (2018) 163–178.
- [22] 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.
- [23] A. Rößler, Runge-Kutta methods for the strong approximation of solutions of stochastic differential equations, *SIAM J. Numer. Anal.* 48 (2010) 922–952.
- [24] J.M. Sanz-Serna, Runge-Kutta schemes for hamiltonian systems, *BIT Numerical Mathematics* 28 (1988) 877–883.
- [25] J.M. Sanz-Serna, L. Abia, Order conditions for canonical Runge-Kutta schemes, *SIAM Journal on Numerical Analysis* 28 (1991) 1081–1096.
- [26] 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: Sample trajectories of (53) with $a = 2$, $\sigma = 0.3$, $p = 1$ and $q = 0$ obtained from the exact solution (blue, solid line) and from numerical methods with time step $h = 0.01$ (red, dashed line): (a), (b) the explicit Milstein first order scheme; (c), (d) the symplectic midpoint scheme; (e), (f) the SRK scheme of order 1 with the tableau 5 with $\alpha_2 = 1/2$; (g) (h) the pseudo-symplectic SRK scheme of order (1,1.5) with the tableau 7 with $a_1 = b_1 = 1/2$.
- Fig. 2: A sample phase trajectory of (53) with $a = 2$, $\sigma = 0.3$, $p = 1$, $q = 0$, $h = 0.01$, on the time interval $0 \leq t \leq 200$: (a) the explicit Milstein first order scheme; (b) the symplectic midpoint scheme; (c) the SRK scheme of order 1 with the tableau 5 with $\alpha_2 = 1/2$; (d) the pseudo-symplectic SRK scheme of order (1,1.5) with the tableau 7 with $a_1 = b_1 = 1/2$.
- Fig. 3: Convergence curves for the symplectic error for SRK of order 1 with the tableau 5 with $\alpha_2 = 1/2$ (blue *); pseudo-symplectic SRK scheme of order (1,1.5) with the tableau 7 with $a_1 = b_1 = 1/2$ (red $\times$)
- Fig. 4: Convergence rate for (53) with $a = 2$, $\sigma = 0.3$, $p = 1$ and $q = 0$, where error is the maximum error of (P, Q) at $T = 100$ averaged over 100 000 trajectories for: red *: Milstein scheme of order 1; blue *times*: symplectic midpoint scheme; black $\diamond$: SRK scheme of order 1 with the tableau 5 with $\alpha_2 = 1/2$; pink $\circ$: Pseudo-symplectic SRK scheme of order (1,1.5) with the tableau 7 with $a_1 = b_1 = 1/2$.
- Fig. 5: Sample trajectories for the solution of (56)-(57) with $h = 0.005$, $\sigma_1 = \sigma_2 = 0.01$: the symplectic method of strong order 0.5 (blue plain line), the pseudo-symplectic SRK method of mean square order (0.5,1) with the tableau 5 with $\alpha_2 = 1/2$ (green, dashed line), and the backward Euler method (red, dot-dashed line)