Commun. Comput. Phys.
doi: 10.4208/cicp.311012.191113a

Vol. **16**, No. 1, pp. 169-200
July 2014

High-Order Symplectic Schemes for Stochastic Hamiltonian Systems

Jian Deng[1,*], Cristina Anton[2] and Yau Shu Wong[1]

[1] *Department of Mathematical and Statistical Sciences, University of Alberta, Edmonton, AB T6G 2G1, Canada.*
[2] *Department of Mathematics and Statistics, Grant MacEwan University, Edmonton, AB T5J 4S2, Canada.*

Received 31 October 2012; Accepted (in revised version) 19 November 2013

Available online 10 April 2014

Abstract. The construction of symplectic numerical schemes for stochastic Hamiltonian systems is studied. An approach based on generating functions method is proposed to generate the stochastic symplectic integration of any desired order. In general the proposed symplectic schemes are fully implicit, and they become computationally expensive for mean square orders greater than two. However, for stochastic Hamiltonian systems preserving Hamiltonian functions, the high-order symplectic methods have simpler forms than the explicit Taylor expansion schemes. A theoretical analysis of the convergence and numerical simulations are reported for several symplectic integrators. The numerical case studies confirm that the symplectic methods are efficient computational tools for long-term simulations.

AMS subject classifications: 60H10, 65C30, 65P10

Key words: Stochastic Hamiltonian systems, symplectic integration, mean-square convergence, high-order schemes.

1 Introduction

The symplectic integration covers a special type of numerical methods which are capable of preserving the symplecticity properties of the Hamiltonian system. The pioneering work on the symplectic integration is due to de Vogelaere [1], Ruth [2] and Kang Feng [3]. Symplectic methods have been applied successfully to deterministic Hamiltonian systems, and numerical simulations consistently show that the most important feature of this approach is that the accuracy of the computed solution is guaranteed even for long

*Corresponding author. *Email addresses:* deng2@ualberta.ca (J. Deng), popescuc@macewan.ca (C. Anton), yaushu.wong@ualberta.ca (Y. S. Wong)

term computation [4,5]. In this paper, we study the symplectic numerical integration for stochastic Hamiltonian systems, and propose an approach to generate symplectic numerical schemes of any desired order.

Consider the stochastic differential equations (SDEs) in the sense of the Stratonovich:

$$dP_i = -\frac{\partial H^{(0)}(P,Q)}{\partial Q_i}dt - \sum_{r=1}^{m}\frac{\partial H^{(r)}(P,Q)}{\partial Q_i}\circ dw_t^r, \qquad P(t_0)=p, \tag{1.1a}$$

$$dQ_i = \frac{\partial H^{(0)}(P,Q)}{\partial P_i}dt + \sum_{r=1}^{m}\frac{\partial H^{(r)}(P,Q)}{\partial P_i}\circ dw_t^r, \qquad Q(t_0)=q, \tag{1.1b}$$

where P, Q, p, q are n-dimensional vectors with components P^i, Q^i, p^i, q^i, $i=1,\cdots,n$ and w_t^r, $r=1,\cdots,m$ are independent standard Wiener Processes. The SDEs (1.1) are called the Stochastic Hamiltonian System (SHS) (see [6]). The SHS (1.1) includes both Hamiltonian systems with additive or multiplicative noise.

A non-autonomous SHS is given by time-dependent Hamiltonian functions $H^{(r)}(t,P,Q)$, $r=0,\cdots,m$. However, it can be rewritten as an autonomous SHS by introducing new variables e_k and f_k. Let

$$df_r = dt, \quad de_r = -\frac{\partial H^{(r)}(t,P,Q)}{\partial t}\circ dw_t^r, \quad (\text{where } dw_t^0 := dt),$$

with the initial condition $e_r(t_0) = -H^{(r)}(t_0,p,q)$ and $f_r(t_0) = t_0$, $r=0,\cdots,m$. Then the new Hamiltonian functions $\bar{H}^{(r)}(\bar{P},\bar{Q}) = H^{(r)}(f_r,P,Q)$, $r=1,\cdots,m$, and $\bar{H}^{(0)}(\bar{P},\bar{Q}) = H^{(0)}(f_r,P,Q)+e_0+\cdots+e_m$, define an autonomous SHS with $\bar{P}=(P^T,e_0,\cdots,e_m)^T$ and $\bar{Q}=(Q^T,f_0,\cdots,f_m)^T$. Hence, in this study, we will only investigate the autonomous case as given in (1.1).

There are growing interests and efforts on the theoretical study and computational implementation of numerical methods for SHS [6,8–11]. Milstein et al. [6,8] introduced the symplectic numerical schemes to SHS, and demonstrated the superiority of the symplectic methods for long time computation. Although they proposed symplectic schemes of orders two or three for special types of SHS, for the general SHS with multiplicative noise given in (1.1), they construct only symplectic schemes of mean square order 0.5. In this paper we apply an approach based on generating functions and we construct symplectic schemes of arbitrary high mean square order. Hong et al. [9] developed a predictor-corrector scheme for a linear SDE with an additive noise, a simple case of SHS. In [11], Wang et al. proposed the variational integrators to construct the stochastic symplectic schemes.

The generating functions associated with the SHS (1.1) were rigorously introduced in [15]. Recently, Wang et al. [12–14] proposed generating functions methods to construct symplectic schemes for SHS. But, in those papers, only the product of one-fold Stratonovich integrals is considered, so their approach cannot be used to construct stochastic symplectic scheme higher than order 1.5. In this paper, the product of multiple Stratonovich integrals (Property 3.1) is studied, such that the generating functions can be used to construct symplectic schemes of arbitrary order.

We follow the rigorous approach presented in [15], and employ the properties of multiple stochastic integrals to derive a recursive formula for determining the coefficients of the generating function. Theoretically, this formula allow us to construct stochastic symplectic schemes of arbitrary high order with corresponding conditions on the Hamiltonian functions. Hence, the major contribution of the work reported here is to present a way for finding the coefficients of the generating functions such that the generating function method could be used to construct stochastic symplectic schemes of any order. Since the computation complexity increases with the order of the numerical schemes, we mainly focus on the symplectic schemes with mean square order 1 for which we also present the convergence analysis. Moreover, for special types of SHSs, such as SHSs with additive noise, SHSs with separable Hamiltonians, or SHS preserving the Hamiltonian functions, we construct computationally attractive symplectic schemes of mean square order 2. The study of high order stochastic symplectic scheme for the general SHS (1.1) helps to construct high order Runge-Kutta type schemes that avoid higher order derivatives.

The paper is organized as follow. In the beginning, we present the theory of generating functions and Hamiltonian-Jacobi partial differential equations in the stochastic setting. The construction of the symplectic schemes based on approximating of the solution of the corresponding Hamilton-Jacobi stochastic partial differential equations is presented in Section 3. In Section 4, we construct higher order symplectic schemes, and in Section 5 we study the convergence of these schemes. Symplectic schemes for special types of SHSs are included in Section 6. Numerical simulations illustrating the performance of the proposed methods are reported in Section 7.

2 Generating function and Stochastic Hamilton-Jacobi partial differential equation

The generating function applied to the deterministic Hamiltonian system has been well studied. However, the extension to stochastic cases is a challenging task.

We denote the solution of the SHS (1.1) by

$$X(t;t_0,x_0;\omega)=\big(P^T(t;t_0,p,q;\omega),Q^T(t;t_0,p,q;\omega)\big)^T,$$

where $t_0\le t\le t_0+T$, and ω is an elementary event. It is known that if $H^{(j)}$, $j=0,\cdots,m$, are sufficiently smooth, then $X(t;t_0,x_0;\omega)$ is a phase flow (diffeomorphism) for almost any ω (see [7]). To simplify the notation, we will remove any mentioning of the dependence on ω unless it is absolutely necessary to avoid confusions, and we make the convention to understand that all the equations involving the solution of the SHS (1.1) are true for almost any ω.

In differential geometry, the differential 1-form of a function $f:\mathbb{R}^{2n}\to\mathbb{R}$ on $\xi\in\mathbb{R}^{2n}$ is defined as:

$$df(\xi):=\sum_{i=1}^{2n}\frac{\partial f}{\partial z_i}\xi_i. \tag{2.1}$$

The exterior product $df\wedge dg$ of $\xi,\eta\in\mathbb{R}^{2n}$ is given by $df\wedge dg(\xi,\eta)=df(\xi)dg(\eta)-df(\eta)dg(\xi)$, and represents the oriented area of the image of the parallelogram with sides $df(\xi)$ and $dg(\eta)$ on the $df(\xi)$, $dg(\eta)$-plane.

The stochastic flow $(p,q)\to(P,Q)$ of the SHS (1.1) preserves the symplectic structure (Theorem 2.1 in [6]) as follows:

$$dP\wedge dQ=dp\wedge dq, \tag{2.2}$$

i.e., the sum over the oriented areas of its projections onto the two dimensional plane (p^i,q^i) is invariant. Here we consider the differential 2-form

$$dp\wedge dq=dp^1\wedge dq^1+\cdots+dp^n\wedge dq^n, \tag{2.3}$$

and the differentiation in (1.1) and (2.2) have different meanings: in (1.1) p, q are fixed parameters and differentiation is done with respect to time t, while in (2.2) differentiation is carried out with respect to the initial data p, q. We say that a method based on the one step approximation $\bar{P}=\bar{P}(t+h;t,p,q)$, $\bar{Q}=\bar{Q}(t+h;t,p,q)$ preserves the symplectic structure if

$$d\bar{P}\wedge d\bar{Q}=dp\wedge dq. \tag{2.4}$$

The previous definition of symplecticity can be extended to general random maps. A random map $\varphi(\omega,x)$ is a map with the property that for any fixed $x\in\mathbb{R}^{2n}$, $\varphi(\cdot,x)$ is a random variable. We denote $\varphi_\omega(\cdot)=\varphi(\omega,\cdot)$. According to (2.2) a random differentiable map $\varphi_\omega:(p,q)\to(P,Q)$ is symplectic if and only if $dP\wedge dQ=dp\wedge dq$ a.s. We can easily prove the following equivalent property.

Proposition 2.1. A differentiable random map $\varphi_\omega:U\to\mathbb{R}^{2n}$ (where $U\subset\mathbb{R}^{2n}$ is an open set) is symplectic if and only if the Jacobian matrix φ'_ω satisfies

$$\varphi'(p,q)^T J\varphi'(p,q)=J\quad\text{with}\ J=\begin{bmatrix}0 & I\\ -I & 0\end{bmatrix}, \tag{2.5}$$

for almost any ω and any $p,q\in\mathbb{R}^n$, $(p,q)^T\in U$, where I is the identity matrix of dimension n.

In the deterministic case, generating functions are powerful tools to study symplectic transformations. The next lemma introduces the generating functions S_ω, S^i_ω, $i=1,2,3$ in the stochastic case.

Lemma 2.1. *Let $\varphi_\omega:(p,q)\to(P,Q)$ be a smooth random map from $\mathbb{R}^{2n}$ to $\mathbb{R}^{2n}$. Then φ_ω is symplectic if any of the following statements is true:*

1. *There exists locally a smooth random map S_ω from $\mathbb{R}^{2n}$ to $\mathbb{R}^{2n}$ such that $\partial(S_\omega)/\partial q\partial Q$ is invertible a.s. and we have*

$$P^i=\frac{\partial S}{\partial Q^i}(q,Q),\quad p^i=-\frac{\partial S}{\partial q^i}(q,Q),\quad i=1,\cdots,n. \tag{2.6}$$

2. *There exists locally a smooth random map S^1_ω from $\mathbb{R}^{2n}$ to $\mathbb{R}^{2n}$ such that $\partial(P^Tq+S^1_\omega)/\partial P\partial q$ is invertible a.s. and we have*

$$p^i=P^i+\frac{\partial S^1_\omega}{\partial q^i}(P,q),\quad Q^i=q^i+\frac{\partial S^1_\omega}{\partial P^i}(P,q),\quad i=1,\cdots,n. \tag{2.7}$$

3. *There exists locally a smooth random map S^2_ω from $\mathbb{R}^{2n}$ to $\mathbb{R}^{2n}$ such that $\partial(p^TQ+S^2_\omega)/\partial p\partial Q$ is invertible a.s. and we have*

$$q^i=Q^i+\frac{\partial S^2_\omega}{\partial p^i}(p,Q),\quad P^i=p^i+\frac{\partial S^2_\omega}{\partial Q^i}(p,Q),\quad i=1,\cdots,n. \tag{2.8}$$

4. *There exists locally a smooth random map S^3_ω from $\mathbb{R}^{2n}$ to $\mathbb{R}^{2n}$ such that $\partial((P+p)^T(Q-q)-2S^3_\omega)/\partial Y\partial y$ is invertible a.s. and we have*

$$Y=y-J\nabla S^3_\omega((y+Y)/2), \tag{2.9}$$

where $Y=(P^T,Q^T)^T$, $y=(p^T,q^T)^T$.

Proof. The proof can be completed easily by adapting the proof of Theorem 3.1 in [8]. □

The previous lemma gives us a powerful tool to analyze the symplectic structure and to construct symplectic methods. For instance, for the SHS (1.1), if in the relation (2.7) we let

$$S^1_\omega=hH^{(0)}(P,q)+\frac{h}{2}\sum_{r=1}^{m}\sum_{j=1}^{n}\frac{\partial H^{(r)}}{\partial q_j}(P,q)\frac{\partial H^{(r)}}{\partial P_j}(P,q)+\sum_{r=1}^{m}\sqrt{h}\xi_rH^{(r)}(P,q), \tag{2.10}$$

where h is the time step and ξ_r are independent bounded random variables such that $E(\xi_r-\xi)^2\le h$, with $\xi\sim N(0,1)$, then we obtain the symplectic Euler scheme proposed by Milstein et al. [6,8]. Lemma 2.1 guarantees that the numerical scheme is symplectic. Moreover the implicit midpoint scheme in [8] is obtained by setting

$$S^3_\omega=hH^{(0)}((y+Y)/2)+\sum_{r=1}^{m}\sqrt{h}\xi_rH^{(r)}((y+Y)/2) \tag{2.11}$$

in relation (2.9).

We introduce the stochastic Hamilton-Jacobi partial differential equation (HJ PDE) associated with the SHS (1.1) following the rigorous approach from [15]. We want to consider the effect of time in the generating function S_ω, so let $S_\omega(x,t)$ be a family of real valued processes with parameters $x\in\mathbb{R}^{2n}$. We can regard it as random field with double parameters x and t. If $S_\omega(x,t)$ is a $\mathcal{C}^\infty$ function of x for almost all ω for any t, we can regard $S_\omega(x,t)$ as a $\mathcal{C}^\infty$ value process [7].

Let assume that the Hamiltonian function $H^{(r)}$ for $r=0,\cdots,m$ in (1.1) belong to $\mathcal{C}^\infty$. In addition, we also suppose that:

$$\begin{aligned}&\sum_{r=0}^{m}(|\nabla_p H^{(r)}(P,Q)-\nabla_p H^{(r)}(p,q)|+|\nabla_q H^{(r)}(P,Q)-\nabla_q H^{(r)}(p,q)|)\\&\le L_1(|P-p|+|Q-q|)\end{aligned}\tag{2.12}$$

and

$$\sum_{r=0}^{m}(|\nabla_p H^{(r)}(p,q)|+|\nabla_q H^{(r)}(p,q)|)\le L_2(1+|p|+|q|).\tag{2.13}$$

So the Lipschitz condition (2.12) and linear growth bound (2.13) guarantees the local existence and uniqueness of the solution $(P(t,\omega)^T,Q(t,\omega)^T)^T$ of the SHS (1.1). Moreover, it is known that $X(t;t_0,x_0;\omega)=\big(P^T(t;t_0,p,q;\omega),\,Q^T(t;t_0,p,q;\omega)\big)^T$, where $t_0\le t\le t_0+T$ is a diffeomorphism a.s. [7]. Thus the generating function, which is a random mapping, becomes a stochastic process $S(q,Q,t,\omega)$, and through Eq. (2.6), this stochastic process generates the symplectic map $(p,q)\to(P(t,\omega),Q(t,\omega))$ of the flow of the SHS.

For the sake of simplification, let us keep the notation S_ω for the stochastic process $S(q,Q,t,\omega)$. The generating function S_ω is connected with the SHS (1.1) by the Hamilton-Jacobi partial differential equation (HJ PDE) (see Theorem 6.14 in [15]):

$$\begin{aligned}dS_\omega=&-H^{(0)}\Big(\frac{\partial S_\omega}{\partial Q_1},\cdots,\frac{\partial S_\omega}{\partial Q_n},Q_1,\cdots,Q_n\Big)dt\\&-\sum_{r=1}^{m}H^{(r)}\Big(\frac{\partial S_\omega}{\partial Q_1},\cdots,\frac{\partial S_\omega}{\partial Q_n},Q_1,\cdots,Q_n\Big)\circ dw_t^r,\end{aligned}\tag{2.14}$$

with the initial condition $S_\omega(q,Q,0)=j(q,Q)$, where j is a C^∞ function. Starting from the flow $X(t;t_0,x_0;\omega)$ of the SHS (1.1) and using the method of characteristics, in Theorem 6.14 and its corollary in [15] it is shown that for any initial point x_0 there exists a stopping time $\tau>t_0$ a.s and a local solution $S_\omega(q,Q,t)$, $t_0\le t<\tau$ of (2.14) for which we have the equations given in (2.6). Moreover, almost sure the flow $X(t;t_0,x_0;\omega)$ is a local Stratonovich semi-martingale, and $S_\omega(q,Q,t)$, $\partial S_\omega(q,Q,t)/\partial Q$ and $\partial S_\omega(q,Q,t)/\partial q$ are local Stratonovich semi-martingale, continuous on (q,Q,t), and $\mathcal{C}^\infty$ value processes (see also Theorem 6.1.5 in [7]).

Theorem 2.1. *Let $S_\omega(q,Q,t)$ be a local solution of the HJ PDE* (2.14) *with initial values satisfying*

$$\frac{\partial j}{\partial q_i}(q,q)+\frac{\partial j}{\partial Q_i}(q,q)=0,\quad i=1,\cdots,n,$$

and such that almost sure $S_\omega(q,Q,t)$, $\partial S_\omega(q,Q,t)/\partial Q$ and $\partial S_\omega(q,Q,t)/\partial q$ are local Stratonovich semi-martingale, continuous on (q,Q,t), and $\mathcal{C}^\infty$ value processes. If there exists a stopping time $\tau' > t_0$ a.s. such that the matrix $(\partial^2(S_\omega)/\partial q\partial Q)$ is a.s. invertible for $t_0 \le t < \tau'$, then the map $(p,q)\to(P(t,\omega),Q(t,\omega))$, $t_0\le t<\tau'$, defined by (2.6) *is the flow of the SHS* (1.1).

Proof. The mapping $(p,q)\to(P_\omega,Q_\omega)$ is well defined by (2.6) because of the invertibility of the matrix $(\partial^2(S_\omega)/\partial q\partial Q)$ for $t_0\le t<\tau'$, and the implicit function theorem.

Differentiation of the second equation of (2.6) (see Theorem 3.3.2 in [7]) yields

$$d\Big(\frac{\partial S_\omega}{\partial q_i}\Big)+\sum_{j=1}^{n}\frac{\partial^2 S_\omega}{\partial q_i\partial Q_j}dQ_j=0. \tag{2.15}$$

Recalling that S_ω is the solution of the stochastic HJ PDE (2.14), the following equation holds by differentiating (2.14) with respect to q_i (see the Corollary of Theorem 6.14 in [15]).

$$d\Big(\frac{\partial S_\omega}{\partial q_i}\Big)+\sum_{j=1}^{n}\frac{\partial H^{(0)}}{\partial P_j}\frac{\partial^2 S_\omega}{\partial q_i\partial Q_j}dt+\sum_{r=1}^{m}\sum_{j=1}^{n}\frac{\partial H^{(r)}}{\partial P_j}\frac{\partial^2 S_\omega}{\partial q_i\partial Q_j}\circ dw_t^r=0. \tag{2.16}$$

Comparing Eqs. (2.15) and (2.16) and using the invertibility of the matrix $(\partial^2 S_\omega/\partial q_i\partial Q_j)$, we have the second equation of (1.1).

The first equation of (1.1) can be obtained using a similar procedure as reported above, i.e., by differentiating the first relation of (2.6) and the HJ PDE with respect to Q_i, then subtracting the obtained equations. Also the initial values guarantee that $(P(t_0,\omega),Q(t_0,\omega))=(p,q)$. $\square$

The HJ PDEs for the coordinate transformations (2) and (4) in Lemma 2.1 can be expressed as

$$S^1_\omega(t,P,q)=\int_{t_0}^{t}H^{(0)}(P,q+\nabla_P S^1_\omega(s,P,q))ds + \int_{t_0}^{t}\sum_{r=1}^{m}H^{(r)}(P,q+\nabla_P S^1_\omega(s,P,q))\circ dw_s^r, \tag{2.17a}$$

$$S^3_\omega(t,w)=\int_{t_0}^{t}H^{(0)}\Big(w+\frac{1}{2}J^{-1}\nabla S^3_\omega(s,w)\Big)ds + \int_{t_0}^{t}\sum_{r=1}^{m}H^{(r)}\Big(w+\frac{1}{2}J^{-1}\nabla S^3_\omega(s,w)\Big)\circ dw_s^r, \tag{2.17b}$$

where $w\in\mathbb{R}^{2n}$, and we consider $S^1_\omega|_{t=t_0}=0$ and $S^3_\omega|_{t=t_0}=0$. It is straightforward to obtain the HJ PDE for the generating function (3) in Lemma 2.1, as it is just the adjoint case of (2).

3 Constructing high-order symplectic schemes

For deterministic problems, the construction of high-order symplectic schemes via generating functions was first proposed by Kang Feng et al. [3,17]. The key idea is to obtain an approximation of the solution of the HJ PDE, and then to construct the symplectic numerical scheme through the relations (2.7)-(2.9).

Following this idea, we now seek an expansion which reflects the stochastic properties of the generating function. Due to the Ito representation theorem, the relation between the Ito integral, the Stratonovich integral and the stochastic Taylor-Stratonovich expansion, it is reasonable to assume that the generating function can be expressed by the following expansion locally:

$$S^1(P,q,t,\theta(t)\omega)=G^1_{(0)}(P,q)J_{(0)}+G^1_{(1)}(P,q)J_{(1)}+G^1_{(0,1)}(P,q)J_{(0,1)}+\cdots=\sum_{\alpha}G^1_{\alpha}J_{\alpha}, \tag{3.1}$$

where $\alpha=(j_1,j_2,\cdots,j_l)$, $j_i\in\{0,1,\cdots,m\}$, $i=1,\cdots,l$ is a multi-index of length l, and J_α is the multiple Stratonovich integral

$$J_\alpha=\int_0^t\int_0^{s_l}\cdots\int_0^{s_2}\circ dw^{j_1}_{s_1}\cdots\circ dw^{j_{l-1}}_{s_{l-1}}\circ dw^{j_l}_{s_l}. \tag{3.2}$$

For convenience, ds is denoted by dw^0_s. Similarly, the multiple Ito stochastic integral I_α is given by

$$I_\alpha=\int_0^t\int_0^{s_l}\cdots\int_0^{s_2}dw^{j_1}_{s_1}\cdots dw^{j_{l-1}}_{s_{l-1}}dw^{j_l}_{s_l}. \tag{3.3}$$

3.1 Properties of multiple stochastic integrals

To prepare for the derivation of the symplectic numerical schemes, we discuss some properties of the multiple stochastic integrals. First, we define some operations for multi-indexes.

If the multi-index $\alpha=(j_1,j_2,\cdots,j_l)$ with $l>1$ then $\alpha-=(j_1,j_2,\cdots,j_{l-1})$, i.e., the last component is deleted. For instance, $(1,3,0)-=(1,3)$. For any two multi-indexes $\alpha=(j_1,j_2,\cdots,j_l)$ and $\alpha'=(j'_1,j'_2,\cdots,j'_{l'})$, we define the concatenation operation $*$ as $\alpha*\alpha'=(j_1,j_2,\cdots,j_l,j'_1,j'_2,\cdots,j'_{l'})$. For example, $(1,3,0)*(1,4)=(1,3,0,1,4)$. The concatenation of a collection Λ of multi-indexes with the multi-index α gives the collection formed by concatenating each element of the collection Λ with the multi-index α, i.e., $\Lambda*\alpha=\{\alpha'*\alpha|\alpha'\in\Lambda\}$. For example, if $\Lambda=\{(1,1),(0,1,2),(1,1)\}$ and $\alpha=(0)$, then $\Lambda*\alpha=\{(1,1,0),(0,1,2,0),(1,1,0)\}$.

Proposition 3.1. For

$$J_\alpha J_{\alpha'}=\sum_{\beta\in\Lambda_{\alpha,\alpha'}}J_\beta, \tag{3.4}$$

where $\alpha=(j_1,j_2,\cdots,j_l)$, $\alpha'=(j'_1,j'_2,\cdots,j'_{l'})$ and $\Lambda_{\alpha,\alpha'}$ is the collection of multi-indexes depending on α and α', and given by the following recurrence relation:

$$\Lambda_{\alpha,\alpha'}=\begin{cases}\{(j_1,j'_1),(j'_1,j_1)\}, & \text{if } l=1 \text{ and } l'=1,\\ \{\Lambda_{(j_1),\alpha'-}*(j'_{l'}),\alpha'*(j_l)\}, & \text{if } l=1 \text{ and } l'\neq 1,\\ \{\Lambda_{\alpha-,(j'_1)}*(j_l),\alpha*(j'_l)\}, & \text{if } l\neq 1 \text{ and } l'=1,\\ \{\Lambda_{\alpha-,\alpha'}*(j_l),\Lambda_{\alpha,\alpha'-}*(j'_{l'})\}, & \text{if } l\neq 1 \text{ and } l'\neq 1.\end{cases} \tag{3.5}$$

Proof. Let consider two stochastic processes

$$X^1_t=\int_0^t b_1(X_s)\circ dw_s^{j_l} \quad \text{and} \quad X^2_t=\int_0^t b_2(X_s)\circ dw_s^{j'_{l'}}. \tag{3.6}$$

Then, for the Stratonovich integrals, we have

$$X^1_tX^2_t=\int_0^t X^2_s b_1(X_s)\circ dw_s^{j_l}+\int_0^t X^1_s b_2(X_s)\circ dw_s^{j'_{l'}}. \tag{3.7}$$

If $l>1$ and $l'>1$, let $b_1(X_s)=J_{\alpha-}$ and $b_2(X_s)=J'_{\alpha'-}$, such that $X^1_t=J_\alpha$ and $X^2_t=J_{\alpha'}$. The product rule (3.7) of the stochastic integrals yields

$$J_\alpha J_{\alpha'}=\int_0^t J_{\alpha'}J_{\alpha-}\circ dw_s^{j_l}+\int_0^t J_\alpha J_{\alpha'-}\circ dw_s^{j'_{l'}}. \tag{3.8}$$

This implies the fourth relation in (3.5).

If $l=1$ (or $l'=1$), the second (or third) relation in the recurrence (3.5) is obtained for $b_1(X_s)=1$ and $b_2(X_s)=J'_{\alpha'-}$ (or $b_1(X_s)=J_{\alpha-}$ and $b_2(X_s)=1$). For the first relation, we take $b_1(X_s)=1$ and $b_2(X_s)=1$. □

For instance, since

$$\begin{aligned}\Lambda_{(2,0),(0,1)}&=\{\Lambda_{(2),(0,1)}*(0),\Lambda_{(2,0),(0)}*(1)\}\\ &=\{\{\Lambda_{(2),(0)}*(1),(0,1,2)\}*(0)\},\{\Lambda_{(2),(0)}*(0),(2,0,0)\}*(1)\}\\ &=\{\{(2,0,1),(0,2,1),(0,1,2)\}*(0),\{(0,2,0),(2,0,0),(2,0,0)\}*(1)\}\\ &=\{(2,0,1,0),(0,2,1,0),(0,1,2,0),(0,2,0,1),(2,0,0,1),(2,0,0,1)\},\end{aligned} \tag{3.9}$$

then we have $J_{(2,0)}J_{(0,1)}=J_{(2,0,1,0)}+J_{(0,2,1,0)}+J_{(0,1,2,0)}+J_{(0,2,0,1)}+2J_{(2,0,0,1)}$.

Remark 3.1. From the recurrence (3.5), we can see that $\Lambda_{\alpha,\alpha'}=\Lambda_{\alpha',\alpha}$, and the length of the multi indexes in $\Lambda_{\alpha,\alpha'}$ is the summation of the lengths of α and α', i.e., if $\beta\in\Lambda_{\alpha,\alpha'}$, then $l(\beta)=l(\alpha)+l(\alpha')$. This will be used to determine the coefficients of the generating function in the next subsection.

Corollary 3.1. *For* $\alpha=(j_1,j_2,\cdots,j_l)$,

$$w_t^j J_\alpha=\sum_{i=0}^{l} J_{(j_1,\cdots,j_i,j,j_{i+1},\cdots,j_l)}. \tag{3.10}$$

Proof. The proof follows by repeatedly applying the second recurrence of (3.5). □

Similarly, we can show that the multiplication of a finite sequence of multiple-indexes can be expressed by the following summation:

$$\prod_{i=1}^{n} J_{\alpha_i}=\sum_{\beta\in\Lambda_{\alpha_1,\cdots,\alpha_n}} J_\beta, \tag{3.11}$$

where the collection $\Lambda_{\alpha_1,\cdots,\alpha_n}$ can be defined recursively by $\Lambda_{\alpha_1,\cdots,\alpha_n}=\{\Lambda_{\beta,\alpha_n}|\beta\in\Lambda_{\alpha_1,\cdots,\alpha_{n-1}}\}$, $n\geq 3$. For example, $\Lambda_{(1),(0),(0)}=\{\Lambda_{\beta,(0)}|\beta\in\Lambda_{(1),(0)}\}=\{\Lambda_{(0,1),(0)},\Lambda_{(1,0),(0)}\}=\{(0,0,1),(0,0,1),(0,1,0),(1,0,0),(0,1,0),(1,0,0)\}$. Thus $J_{(1)}J_{(0)}^2=2J_{(0,0,1)}+2J_{(1,0,0)}+2J_{0,1,0)}$.

In addition to the recurrence relation (3.5), we also propose an explicit way to calculate the collection $\Lambda_{\alpha,\alpha'}$. First, for any multi-index $\alpha=(j_1,j_2,\cdots,j_l)$ with no duplicated elements (i.e., $j_m\neq j_n$ if $m\neq n$, $m,n=1,\cdots,l$), we define the set $R(\alpha)$ to be the empty set $R(\alpha)=\Phi$ if $l=1$ and $R(\alpha)=\{(j_m,j_n)|m<n,m,n=1,\cdots,l\}$ if $l\geq 2$. $R(\alpha)$ defines a partial order on the set formed with the numbers included in the multi-index α, defined by $i\prec j$ if and only if $(i,j)\in R(\alpha)$. We suppose that there are no duplicated elements in or between the multi-indexes $\alpha=(j_1,j_2,\cdots,j_l)$ and $\alpha'=(j'_1,j'_2,\cdots,j'_{l'})$.

Lemma 3.1. *If there is no duplicated elements in or between the multi-indexes* $\alpha=(j_1,j_2,\cdots,j_l)$ *and* $\alpha'=(j'_1,j'_2,\cdots,j'_{l'})$, *then*

$$\begin{aligned}\Lambda_{\alpha,\alpha'}=\{\beta\in\mathcal{M}|\,&l(\beta)=l(\alpha)+l(\alpha'),R(\alpha)\cup R(\alpha')\subseteq R(\beta)\\ &\textit{and } \beta \textit{ has no duplicated elements}\},\end{aligned} \tag{3.12}$$

where $\mathcal{M}=\{(\hat{j}_1,\hat{j}_2,\cdots,\hat{j}_{l+l'})|\hat{j}_i\in\{j_1,j_2,\cdots,j_l,j'_1,j'_2,\cdots,j'_{l'}\},\ i=1,\cdots,l+l'\}$.

Proof. Let denote $\Lambda'_{\alpha,\alpha'}=\{\beta\in\mathcal{M}|l(\beta)=l(\alpha)+l(\alpha'),\ R(\alpha)\cup R(\alpha')\subseteq R(\beta)$ and β has no duplicated elements $\}$. Since there are no duplicated elements in β and $l(\beta)=l(\alpha)+l(\alpha')$, each element of $\{j_1,j_2,\cdots,j_l,j'_1,j'_2,\cdots,j'_{l'}\}$ must appear in β only once.

We prove that $\Lambda_{\alpha,\alpha'}=\Lambda'_{\alpha,\alpha'}$ by induction on $l(\alpha)+l(\alpha')$. If $l(\alpha)+l(\alpha')=2$, then $l(\alpha)=l(\alpha')=1$ and $R(\alpha)=R(\alpha')=\Phi$. Hence $R(\beta)$ contains any pair with distinct components from $\mathcal{M}=\{(\hat{j}_1,\hat{j}_2)|\hat{j}_1,\hat{j}_2\in\{j_1,j'_1\}\}$, so $\Lambda'_{\alpha,\alpha'}=\{(j_1,j'_1),(j'_1,j_1)\}$, and from the first equation in the recurrence (3.5), $\Lambda'_{\alpha,\alpha'}=\Lambda_{\alpha,\alpha'}$.

We suppose that $\Lambda_{\alpha,\alpha'}=\Lambda'_{\alpha,\alpha'}$ for any multi-indexes α and α' such that $l(\alpha)+l(\alpha')<k$ and we prove that $\Lambda_{\alpha,\alpha'}=\Lambda'_{\alpha,\alpha'}$ for any multi-indexes α and α' with $l(\alpha)+l(\alpha')=k$.

First, we now prove that $\Lambda'_{\alpha,\alpha'}\subseteq\Lambda_{\alpha,\alpha'}$. For any element $\beta=\{\hat{j}_1,\hat{j}_2,\cdots,\hat{j}_k\}$ in $\Lambda'_{\alpha,\alpha'}$, because j_l is the largest element with respect to the partial order defined by $R(\alpha)$, and $j'_{l'}$ is the largest element with respect to the partial order defined by $R(\alpha')$, then $\hat{j}_k$ can only be j_l or $j'_{l'}$. This leads to the following cases:

1. If $\hat{j}_k = j_l$ and $l=1$, then $\beta = \alpha' * j_1 \in \Lambda_{\alpha,\alpha'}$, by the second equation in recurrence (3.5).
2. If $\hat{j}_k = j_l$ and $l>1$, then $\beta- \in \Lambda'_{\alpha-,\alpha'} = \Lambda_{\alpha-,\alpha'}$ by the induction assumption because $l(\alpha-)+l(\alpha')=k-1<k$. Hence $\beta=\beta-*(j_l)\in\Lambda_{\alpha-,\alpha'}*(j_l)$, and from the fourth equation in recurrence (3.5) we get $\beta\in\Lambda_{\alpha,\alpha'}$.
3. If $\hat{j}_k = j_{l'}$ and $l'=1$, then $\beta=\alpha*j'_1\in\Lambda_{\alpha,\alpha'}$, by the third equation in recurrence (3.5).
4. If $\hat{j}_k = j_{l'}$ and $l'>1$, then $\beta-\in\Lambda'_{\alpha,\alpha'-}=\Lambda_{\alpha,\alpha'-}$. Hence $\beta=\beta-*(j'_l)\in\Lambda_{\alpha,\alpha'-}*(j'_l)$, and from the fourth equation in recurrence (3.5) we get $\beta\in\Lambda_{\alpha,\alpha'}$.

Similarly, using the recurrence (3.5), we can prove that $\Lambda_{\alpha,\alpha'}\subseteq\Lambda'_{\alpha,\alpha'}$. Thus $\Lambda_{\alpha,\alpha'}=\Lambda'_{\alpha,\alpha'}$ and the lemma is proved. ☐

The lemma can be easily extended to determine the collection $\Lambda_{\alpha_1,\cdots,\alpha_n}$.

Lemma 3.2. *If there are no duplicated elements in or between any of the multi-indexes* $\alpha=(j_1^{(1)},j_2^{(1)},\cdots,j_{l_1}^{(1)})$, $\cdots$, $\alpha_n=(j_1^{(n)},j_2^{(n)},\cdots,j_{l_n}^{(n)})$, *then*

$$\Lambda_{\alpha_1,\cdots,\alpha_n}=\Big\{\beta\in\mathcal{M}\,|\,l(\beta)=\sum_{k=1}^{n} l(\alpha_k)\text{ and }\cup_{k=1}^{n}R(\alpha_k)\subseteq R(\beta)$$
$$\text{and there are no duplicated elements in }\beta\Big\},\tag{3.13}$$

where $\mathcal{M}=\{(\hat{j}_1,\hat{j}_2,\cdots,\hat{j}_{\hat{l}})\,|\,\hat{j}_i\in\{j_1^{(1)},j_2^{(1)},\cdots,j_{l_1}^{(1)},\cdots,j_1^{(n)},j_2^{(n)},\cdots,j_{l_n}^{(n)}\},\ i=1,\cdots,\hat{l},\hat{l}=l_1+\cdots+l_n\}$.

To extend the previous two lemmas to multi-indexes with duplicated elements, we just need to assign a different subscript to each duplicated element, for example, $\Lambda_{(2,0),(0,1)}=\Lambda_{(2,0_1),(0_2,1)}=\{(2,0_2,1,0_1),(0_2,2,1,0_1),(0_1,1,2,0_2),(0_2,2,0_1,1),(2,0_1,0_2,1),(2,0_2,0_1,1)\}$.

3.2 Higher order symplectic scheme

Inserting (3.1) into the HJ PDE (2.17a), and using the proposition (3.1), we get

$$\begin{aligned}
S_\omega^1&=\int_0^t H^{(0)}\Big(P,q+\sum_\alpha\frac{\partial G_\alpha^1}{\partial P}J_\alpha\Big)ds+\sum_{r=1}^m\int_0^t H^{(r)}\Big(P,q+\sum_\alpha\frac{\partial G_\alpha^1}{\partial P}J_\alpha\Big)\circ dw_s^r\\
&=\sum_{r=0}^m\int_0^t H^{(r)}\Big(P,q+\sum_\alpha\frac{\partial G_\alpha^1}{\partial P}J_\alpha\Big)\circ dw_s^r\\
&=\sum_{r=0}^m\int_0^t\sum_{i=0}^\infty\frac{1}{i!}\sum_{k_1,\cdots,k_i=1}^n\frac{\partial^i H^{(r)}}{\partial q_{k_1}\cdots\partial q_{k_i}}\Big(\sum_\alpha\frac{\partial G_\alpha^1}{\partial P}J_\alpha\Big)_{k_1}\cdots\Big(\sum_\alpha\frac{\partial G_\alpha^1}{\partial P}J_\alpha\Big)_{k_i}\circ dw_s^r\\
&=\sum_{r=0}^m\sum_{i=0}^\infty\frac{1}{i!}\sum_{k_1,\cdots,k_i=1}^n\frac{\partial^i H^{(r)}}{\partial q_{k_1}\cdots\partial q_{k_i}}\sum_{\alpha_1,\cdots,\alpha_i}\frac{\partial G_{\alpha_1}^1}{\partial P_{k_1}}\cdots\frac{\partial G_{\alpha_i}^1}{\partial P_{k_i}}\int_0^t\prod_{k=1}^i J_{\alpha_k}\circ dw_s^r
\end{aligned}$$

$$=\sum_{r=0}^{m}\sum_{i=0}^{\infty}\frac{1}{i!}\sum_{k_1,\cdots,k_i=1}^{n}\frac{\partial^i H^{(r)}}{\partial q_{k_1}\cdots\partial q_{k_i}}\sum_{\alpha_1,\cdots,\alpha_i}\frac{\partial G^1_{\alpha_1}}{\partial P_{k_1}}\cdots\frac{\partial G^1_{\alpha_i}}{\partial P_{k_i}}\sum_{\beta\in\Lambda_{\alpha_1,\cdots\alpha_i}}J_{\beta*(r)}$$

$$=\sum_{r=0}^{m}\sum_{i=0}^{\infty}\sum_{k_1,\cdots,k_i=1}^{n}\sum_{\alpha_1,\cdots,\alpha_i}\sum_{\beta\in\Lambda_{\alpha_1,\cdots\alpha_i}}\frac{1}{i!}\frac{\partial^i H^{(r)}}{\partial q_{k_1}\cdots\partial q_{k_i}}\frac{\partial G^1_{\alpha_1}}{\partial P_{k_1}}\cdots\frac{\partial G^1_{\alpha_i}}{\partial P_{k_i}}J_{\beta*(r)}, \tag{3.14}$$

where $(\sum_\alpha \partial G^1_{\alpha_i}/\partial P)_{k_i}$ is the k_i-th component of the column vector $\sum_\alpha \partial G^1_{\alpha_i}/\partial P$. Equating the coefficients of J_α in (3.1) and (3.14), we get the recurrence formula for determining G^1_α.

For instance, for the SHS (1.1) with $m=1$, we have

$$G^1_{(0)}=H^{(0)},\quad G^1_{(1)}=H^{(1)}. \tag{3.15}$$

To find $G^1_{(0,0)}$, since $l((0,0))$ =2 we only need to consider the values $i=1$, $\alpha=(0)$ and $r=0$, so that

$$G^1_{(0,0)}=\sum_{k=1}^{n}\frac{\partial H^{(0)}}{\partial q_k}\frac{\partial G^1_{(0)}}{\partial P_k}=\sum_{k=1}^{n}\frac{\partial H^{(0)}}{\partial q_k}\frac{\partial H^{(0)}}{\partial P_k}. \tag{3.16}$$

Similarly, using $i=1$, $\alpha=(1)$ and $r=0$ for $G^1_{(1,0)}$, $i=1$, $\alpha=(0)$ and $r=1$ for $G^1_{(0,1)}$, and $i=1$, $\alpha=(1)$ and $r=1$ for $G^1_{(1,1)}$, we obtain

$$G^1_{(1,1)}=\sum_{k=1}^{n}\frac{\partial H^{(1)}}{\partial q_k}\frac{\partial H^{(1)}}{\partial P_k},\quad G^1_{(1,0)}=\sum_{k=1}^{n}\frac{\partial H^{(0)}}{\partial q_k}\frac{\partial H^{(1)}}{\partial P_k},\quad G^1_{(0,1)}=\sum_{k=1}^{n}\frac{\partial H^{(1)}}{\partial q_k}\frac{\partial H^{(0)}}{\partial P_k}. \tag{3.17}$$

Because $l((0,0,0))=3$, the cases $i=1$, $\alpha=(0,0)$, $r=0$ and $i=2$, $\alpha_1=(0)$, $\alpha_2=(0)$, $r=0$ both contribute to the coefficient of $J_{(0,0,0)}$:

$$\begin{aligned}&G^1_{(0,0,0)}\\&=\sum_{k_1=1}^{n}\frac{\partial H^{(0)}}{\partial q_{k_1}}\frac{\partial G^1_{(0,0)}}{\partial P_{k_1}}+\sum_{k_1,k_2=1}^{n}\frac{1}{2}\frac{\partial^2 H^{(0)}}{\partial q_{k_1}\partial q_{k_2}}2\frac{\partial G^1_{(0)}}{\partial P_{k_1}}\frac{\partial G^1_{(0)}}{\partial P_{k_2}}\\&=\sum_{k_1,k_2=1}^{n}\Big(\frac{\partial^2 H^{(0)}}{\partial q_{k_1}\partial q_{k_2}}\frac{\partial H^{(0)}}{\partial P_{k_1}}\frac{\partial H^{(0)}}{\partial P_{k_2}}+\frac{\partial H^{(0)}}{\partial q_{k_1}}\frac{\partial H^{(0)}}{\partial P_{k_2}}\frac{\partial^2 H^{(0)}}{\partial q_{k_2}\partial P_{k_1}}+\frac{\partial H^{(0)}}{\partial q_{k_1}}\frac{\partial H^{(0)}}{\partial q_{k_2}}\frac{\partial^2 H^{(0)}}{\partial P_{k_1}\partial P_{k_2}}\Big).\end{aligned} \tag{3.18}$$

Similarly,

$$G^1_{(1,1,1)}=\sum_{k_1,k_2=1}^{n}\Big(\frac{\partial^2 H^{(1)}}{\partial q_{k_1}\partial q_{k_2}}\frac{\partial H^{(1)}}{\partial P_{k_1}}\frac{\partial H^{(1)}}{\partial P_{k_2}}+\frac{\partial H^{(1)}}{\partial q_{k_1}}\frac{\partial H^{(1)}}{\partial P_{k_2}}\frac{\partial^2 H^{(1)}}{\partial q_{k_2}\partial P_{k_1}}+\frac{\partial H^{(1)}}{\partial q_{k_1}}\frac{\partial H^{(1)}}{\partial q_{k_2}}\frac{\partial^2 H^{(1)}}{\partial P_{k_1}\partial P_{k_2}}\Big), \tag{3.19a}$$

$$G^1_{(1,1,0)}=\sum_{k_1,k_2=1}^{n}\Big(\frac{\partial^2 H^{(0)}}{\partial q_{k_1}\partial q_{k_2}}\frac{\partial H^{(1)}}{\partial P_{k_1}}\frac{\partial H^{(1)}}{\partial P_{k_2}}+\frac{\partial H^{(0)}}{\partial q_{k_1}}\frac{\partial H^{(1)}}{\partial P_{k_2}}\frac{\partial^2 H^{(1)}}{\partial q_{k_2}\partial P_{k_1}}+\frac{\partial H^{(0)}}{\partial q_{k_1}}\frac{\partial H^{(1)}}{\partial q_{k_2}}\frac{\partial^2 H^{(1)}}{\partial P_{k_1}\partial P_{k_2}}\Big), \tag{3.19b}$$

$$G^1_{(1,0,1)}=\sum_{k_1,k_2=1}^{n}\Big(\frac{\partial^2 H^{(1)}}{\partial q_{k_1}\partial q_{k_2}}\frac{\partial H^{(0)}}{\partial P_{k_1}}\frac{\partial H^{(1)}}{\partial P_{k_2}}+\frac{\partial H^{(1)}}{\partial q_{k_1}}\frac{\partial H^{(1)}}{\partial P_{k_2}}\frac{\partial^2 H^{(0)}}{\partial q_{k_2}\partial P_{k_1}}+\frac{\partial H^{(1)}}{\partial q_{k_1}}\frac{\partial H^{(0)}}{\partial q_{k_2}}\frac{\partial^2 H^{(1)}}{\partial P_{k_1}\partial P_{k_2}}\Big), \tag{3.19c}$$

$$G^1_{(0,1,1)}=\sum_{k_1,k_2=1}^{n}\Big(\frac{\partial^2 H^{(1)}}{\partial q_{k_1}\partial q_{k_2}}\frac{\partial H^{(1)}}{\partial P_{k_1}}\frac{\partial H^{(0)}}{\partial P_{k_2}}+\frac{\partial H^{(1)}}{\partial q_{k_1}}\frac{\partial H^{(0)}}{\partial P_{k_2}}\frac{\partial^2 H^{(1)}}{\partial q_{k_2}\partial P_{k_1}}+\frac{\partial H^{(1)}}{\partial q_{k_1}}\frac{\partial H^{(1)}}{\partial q_{k_2}}\frac{\partial^2 H^{(0)}}{\partial P_{k_1}\partial P_{k_2}}\Big). \tag{3.19d}$$

For $m\geq 1$, we apply Lemma 3.2 to obtain a recurrence formula for G^1_α. If $\alpha=(r)$, $r=1,\cdots,m$ then $G^1_\alpha=H^{(r)}$. If $\alpha=(i_1,\cdots,i_{l-1},r)$, $l>1$, $i_1,\cdots,i_{l-1}$, $r=1,\cdots,m$ has no duplicates then

$$G^1_\alpha=\sum_{i=1}^{l(\alpha)-1}\frac{1}{i!}\sum_{k_1,\cdots,k_i=1}^{n}\frac{\partial^i H^{(r)}}{\partial q_{k_1}\cdots\partial q_{k_i}}\sum_{\substack{l(\alpha_1)+\cdots+l(\alpha_i)=l(\alpha)-1\\ R(\alpha_1)\cup\cdots\cup R(\alpha_i)\subseteq R(\alpha-)}}\frac{\partial G^1_{\alpha_1}}{\partial P_{k_1}}\cdots\frac{\partial G^1_{\alpha_i}}{\partial P_{k_i}}. \tag{3.20}$$

If the multi-index α contains any duplicates, then we apply formula (3.20) after associating different subscripts to the repeating numbers.

We can use the same approach, for the HJPDE (2.17b). For example, for the SHS (1.1) with $m=1$, for S^3_ω we get

$$G^3_{(0)}=H^{(0)},\qquad G^3_{(1)}=H^{(1)},\qquad G^3_{(0,0)}=0,\qquad G^3_{(1,1)}=0, \tag{3.21a}$$

$$G^3_{(1,0)}=\frac{1}{2}(\nabla H^{(0)})^TJ^{-1}\nabla H^{(1)},\qquad G^3_{(0,1)}=\frac{1}{2}(\nabla H^{(1)})^TJ^{-1}\nabla H^{(0)}, \tag{3.21b}$$

$$G^3_{(0,0,0)}=\frac{1}{4}(J^{-1}\nabla H^{(0)})^T\nabla^2H^{(0)}(J^{-1}\nabla H^{(0)}),\cdots. \tag{3.21c}$$

Using (2.7) and a truncated series for S^1_ω, or using (2.9) and a truncated series for S^3_ω, we obtain various symplectic schemes for the SHS (1.1). In this paper we study only the strong schemes, but a similar approach can be applied to construct the weak schemes, and it will be reported in a different paper [18]. Let define $\mathcal{A}_\gamma=\{\alpha: l(\alpha)+n(\alpha)\leq 2\gamma\}$ and $\mathcal{B}_\gamma=\{\alpha: l(\alpha)+n(\alpha)\leq 2\gamma \text{ or } l(\alpha)=n(\alpha)=\gamma+0.5\}$, where $n(\alpha)$ is the number of zero components of the multi-index α (e.g., $n((0,0,1))=2$).

The implicit midpoint scheme in [6] is the numerical scheme of order 1 obtained from (2.9) using the truncated series $S^3_\omega\approx\sum_{\alpha\in\mathcal{A}_1}G^3_\alpha J_\alpha=\sum_{r=1}^m G^3_{(r)}J_{(r)}$ (see also Eq. (2.11) where bounded random variables are used to approximate $J_{(r)}$ because the scheme is implicit). A first order symplectic implicit scheme is also obtained if we truncate the Stratonovich expansion for S^1_ω according to $\mathcal{A}_1$:

$$S^1_\omega\approx G^1_{(0)}J_{(0)}+\sum_{r=1}^m\left(G^1_{(r)}J_{(r)}+G^1_{(r,r)}J_{(r,r)}\right)+\sum_{i,j=1,i\neq j}^m G^1_{(i,j)}J_{(i,j)}. \tag{3.22}$$

In the next section, we study the convergence and we prove the first order mean square convergence for the scheme based on the generating function given in (3.22).

To obtain the symplectic Euler scheme of order 0.5 in [8], we use the relation of the Ito stochastic multiple integrals and the Stratonovich stochastic multiple integrals (see [19]), and we replace in the expansion (3.1) of S^1_ω 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 $\alpha\in\mathcal{B}_{0.5}$, and for $m=1$ we have

$$S^1_\omega\approx G^1_{(1)}I_{(1)}+\left(G^1_{(0)}+\frac{1}{2}G^1_{(1,1)}\right)I_{(0)}. \tag{3.23}$$

We notice that the generating function in (2.10) was obtained from the previous equation, using (3.15)-(3.16) and bounded random variables to approximate $I_{(1)}$.

For the 1.5 order scheme, we truncate according to $\mathcal{B}_{1.5}$, so, for $m=1$ we get :

$$\begin{aligned} S^1_\omega \approx & G^1_{(1)} I_{(1)} + \Big(G^1_{(0)} + \frac{1}{2} G^1_{(1,1)}\Big) I_{(0)} + \Big(G^1_{(0,1)} + \frac{1}{2} G^1_{(1,1,1)}\Big) I_{(0,1)} + \Big(G^1_{(1,0)} + \frac{1}{2} G^1_{(1,1,1)}\Big) \\ & I_{(1,0)} + G^1_{(1,1)} I_{(1,1)} + \Big(G^1_{(0,0)} + \frac{1}{2}(G^1_{(0,1,1)} + G^1_{(1,1,0)}) + \frac{1}{4} G^1_{(1,1,1,1)}\Big) I_{(0,0)}. \end{aligned} \tag{3.24}$$

The formulas for the coefficients G^1_α included in (3.24) are given in (3.15)-(3.19), and the Ito integrals $I_{(0,1)}$, $I_{(1,0)}$, and $I_{(1,1)}$ should be approximated using bounded random variables (see [8, 19]).

Remark 3.2. If we consider the deterministic cases, i.e., $m=0$, then $J_\alpha = t^n/n!$ with $l(\alpha)=n$. The coefficients (3.16)-(3.18) and (3.21) of the approximations of the generating function proposed in this paper, are consistent with those of Type (II) and Type (III) generating functions in [17]. In other words, the proposed construction of the stochastic symplectic numerical schemes via generating function is an extension of the methods introduced by Kang Feng [17].

4 Convergence analysis

In this section, we study the convergence of the first order symplectic implicit scheme constructed using the generating function given in (3.22). As we have mentioned early, since this is an implicit scheme we need to use bounded random variables. To keep the notation simple, we consider the SHS (1.1) with $n=1$ and $m=1$, but the same approach can be easily extended to the general case. Also for notation convenience, $\partial H/\partial p$ and $\partial H/\partial q$ are denoted as H_p and H_q, respectively.

As in [8], 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:

$$\xi_h = \begin{cases} -A_h, & \text{if } \xi < -A_h, \\ \xi, & \text{if } |\xi| \le A_h, \\ A_h, & \text{if } \xi > A_h, \end{cases} \tag{4.1}$$

where $A_h = 2\sqrt{|\ln h|}$. From [8], we know that

$$E(\xi - \xi_h)^2 \le h^2, \tag{4.2a}$$

$$0 \le E(\xi^2 - \xi_h^2) \le \Big(1 + 4\sqrt{|\ln h|}\Big) h^2 \le 5h^{3/2}. \tag{4.2b}$$

Carrying out similar calculations, we get

$$E(\zeta^2-\zeta_h^2)^2=\frac{2}{\sqrt{2\pi}}\int_{A_h}^{\infty}(x^2-A_h^2)^2e^{-x^2/2}dx=\frac{2}{\sqrt{2\pi}}\int_0^{\infty}(y^2+2A_hy)^2, \tag{4.3a}$$

$$e^{-\frac{(y+A_h)^2}{2}}dy\le\frac{2e^{-\frac{A_h^2}{2}}}{\sqrt{2\pi}}\int_0^{\infty}(y^2+2A_hy)^2e^{-\frac{y^2}{2}}dy=e^{-\frac{A_h^2}{2}}\Big(3+4A_h^2+\frac{8A_h}{\sqrt{2\pi}}\Big)\le 27h. \tag{4.3b}$$

From (4.1), for any non-negative integer k, we can easily verify

$$E(\zeta_h^{2k+1})=E(\zeta^{2k+1})=0,\quad E(|\zeta_h|^k)\le E(|\zeta|^k)<\infty. \tag{4.4}$$

Using (2.7) and (3.22), for the SHS (1.1) with $n=1$ and $m=1$, we construct an implicit symplectic scheme corresponding to the following one step approximation:

$$P=p-(H_q^{(0)}(P,q)J_{(0)}+H_q^{(1)}(P,q)J_{(1)}^h+(H_p^{(1)}(P,q)H_q^{(1)}(P,q))_qJ_{(1,1)}^h), \tag{4.5a}$$

$$Q=q+(H_p^{(0)}(P,q)J_{(0)}+H_p^{(1)}(P,q)J_{(1)}^h+(H_p^{(1)}(P,q)H_q^{(1)}(P,q))_pJ_{(1,1)}^h), \tag{4.5b}$$

where $J_{(0)}=h$, $J_{(1)}^h=\sqrt{h}\zeta_h$ and $J_{(1,1)}^h=\zeta_h^2/2$.

We suppose that the Hamiltonian functions $H^{(0)}$ and $H^{(1)}$ and their partial derivatives up to order four are continuous, and the following inequalities hold for some positive constants L_i, $i=1,\cdots,5$,

$$\sum_{r=0}^{1}(|H_p^{(r)}(P,Q)-H_p^{(r)}(p,q)|+|H_q^{(r)}(P,Q)-H_q^{(r)}(p,q)|)\le L_1(|P-p|+|Q-q|), \tag{4.6a}$$

$$\sum_{r=0}^{1}(|H_p^{(r)}(p,q)|+|H_q^{(r)}(p,q)|)\le L_2(1+|p|+|q|), \tag{4.6b}$$

$$\sum_{r=0}^{1}(|H_{pp}^{(r)}(p,q)|+|H_{pq}^{(r)}(p,q)|+|H_{qq}^{(r)}(p,q)|)\le L_3, \tag{4.6c}$$

$$\sum_{r=0}^{1}(|H_{ppq}^{(r)}(p,q)|+|H_{pqq}^{(r)}(p,q)|+|H_{ppp}^{(r)}(p,q)|)\le\frac{L_4}{1+|p|+|q|}, \tag{4.6d}$$

$$|H_{pppq}^{(1)}(p,q)|+|H_{ppqq}^{(1)}(p,q)|+|H_{pppp}^{(1)}(p,q)|\le\frac{L_5}{(1+|p|+|q|)^2}, \tag{4.6e}$$

$$\begin{aligned}(|(H_p^{(1)}H_q^{(1)})_p(P,Q)-(H_p^{(1)}H_q^{(1)})_p(p,q)|+|(H_p^{(1)}H_q^{(1)})_q(P,Q)\\ -(H_p^{(1)}H_q^{(1)})_q(p,q)|)\le L_1(|P-p|+|Q-q|).\end{aligned} \tag{4.6f}$$

The first equation in (4.5) is implicit, so in the following lemma we show that the scheme (4.5) is well-defined.

Lemma 4.1. *There exists constants $K_0>0$ and $h_0>0$, such that for any $h<h_0$ the first equation in* (4.5) *has a unique solution P which satisfies*

$$|P-p|\le K_0(1+|p|+|q|)\Big(|\xi_h|\sqrt{h}+h+\frac{1}{2}\xi_h^2h\Big),\quad k=1,2,\cdots. \tag{4.7}$$

Proof. The proof can be completed similarly with the proof of Lemma 2.4 in [8], using the Assumptions (4.6a)-(4.6f) and the contraction principle. □

Corollary 4.1. *There exists constants $K>0$ and $h_0>0$, such that for any $h<h_0$, we have*

$$E(|P-p|^i+|Q-q|^i)\le K(1+|p|+|q|)^ih^{\frac{i}{2}},\quad i=1,2,\cdots. \tag{4.8}$$

To prove the first order mean square convergence for the scheme based on the one step approximation (4.5), we apply the following general result (Theorem 1.1 in [20]):

Theorem 4.1. *Let $\bar{X}_{t,x}(t+h)$ be a one step approximation for the solution $X_{t,x}(t+h)$ of the SHS* (1.1)*. If for arbitrary $t_0\le t\le t_0+T-h$, $x\in R^{2n}$ the following inequalities hold:*

$$\big|E(X_{t,x}(t+h)-\bar{X}_{t,x}(t+h))\big|\le K(1+|x|^2)^{1/2}h^{p_1}, \tag{4.9a}$$

$$\big[E|X_{t,x}(t+h)-\bar{X}_{t,x}(t+h)|^2\big]^{1/2}\le K(1+|x|^2)^{1/2}h^{p_2}, \tag{4.9b}$$

with $p_2\ge 1/2$ and $p_1\ge p_2+1/2$, then the mean square order of accuracy of the method constructed using the one step approximation $\bar{X}_{t,x}(t+h)$ is $p_2-1/2$.

Before proving the main convergence theorem, we include some preliminary results in the following lemma.

Lemma 4.2. *There exists constants $K_1,K_2,K_3>0$ and $h_0>0$, such that for any $h<h_0$, we have*

$$|E(P-p)|+|E(Q-q)|\le K_1(1+|p|+|q|)h, \tag{4.10a}$$

$$|E((P-p)J^h_{(11)})|+|E((Q-q)J^h_{(11)})|\le K_2(1+|p|+|q|)h^2, \tag{4.10b}$$

$$|E((P-p)^2J^h_{(1)})|+|E((Q-q)^2J^h_{(1)})|\le K_3(1+|p|+|q|)^2h^{5/2}. \tag{4.10c}$$

Proof. For $r=0,1$, $z=p$ or q, and with sufficiently small h, from (4.6b) and (4.7), we have

$$\begin{aligned}|H_z^{(r)}(P,q)|&\le|H_z^{(r)}(P,q)-H_z^{(r)}(p,q)|+|H_z^{(r)}(p,q)|\le L_1|P-p|+|H_z^{(r)}(p,q)|\\&\le(1+|p|+|q|)\Big(K|\xi_h|\sqrt{h}+Kh+K\frac{1}{2}\xi_h^2h+L_3\Big).\end{aligned} \tag{4.11}$$

Hence, using (4.4) we show that there exists constants $KC_i>0$, $i=1,2,\cdots$, such that

$$E|H_z^{(r)}(P,q)|^i\le KC_i(1+|p|+|q|)^i,\quad i=1,2,\cdots. \tag{4.12}$$

Using the Taylor expansion, we rewrite the first relation of (4.5) as $P-p=-H_q^{(1)}(p,q)J^h_{(1)}+R_1$ with $R_1=-H_q^{(0)}(P,q)J_{(0)}-H_{qp}^{(1)}(\bar{p}_1,q)(P-p)J^h_{(1)}-(H_p^{(1)}(P,q)H_q^{(1)}(P,q))_q J^h_{(1,1)}$, where $\bar{p}_1$ is between p and P. Hence, using the Cauchy-Schwarz inequality, (4.6b), (4.6c), (4.8), and (4.12) imply that there is a constant $K_1>0$, such that

$$\begin{aligned}|E(R_1)|\le E|R_1|\le & E|H_q^{(0)}(P,q)|J_{(0)}+L_3\sqrt{E|P-p|^2}\sqrt{E|J^h_{(1)}|^2}\\ &+\Big(\sqrt{E|H_q^{(1)}(P,q)|^2}+\sqrt{E|H_p^{(1)}(P,q)|^2}\Big)L_3\sqrt{E|J^h_{(11)}|^2}\\ \le&\frac{K_1}{2}(1+|p|+|q|)h. \end{aligned}\tag{4.13}$$

Moreover, since we have

$$\begin{aligned}R_1^2\le 2\Big(&(H_q^{(0)}(P,q))^2(J_{(0)})^2+(H_{qp}^{(1)}(\bar{p}_1,q))^2(P-p)^2(J^h_{(1)})^2\\ &+\big((H_p^{(1)}(P,q))^2(H_{qq}^{(1)}(P,q))^2+(H_{pq}^{(1)}(P,q))^2(H_q^{(1)}(P,q))^2\big)(J^h_{(1,1)})^2\Big),\end{aligned}\tag{4.14}$$

proceeding similarly we can show that there exists constants $K_2>0$ and $K_3'>0$, such that

$$E(R_1^2)\le\frac{1}{3}K_2^2(1+|p|+|q|)^2h^2,\qquad |E(R_1^2J^h_{(1)})|\le K_3'(1+|p|+|q|)^2h^{5/2},\tag{4.15a}$$

$$E(R_1^2(J^h_{(1)})^2)\le K_3'(1+|p|+|q|)^2h^3.\tag{4.15b}$$

Using the Cauchy-Schwarz inequality, (4.13), (4.15a) and (4.4) imply that there exists a constant $K_3>0$ such that:

$$|E(P-p)|\le\Big|H_q^{(1)}(p,q)E(J^h_{(1)})\Big|+|E(R_1)|\le\frac{K_1}{2}(1+|p|+|q|)h,\tag{4.16a}$$

$$\begin{aligned}|E((P-p)J^h_{(11)})|&\le\Big|H_q^{(1)}(p,q)E(J^h_{(1)}J^h_{(11)})\Big|+|E(R_1J^h_{(11)})|\le\sqrt{E(R_1^2)E|J^h_{(11)}|^2}\\ &\le\frac{K_2}{2}(1+|p|+|q|)h^2,\end{aligned}\tag{4.16b}$$

$$\begin{aligned}|E((P-p)^2J^h_{(1)})|&\le(H_q^{(1)}(p,q))^2\big|E((J^h_{(1)})^3)\big|+|E(R_1^2J^h_{(1)})|+2\Big|H_q^{(1)}(p,q)E(R_1(J^h_{(1)})^2)\Big|\\ &\le K_3'(1+|p|+|q|)^2h^{5/2}+2\big|H_q^{(1)}(p,q)\big|\sqrt{E(R_1^2)E|J^h_{(1)}|^4}\le\frac{K_3}{2}(1+|p|+|q|)^2h^{5/2}.\end{aligned}\tag{4.16c}$$

Similarly,

$$|E(Q-q)|\le\frac{K_1}{2}(1+|p|+|q|)h,$$

$$|E((Q-q)J^h_{(11)})|\le\frac{K_2}{2}(1+|p|+|q|)h^2,$$

and

$$|E((Q-q)^2J^h_{(1)})|\le\frac{K_3}{2}(1+|p|+|q|)^2h^{5/2},$$

so (4.10a)-(4.10c) are proved. □

Remark 4.1. Notice that for h sufficiently small, using Taylor expansions, inequalities (4.6b)-(4.6d), (4.8) and the characterization of the moments in (4.4), we can also show that there exist a constant $K_4>0$, such that

$$\begin{aligned}
|E(R_1 J^h_{(1)})| \le & \Big| E\Big((H_q^{(0)}(p,q)J_{(0)} + (H_p^{(1)}H_q^{(1)})_q(p,q)J_{(1,1)}) J^h_{(1)} \Big) \Big| \\
& + E\Big| J^h_{(1)} H^{(0)}_{pq}(\bar{p}_0,q) J_{(0)}(P-p)\Big| + E\Big| J^h_{(1)} (H_p^{(1)}H_q^{(1)})_{pq}(\bar{p}_{11},q)) J^h_{(1,1)}(P-p)\Big| \\
& + \Big| E\Big(J^h_{(1)} H^{(1)}_{pq}(\bar{p}_1,q) J^h_{(1)}(P-p)\Big)\Big| \\
\le & L_3K(1+|p|+|q|)h^2 + (3L_4L_2+L_3^2)(1+|p|+|q|)h^2 \\
& + 2|H^{(1)}_{pq}(p,q)||E(J^h_{(11)}(P-p))| + E|J^h_{(1)} H^{(1)}_{ppq}(\hat{\bar{p}}_1,q) J^h_{(1)}(\bar{p}_1-p)(P-p)| \\
\le & L_3K(1+|p|+|q|)h^2 + (3L_4L_2+L_3^2)(1+|p|+|q|)h^2 + 2K_2L_3(1+|p|+|q|)h^2 \\
& + L_4(1+|p|+|q|)h^2 \\
\le & K_4(1+|p|+|q|)h^2, \qquad (4.17)
\end{aligned}$$

where $\bar{p}_0$, $\bar{p}_1$ and $\bar{p}_{11}$ are values between P and p and $\hat{\bar{p}}_1$ is a value between $\bar{p}_1$ and p. Here we have also used $2J^h_{(11)} = (J^h_{(1)})^2$ and the fact that for sufficiently small h, Lemma 4.1 implies that there exists positive constants C_1 and C_2 such that

$$1+|\bar{p}|+|q| \le 1+|P-p|+|p|+|q| \le C_1(1+|p|+|q|), \qquad (4.18a)$$

$$\frac{1}{1+|\bar{p}|+|q|} \le \frac{1}{1+|p|-|\bar{p}-p|+|q|} \le \frac{1}{1+|p|-|P-p|+|q|} \le \frac{C_2}{1+|p|+|q|}, \qquad (4.18b)$$

for any $\bar{p}$ between P and p.

Theorem 4.2. *If the conditions* (4.6a)-(4.6d) *are satisfied, the scheme* (4.5) *converges with the mean square order 1.*

Proof. Applying Taylor expansions for the first part of the one step approximation (4.5), we have

$$\begin{aligned}
P-p = & -H_q^{(0)}(p,q)J_{(0)} - H_q^{(1)}(p,q)J^h_{(1)} - (H_p^{(1)}H_q^{(1)})_q(p,q)J^h_{(1,1)} \\
& - H^{(0)}_{pq}(p,q)J_{(0)}(P-p) - (H_p^{(1)}H_q^{(1)})_{pq}(p,q)J^h_{(1,1)}(P-p) \\
& - H^{(1)}_{pq}(p,q)(P-p)J^h_{(1)} - \frac{1}{2}H^{(0)}_{ppq}(\bar{p}_{00},q)(P-p)^2 J^h_{(0)} \\
& - \frac{1}{2}H^{(1)}_{ppq}(p,q)(P-p)^2 J^h_{(1)} - \frac{1}{2}(H_p^{(1)}H_q^{(1)})_{ppq}(\bar{p}_{011},q)(P-p)^2 J^h_{(11)} \\
& - \frac{1}{6}H^{(1)}_{pppq}(\bar{p}_{01},q)(P-p)^3 J^h_{(1)}
\end{aligned}$$

$$
\begin{aligned}
&= -H_q^{(0)}(p,q)J_{(0)} - H_q^{(1)}(p,q)J_{(1)}^h - (H_p^{(1)}(p,q)H_q^{(1)}(p,q))_q J_{(1,1)}^h \\
&\quad - H_{pq}^{(1)}(p,q)(P-p)J_{(1)}^h + R_2,
\end{aligned} \tag{4.19}
$$

where $\bar{p}_{00}$, $\bar{p}_{01}$ and $\bar{p}_{011}$ are values between P and p. Since

$$
\begin{aligned}
R_2^2 \le\, & 2(H_{pq}^{(0)}(p,q))^2(J_{(0)})^2(P-p)^2 + 2\big((H_p^{(1)}H_q^{(1)})_{pq}(p,q)\big)^2(J_{(1,1)}^h)^2(P-p)^2 \\
&+ \frac{1}{2}\big(H_{ppq}^{(0)}(\bar{p}_{00},q)\big)^2(P-p)^4(J_{(0)}^h)^2 + \frac{1}{2}\big(H_{ppq}^{(1)}(p,q)\big)^2(P-p)^4(J_{(1)}^h)^2 \\
&+ \frac{1}{2}\big((H_p^{(1)}H_q^{(1)})_{ppq}(\bar{p}_{011},q)\big)^2(P-p)^4(J_{(11)}^h)^2 \\
&+ \frac{1}{18}\big(H_{pppq}^{(1)}(\bar{p}_{01},q)\big)^2(P-p)^6(J_{(1)}^h)^2,
\end{aligned} \tag{4.20}
$$

the Assumptions (4.6b)-(4.6e), Cauchy-Schwarz inequality, inequalities (4.8), Lemma 4.2, and the characterization of the moments in (4.4) implies that for h sufficiently small there exists a positive constant K_6, such that

$$
|E(R_2)| \le K_6(1+|p|+|q|)h^2, \quad E|R_2|^2 \le K_6(1+|p|+|q|)^2h^3. \tag{4.21}
$$

Substituting $P-p = H_q^1(p,q)J_{(1)}^h + R_1$, where R_1 is defined in the proof of Lemma 4.2, into $H_{pq}^{(1)}(p,q)(P-p)J_{(1)}^h$, we obtain

$$
\begin{aligned}
P-p =& -H_q^{(0)}(p,q)J_{(0)} - H_q^{(1)}(p,q)J_{(1)}^h - (H_p^{(1)}(p,q)H_q^{(1)}(p,q))_q J_{(1,1)}^h \\
&- H_{pq}^{(1)}(p,q)H_q^1(p,q)(J_{(1)}^h)^2 - H_{pq}^{(1)}(p,q)J_{(1)}^h R_1 + R_2.
\end{aligned} \tag{4.22}
$$

It is easy to verify that assumption (4.6b) and inequalities (4.15b), (4.17) and (4.21) imply that there exists a positive constant K_7, such that

$$
\big|E(H_{pq}^{(1)}(p,q)J_{(1)}^h R_1 - R_2)\big| \le K_7(1+|p|+|q|)h^2, \tag{4.23a}
$$

$$
E\big(H_{pq}^{(1)}(p,q)J_{(1)}^h R_1 - R_2\big)^2 \le K_7(1+|p|+|q|)^2h^3. \tag{4.23b}
$$

Recall that the Milstein scheme [19] for the stochastic Hamiltonian system (1.1) satisfying conditions (4.6a)-(4.6e) has the mean square order 1 and satisfies the inequalities (4.9a)-(4.9b) with $p_1 = 2$, $p_2 = 1.5$. The one step approximation corresponding to the Milstein scheme is given by

$$
\begin{aligned}
\tilde{P} =\,& p - H_q^{(0)}(p,q)J_{(0)} - H_q^{(1)}(p,q)J_{(1)} \\
&+ (H_{pq}^{(1)}(p,q)H_q^{(1)}(p,q) - H_{qq}^{(1)}(p,q)H_p^{(1)}(p,q))J_{(1,1)},
\end{aligned} \tag{4.24a}
$$

$$
\begin{aligned}
\tilde{Q} =\,& q + H_p^{(0)}(p,q)J_{(0)} + H_p^{(1)}(p,q)J_{(1)} \\
&+ (H_{pq}^{(1)}(p,q)H_p^{(1)}(p,q) - H_{pp}^{(1)}(p,q)H_q^{(1)}(p,q))J_{(1,1)}.
\end{aligned} \tag{4.24b}
$$

Comparing the one step approximation corresponding to the Milstein scheme with (4.5), we obtain

$$\begin{aligned} P-\tilde{P}=&H_q^{(1)}(p,q)(J_{(1)}-J_{(1)}^h)+(H_{pq}^{(1)}(p,q)H_q^{(1)}(p,q)-H_{qq}^{(1)}(p,q)H_p^{(1)}(p,q))(J_{(1,1)}-J_{(1,1)}^h)\\ &-H_{pq}^{(1)}(p,q)J_{(1)}^h R_1+R_2. \end{aligned} \tag{4.25}$$

Thus, from (4.2a)-(4.2a), assumptions (4.6b), (4.6c) and (4.23a), (4.23b), we get

$$E(P-\tilde{P})^2 \le (1+|p|+|q|)^2 h^3\big(2L_2^2+54L_3^2L_2^2+K_7\big), \tag{4.26a}$$

$$|E(P-\tilde{P})| \le (1+|p|+|q|)h^2\big(L_3L_2h^{\frac{1}{2}}+K_7\big). \tag{4.26b}$$

The proof for the $Q-\tilde{Q}$ follows similarly by repeating the same procedure for the second relation of (4.5), so the scheme corresponding to the one step approximation (4.5) satisfies the inequalities (4.9a)-(4.9b) with $p_1=2$, $p_2=1.5$. □

Remark 4.2. Using the same approach, we were able to prove that the symplectic schemes based on truncations of S_ω^1 or S_ω^3 for multi-indexes $\alpha\in\mathcal{B}_k$ or $\alpha\in\mathcal{A}_k$ have mean square order k, for $k=1,1.5,2$. Higher order schemes include Ito multiple stochastic integrals I_α with multi-indexes $\alpha\in\mathcal{B}_k$ or Stratonovich multiple stochastic integrals J_α with multi-indexes $\alpha\in\mathcal{A}_k$, for which is computationally expensive to simulate bounded approximations when $k>2$.

5 Symplectic schemes for special types of stochastic Hamiltonian systems

As mentioned in Section 3.2, the idea of construction of stochastic symplectic schemes of mean square order k is to substitute the truncation of S_ω^1 or S_ω^3 with multi-indexes $\alpha\in\mathcal{B}_k$ for $k=1,2,\cdots$ or $\alpha\in\mathcal{A}_k$ for $k=0.5,1.5,\cdots$, into (2.7) or (2.9).

Take S_ω^1 with $k=1.5$ on the SHS with one noise, $m=1$, for instances, the truncation of S_ω^1 with $k=1.5$ is presented as

$$\begin{aligned} S_\omega^1=&G_{(0)}^1I_{(0)}+G_{(1)}^1\bar{I}_{(1)}+G_{(1,1)}^1I_{(1,1)}+G_{(1,1,1)}^1\bar{I}_{(1,1,1)}+G_{(1,0)}^1\bar{I}_{(1,0)}+G_{(0,1)}^1\bar{I}_{(0,1)}\\ &+\left(G_{(0,0)}^1+\frac{G_{(0,1,1)}^1}{2}+\frac{G_{(1,1,0)}^1}{2}+\frac{G_{(1,1,1,1)}^1}{4}\right)\bar{I}_{(0,0)}, \end{aligned} \tag{5.1}$$

where

$$I_{(0)}=h,\qquad \bar{I}_{(1)}=\sqrt{h}\xi_h,\qquad I_{(1,1)}=\frac{h\xi_h^2}{2},$$

$$\bar{I}_{(1,1,1)}=\frac{(\sqrt{h}\xi_h)^3}{6},\quad \bar{I}_{(1,0)}^h=\frac{h^{\frac{3}{2}}}{2}\Big(\xi_h+\frac{\eta_h}{\sqrt{3}}\Big),\quad \bar{I}_{(0,1)}=\xi_h h^{\frac{3}{2}}-\bar{I}_{(1,0)},$$

and

$$I_{(0,0)}=\frac{h^2}{2}.$$

Here ξ_h and η_h are independent bounded Gaussian random variables given by (4.1). It is noticed that the transformation of multiple Ito integrals and multiple Stratonovich integrals,

$$I_{(0,0)}=J_{(0,0)}+\frac{1}{2}J_{(0,1,1)}+\frac{1}{2}J_{(1,1,0)}+\frac{1}{4}J_{(1,1,1,1)},$$

are used (see [19] and [18] for more details). Then the stochastic S_ω^1 scheme of mean square order 1.5 is obtained by substituting (5.1) into (2.7).

In this framework, the construction of the stochastic symplectic schemes is reduced to the calculation of G_α^1 or G_α^3. Three special types of SHSs will be considered in the following subsection.

5.1 SHS with additive noise

First, we consider the special case of SHS with additive noise

$$dP_i=-\frac{\partial H^{(0)}(P,Q)}{\partial Q_i}dt-\sum_{r=1}^m\sigma_r\circ dw_t^r,\qquad P(t_0)=p, \tag{5.2a}$$

$$dQ_i=\frac{\partial H^{(0)}(P,Q)}{\partial P_i}dt+\sum_{r=1}^m\tau_r\circ dw_t^r,\qquad Q(t_0)=q, \tag{5.2b}$$

where $i=1,\cdots,n$. Notice that $H^{(r)}=\sum_{i=1}^n(P_i\tau_r+Q_i\sigma_r)$, where σ_r and τ_r are constants.

To calculate the coefficients of S_ω^1, we replace in (3.20) and get,

$$G^1_{(0,0)}=\sum_{k=1}^n\frac{\partial H^{(0)}}{\partial q_k}\frac{\partial H^{(0)}}{\partial P_k},\qquad G^1_{(r_1,0)}=\tau_{r_1}\sum_{k=1}^n\frac{\partial H^{(0)}}{\partial q_k},\qquad G^1_{(0,r_1)}=\sigma_{r_1}\sum_{k=1}^n\frac{\partial H^{(0)}}{\partial P_k}, \tag{5.3a}$$

$$G^1_{(r_1,r_2)}=\sigma_{r_2}\tau_{r_1},\qquad G^1_{(0,r_1,r_2)}=\sigma_{r_2}\sigma_{r_1}\sum_{k_1,k_2=1}^n\frac{\partial^2 H^{(0)}}{\partial P_{k_1}\partial P_{k_2}}, \tag{5.3b}$$

$$G^1_{(r_1,0,r_2)}=\sigma_{r_2}\tau_{r_1}\sum_{k_1,k_2=1}^n\frac{\partial^2 H^{(0)}}{\partial q_{k_1}\partial P_{k_2}},\qquad G^1_{(r_1,r_2,0)}=\tau_{r_2}\tau_{r_1}\sum_{k_1,k_2=1}^n\frac{\partial^2 H^{(0)}}{\partial q_{k_1}\partial q_{k_2}}, \tag{5.3c}$$

$$G^1_{(r_1,r_2,r_3)}=0,\qquad G^1_{(r_1,r_2,r_3,r_4)}=0, \tag{5.3d}$$

where $1\le r_1,\cdots,r_4\le m$. The 1.5 order schemes are obtained by truncating the generating functions to multi-indexes $\alpha\in\mathcal{B}_{1.5}$. Using the approximation of S_ω^1 given in Eq. (3.24), we have the following symplectic implicit scheme of mean square order 1.5:

$$\begin{aligned}P_i(k+1)=&P_i(k)-\frac{\partial H^{(0)}}{\partial Q_i}h-\sum_{r=1}^{m}\Bigg(\sigma_r\sqrt{h}\xi_{hk}^{(r)}+\frac{\partial G^1_{(0,r)}}{\partial Q_i}\bar{I}^h_{(0,r)}+\frac{\partial G^1_{(r,0)}}{\partial Q_i}\bar{I}^h_{(r,0)}\\&+\frac{1}{4}\left(2\frac{\partial G^1_{(0,0)}}{\partial Q_i}+\frac{\partial G^1_{(0,r,r)}}{\partial Q_i}+\frac{\partial G^1_{(r,r,0)}}{\partial Q_i}\right)h^2\Bigg),\end{aligned}\tag{5.4a}$$

$$\begin{aligned}Q_i(k+1)=&Q_i(k)+\frac{\partial H^{(0)}}{\partial P_i}h+\sum_{r=1}^{m}\Bigg(\tau_r\sqrt{h}\xi_{hk}^{(r)}+\frac{\partial G^1_{(0,r)}}{\partial P_i}\bar{I}^h_{(0,r)}+\frac{\partial G^1_{(r,0)}}{\partial P_i}\bar{I}^h_{(r,0)}\\&+\frac{1}{4}\left(2\frac{\partial G^1_{(0,0)}}{\partial P_i}+\frac{\partial G^1_{(0,r,r)}}{\partial P_i}+\frac{\partial G^1_{(r,r,0)}}{\partial P_i}\right)h^2\Bigg),\end{aligned}\tag{5.4b}$$

where $i=1,\cdots,n$ and all the functions have $(P(k+1),Q(k))$ as their arguments. Here

$$\bar{I}^h_{(r,0)}=\frac{h^{\frac{3}{2}}}{2}\left(\xi_{hk}^{(r)}+\frac{\eta_{hk}^{(r)}}{\sqrt{3}}\right)\quad\text{and}\quad\bar{I}^h_{(0,r)}=\xi_{hk}^{(r)}h^{\frac{3}{2}}-\bar{I}^h_{(r,0)},$$

where at each time step k, $\xi_{hk}^{(r)}$ and $\eta_{hk}^{(r)}$ are independent bounded random variables as given in (4.1).

Analogously, for S^3_ω, we obtain,

$$G^3_{(0,0)}=0,\qquad G^3_{(r_1,r_2)}=0,\qquad G^3_{(r_1,0)}=-G^3_{(0,r_1)}=\frac{1}{2}T_r^T\nabla H^{(0)},\tag{5.5a}$$

$$G^3_{(r_1,r_2,r_3)}=0,\qquad G^3_{(r_1,r_2,0)}=G^3_{(0,r_1,r_2)}=-G^3_{(r_1,0,r_2)}=\frac{1}{4}T_{r_1}^T\nabla^2 H^{(0)}T_{r_2},\tag{5.5b}$$

$$G^3_{(r_1,r_2,r_3,r_4)}=0,\tag{5.5c}$$

where $T_r=J^{-1}\nabla H^{(r)}=(-\sigma_r,\cdots,-\sigma_r,\tau_r,\cdots,\tau_r)^T$ and $1\le r_1,\cdots r_4\le m$. The following 1.5 order scheme is derived based on the truncation of S^3_ω according to multi-indexes $\alpha\in\mathcal{B}_{1.5}$:

$$\begin{aligned}Y_{k+1}=&Y_k+J^{-1}\nabla H^{(0)}(Y_{k+\frac{1}{2}})h+\sum_{r=1}^{m}\Bigg(T_r\xi_{hk}^{(r)}+J^{-1}\nabla G^3_{(r,0)}(Y_{k+\frac{1}{2}})(\bar{I}^h_{(r,0)}-\bar{I}^h_{(0,r)})\\&+J^{-1}\nabla G^3_{(r,r,0)}(Y_{k+\frac{1}{2}})\frac{h^2}{2}\Bigg),\end{aligned}\tag{5.6}$$

where for each time step k, we have $Y_k=(P_k^T,Q_k^T)^T$ and the arguments are everywhere $Y_{k+1/2}=(Y_{k+1}+Y_k)/2$. The random variables $\bar{I}^h_{(r,0)}$, $\bar{I}^h_{(0,r)}$, $\xi_{hk}^{(r)}$ and $\eta_{hk}^{(r)}$ are the same as for (5.4).

Notice that the 1.5 symplectic methods (5.4) and (5.6) are implicit. These methods have a similar computational complexity as the 1.5 symplectic implicit Runge-Kutta method proposed in [6].

5.2 Separable SHS

Let consider the general autonomous SHS (1.1) with separable Hamiltonian functions such that

$$H^{(0)}(P,Q)=V_0(P)+U_0(Q), \quad H^{(r)}(P,Q)=U_r(Q), \quad r=1,\cdots,m. \tag{5.7}$$

In this case, the coefficients of S^1_ω become:

$$G^1_{(r_1,r_2)}=0, \qquad G^1_{(r_1,0)}=0, \qquad G^1_{(0,r1)}=\sum_{k=1}^{n}\frac{\partial U^{(r_1)}}{\partial q_k}\frac{\partial V^{(0)}}{\partial P_k}, \tag{5.8a}$$

$$G^1_{(0,0)}=\sum_{k=1}^{n}\frac{\partial U^{(0)}}{\partial q_k}\frac{\partial V^{(0)}}{\partial P_k}, \qquad G^1_{(r_1,r_2,r_3)}=G^1_{(r_1,r_2,0)}=G^1_{(r_1,0,r_2)}=0, \tag{5.8b}$$

$$G^1_{(0,r_1,r_2)}=\sum_{k_1,k_2=1}^{n}\frac{\partial U^{(r_2)}}{\partial q_{k_1}}\frac{\partial U^{(r_1)}}{\partial q_{k_2}}\frac{\partial^2 V^{(0)}}{\partial P_{k_1}\partial P_{k_2}}, \qquad G^1_{(r_1,r_2,r_3)}=0, \qquad G^1_{(r_1,r_2,r_3,r_4)}=0, \tag{5.8c}$$

where $1\le r_1,\cdots,r_4\le m$.

The following symplectic first order scheme based on S^1_ω is explicit, and it is different from the two explicit symplectic first mean square order partitioned Runge-Kutta methods presented in [8]:

$$P_i(k+1)=P_i(k)-\frac{\partial U^{(0)}}{\partial Q_i}(Q(k))h-\sum_{r=1}^{m}\frac{\partial U^{(r)}}{\partial Q_i}(Q(k))\sqrt{h}\xi^{(r)}_{hk}, \tag{5.9a}$$

$$Q_i(k+1)=Q_i(k)+\frac{\partial V^{(0)}}{\partial P_i}(P(k+1))h, \tag{5.9b}$$

where $i=1,\cdots,n$. An explicit 1.5 mean square order partitioned Runge-Kutta method was reported in [8].However, when the order increases to 1.5 or higher, the symplectic schemes based on the generating function S^1_ω are implicit. The 1.5 order scheme is provided below:

$$\begin{aligned}P_i(k+1)=P_i(k)-\frac{\partial U^{(0)}}{\partial Q_i}(Q(k))h-\sum_{r=1}^{m}\Bigg(&\frac{\partial U^{(r)}}{\partial Q_i}(Q(k))\sqrt{h}\xi^{(r)}_{hk}+\frac{\partial G^1_{(0,r)}}{\partial Q_i}(P(k+1),Q(k))\bar{I}^h_{(0,r)}\\&+\frac{1}{4}\left(2\frac{\partial G^1_{(0,0)}}{\partial Q_i}(P(k+1),Q(k))+\frac{\partial G^1_{(0,r,r)}}{\partial Q_i}(P(k+1),Q(k))\right)h^2\Bigg),\end{aligned} \tag{5.10a}$$

$$\begin{aligned}Q_i(k+1)=Q_i(k)+\frac{\partial V^{(0)}}{\partial P_i}(P(k+1))h+\sum_{r=1}^{m}\Bigg(&\frac{\partial G^1_{(0,r)}}{\partial P_i}(P(k+1),Q(k))\bar{I}^h_{(0,r)}\\&+\frac{1}{4}\left(2\frac{\partial G^1_{(0,0)}}{\partial P_i}(P(k+1),Q(k))+\frac{\partial G^1_{(0,r,r)}}{\partial P_i}(P(k+1),Q(k))\right)h^2\Bigg),\end{aligned} \tag{5.10b}$$

where $i=1,\cdots,n$ and the random variables are generated following the same procedure as for (5.4).

5.3 SHS preserving Hamiltonian functions

Unlike the deterministic cases, in general the SHSs no longer preserve with respect to time the Hamiltonian functions H_i, $i=0,\cdots,n$, even when the SHS is autonomous. However, using the chain rule of the Stratonovich stochastic integration, it is easy to verify for the Hamiltonian system (1.1) that the Hamiltonian functions $H^{(i)}$, $i=0,\cdots,m$ are invariant (i.e., $dH^{(i)}=0$), if and only if $\{H^{(i)},H^{(j)}\}=0$ for any $i,j=0,\cdots,m$, where the Poisson bracket is defined as

$$\{H^{(i)},H^{(j)}\}=\sum_{k=1}^{n}\Big(\frac{\partial H^{(j)}}{\partial Q_k}\frac{\partial H^{(i)}}{\partial P_k}-\frac{\partial H^{(i)}}{\partial Q_k}\frac{\partial H^{(j)}}{\partial P_k}\Big).$$

For systems preserving the Hamiltonian functions, the coefficients G^1_α of S^1_ω are invariant under the permutations on α, when $l(\alpha)=2$ because for any $r_1,r_2=0,\cdots,m$, we have

$$G^1_{(r_1,r_2)}=G^1_{(r_2,r_1)}=\sum_{k=1}^{n}\frac{\partial H^{(r_1)}}{\partial q_k}\frac{\partial H^{(r_2)}}{\partial P_k}. \tag{5.11}$$

Moreover, for $l(\alpha)=3$, from the formula (3.20) we easily see that $G^1_{(r_1,r_2,r_3)}=G^1_{(r_2,r_1,r_3)}$ for any $r_1,r_2,r_3=0,\cdots,m$. Also, since for any $k_1,k_2=1,\cdots,n$ and any $r_1,r_2=0,\cdots,m$, we have

$$\frac{\partial}{\partial q_{k_2}}\Big(\sum_{k_1=1}^{n}\frac{\partial H^{(r_2)}}{\partial q_{k_1}}\frac{\partial H^{(r_3)}}{\partial P_{k_1}}\Big)=\frac{\partial}{\partial q_{k_2}}\Big(\sum_{k_1=1}^{n}\frac{\partial H^{(r_3)}}{\partial q_{k_1}}\frac{\partial H^{(r_2)}}{\partial P_{k_1}}\Big), \tag{5.12}$$

a simple calculation confirms that $G^1_{(r_1,r_2,r_3)}=G^1_{(r_1,r_3,r_2)}$. Hence, G^1_α is also invariant under the permutation on α when $l(\alpha)=3$.

These properties are helpful not only to reduce the calculations for G^1_α, but the need of using approximation of high-order stochastic multiple integrals in the symplectic schemes based on the generating function S^1_ω is also avoided. For instance, when $m=1$, we have the following symplectic second order scheme:

$$\begin{aligned}P_i(k+1)=P_i(k)-\Big(&\frac{\partial G^1_{(0)}}{\partial Q_i}h+\frac{\partial G^1_{(1)}}{\partial Q_i}\sqrt{h}\xi_h+\frac{\partial G^1_{(0,0)}}{\partial Q_i}\frac{h^2}{2}+\frac{\partial G^1_{(1,1)}}{\partial Q_i}\frac{h\xi_h^2}{2}+\frac{\partial G^1_{(1,0)}}{\partial Q_i}\xi_h h^{\frac{3}{2}}\\&+\frac{\partial G^1_{(1,1,1)}}{\partial Q_i}\frac{h^{\frac{3}{2}}\xi_h^3}{6}+\frac{\partial G^1_{(1,1,0)}}{\partial Q_i}\frac{\xi_h^2h^2}{2}+\frac{\partial G^1_{(1,1,1,1)}}{\partial Q_i}\frac{h^2\xi_h^4}{24}\Big),\end{aligned} \tag{5.13a}$$

$$\begin{aligned}Q_{k+1}=Q_k+\Big(&\frac{\partial G^1_{(0)}}{\partial P_i}h+\frac{\partial G^1_{(1)}}{\partial P_i}\sqrt{h}\xi_h+\frac{\partial G^1_{(0,0)}}{\partial P_i}\frac{h^2}{2}+\frac{\partial G^1_{(1,1)}}{\partial P_i}\frac{h\xi_h^2}{2}+\frac{\partial G^1_{(1,0)}}{\partial P_i}\xi_h h^{\frac{3}{2}}\\&+\frac{\partial G^1_{(1,1,1)}}{\partial P_i}\frac{h^{\frac{3}{2}}\xi_h^3}{6}+\frac{\partial G^1_{(1,1,0)}}{\partial P_i}\frac{\xi_h^2h^2}{2}+\frac{\partial G^1_{(1,1,1,1)}}{\partial P_i}\frac{h^2\xi_h^4}{24}\Big),\end{aligned} \tag{5.13b}$$

where everywhere the arguments are (P_{k+1},Q_k), and we have used $J_{(0,1)}+J_{(1,0)}=J_{(1)}J_{(0)}$ and $J_{(0,1,1)}+J_{(1,0,1)}+J_{(1,1,0)}=J_{(1,1)}J_{(0)}$ (see the Corollary 3.1).

For the coefficients of S^3_ω, $\{H^{(r_1)},H^{(r_2)}\}=0$ for any $0\leq r_1,r_2\leq m$ implies that $G^3_{(r_1,r_2)}=0$ and $G^3_{(r_1,r_2,r_3,r_4)}=0$, $r_1,r_2,r_3,r_4=1,\cdots,m$. Moreover, a simple calculation shows that G^3_α is also invariant under the permutation on α, when $l(\alpha)=3$. Hence the second order midpoint symplectic scheme when $m=1$ is given by

$$\begin{aligned}Y_{k+1}=&Y_k+J^{-1}\nabla G^3_{(0)}(Y_{k+\frac{1}{2}})h+J^{-1}\nabla G^3_{(1)}(Y_{k+\frac{1}{2}})\sqrt{h}\xi_h\\&+J^{-1}\nabla G^3_{(1,1,1)}(Y_{k+\frac{1}{2}})\frac{h^{\frac{3}{2}}\xi_h^3}{6}+J^{-1}\nabla G^3_{(1,1,0)}(Y_{k+\frac{1}{2}})\frac{\xi_h^2h^2}{2},\end{aligned}\tag{5.14}$$

where $Y_{k+1/2}=(Y_{k+1}+Y_k)/2$.

It can be verified that G^3_α and G^1_α is invariant under the permutation on α for any $l(\alpha)$ (see [21]), and this property makes the higher order symplectic schemes computationally attractive for the SHS preserving Hamiltonian functions.

6 Numerical simulations and conclusions

To validate the high-order symplectic schemes proposed in this study, and to compare the performance with the lower order schemes, we consider three test cases. The cases have been used as the test examples in [6, 8, 9], and the last example is a nonlinear problem which is often used for testing numerical algorithms for stochastic computations [8].

6.1 SHS with additive noise

We now consider the following SHS with additive noise:

$$dP=Qdt+\sigma dw^1_t,\qquad P(0)=p,\tag{6.1a}$$

$$dQ=-Pdt+\gamma dw^2_t,\qquad Q(0)=q,\tag{6.1b}$$

where σ and γ are constant.

The exact solution can be expressed in the following form using the equal-distance time discretization $0=t_0<t_1<\cdots<t_N=T$, where the time-step h ($h=t_{k+1}-t_k$) is a small positive number:

$$X(t_{k+1})=FX(t_k)+u_k,\quad X(0)=X_0,\quad k=0,1,\cdots,N-1,\tag{6.2}$$

where

$$X(t_k)=\begin{bmatrix}P(t_k)\\Q(t_k)\end{bmatrix},\quad X_0=\begin{bmatrix}p\\q\end{bmatrix},\quad F=\begin{bmatrix}\cos h&\sin h\\-\sin h&\cos h\end{bmatrix},\tag{6.3a}$$

$$u_k=\begin{bmatrix}\sigma\int_{t_k}^{t_{k+1}}\cos(t_{k+1}-s)dw^1_s+\gamma\int_{t_k}^{t_{k+1}}\sin(t_{k+1}-s)dw^2_s\\-\sigma\int_{t_k}^{t_{k+1}}\sin(t_{k+1}-s)dw^1_s+\gamma\int_{t_k}^{t_{k+1}}\cos(t_{k+1}-s)dw^2_s\end{bmatrix}.\tag{6.3b}$$

The mean square order two symplectic scheme based on a truncation of S^1_ω according to multi-indexes $\alpha \in \mathcal{A}_2$ is given by

$$\begin{bmatrix} 1+\dfrac{h^2}{2} & 0 \\ h & 1 \end{bmatrix} X_{k+1} = \begin{bmatrix} 1 & h \\ 0 & 1+\dfrac{h^2}{2} \end{bmatrix} X_k + \begin{bmatrix} \sigma J_{(1)} + \gamma J_{(2,0)} \\ \gamma J_{(2)} + \sigma J_{(0,1)} \end{bmatrix}. \tag{6.4}$$

We have the following proposition about the long time error of the symplectic second order scheme.

Proposition 6.1. If T and h are positive values such that Th^2 and h are sufficiently small, and $E|X_0|^2$ is finite, then the mean square error is bounded by

$$\sqrt{E|X(t_k)-X_k|^2} \le K(\sqrt{T}h^2 + \sqrt{T^3h^4}), \quad k=1,2,\cdots,N. \tag{6.5}$$

Proof. As in the proof of Propositions 6.1 in [6], we can show that if T and h are positive values such that Th^2 and h are sufficiently small, then for $k=0,1,\cdots,N$, $T=Nh$, there exists a constant K_1 such that the following inequality holds:

$$\|H^k - F^k\| \le K_1(h^3+Th^2), \quad H = \begin{bmatrix} 1+\dfrac{h^2}{2} & 0 \\ h & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1 & h \\ 0 & 1+\dfrac{h^2}{2} \end{bmatrix}. \tag{6.6}$$

The proof then follows from the previous inequality, proceeding as in the proof of Propositions 6.2 in [6]. □

The corresponding error for the first-order scheme proposed in [6] is given by $\mathcal{O}(T^{1/2}h + T^{3/2}h^2)$. Clearly, a better performance is expected using the second-order scheme.

In numerical simulations, to guarantee that the exact solution, Euler scheme, first-order and second-order schemes have the same sample paths, eight independent standard normal distributed random variables, $\xi_{1,k}$, $\xi_{2,k}$, $\eta_{1,k}$, $\eta_{2,k}$, $\zeta_{1,k}$, $\zeta_{2,k}$, $\varepsilon_{1,k}$, $\varepsilon_{2,k}$ are used at every time step k. The random variables in (6.3) and (6.4) are evaluated as:

$$J_{(i)} = \sqrt{h}\xi_{1,k}, \quad \int_{t_k}^{t_{k+1}} \cos(t_{k+1}-s)dw^i_s = \frac{\sin h}{\sqrt{h}}\xi_{i,k} + a_1\eta_{i,k}, \tag{6.7a}$$

$$\int_{t_k}^{t_{k+1}} \sin(t_{k+1}-s)dw^i_s = \frac{2}{\sqrt{h}}\sin^2\frac{h}{2}\xi_{i,k} + a_2\eta_{i,k} + a_3\zeta_{i,k}, \tag{6.7b}$$

$$J_{(0,i)} = \frac{h^{\frac{3}{2}}}{2}\xi_{i,k} + a_4\eta_{i,k} + a_5\zeta_{i,k} + a_6\varepsilon_{i,k}, \quad J_{(i,0)} = hJ_{(i)} - J_{(0,i)}, \quad i=1,2, \tag{6.7c}$$

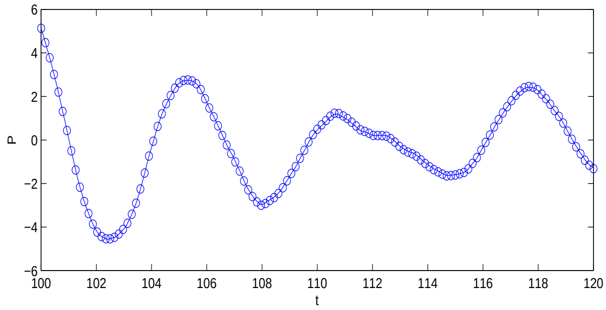

Figure 1: A sample trajectory of the solution to (6.1) for $\sigma=0$, $\tau=1$, $p=1$ and $q=0$: exact solution (solid line), S^1_ω second order scheme with time step $h=2^{-6}$ (circle). The circle of different scheme are plotted once per 10 steps.

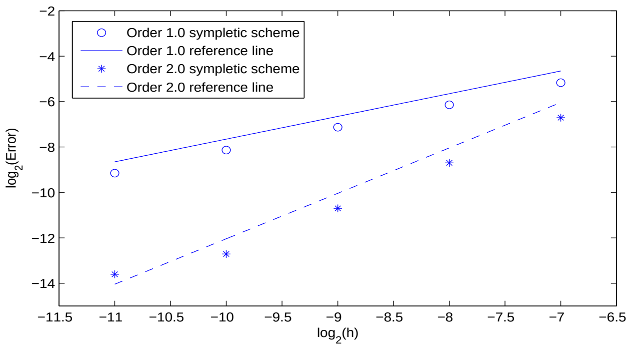

Figure 2: Convergence rate of different order S^1_ω symplectic scheme for (6.1), where error is the maximum error of (P,Q) at $T=100$.

where

$$a_1=\sqrt{\frac{h}{2}+\frac{\sin 2h}{4}-\frac{\sin^2 h}{h}}, \qquad a_2=\frac{1}{a_1}\left(\frac{\sin^2 h}{2}-\frac{2\sinh}{h}\sin^2\frac{h}{2}\right), \tag{6.8a}$$

$$a_3=\sqrt{\frac{h}{2}-\frac{\sin 2h}{4}-\frac{4}{h}\sin^4\frac{h}{2}-a_2^2}, \qquad a_4=\frac{1}{a_1}\left(1-\cos h-\frac{h\sin h}{2}\right), \tag{6.8b}$$

$$a_5=\frac{1}{a_3}\left(h-\sin h-h\sin^2\frac{h}{2}-a_2a_4\right), \qquad a_6=\sqrt{\frac{h^3}{12}-a_4^2-a_5^2}. \tag{6.8c}$$

Fig. 1 displays the results obtained using the symplectic schemes for long time simulations. An excellent agreement with the exact solution is observed when the symplectic scheme is implemented.

Fig. 2 presents the estimations of the convergence rate for various order symplectic schemes based on S^1_ω. We notice that the numerical results agree with the prediction based on the theoretical study. It is also clear that the second order symplectic scheme provides a more accuracy estimation than the first order scheme with the same time step.

6.2 Kubo oscillator

In [8], the following SDEs (the Kubo oscillator) in the sense of Stratonovich are used to demonstrate the advantage of the stochastic symplectic scheme for long time computation.

$$dP=-aQdt-\sigma Q\circ dw_t, \qquad P(0)=p, \tag{6.9a}$$
$$dQ=aPdt+\sigma P\circ dw_t, \qquad Q(0)=q, \tag{6.9b}$$

where a and σ are constants.

As illustrated in [8], the Hamiltonian functions

$$H^{(0)}(P(t),Q(t))=a\frac{P(t)^2+Q(t)^2}{2} \quad \text{and} \quad H^{(1)}(P(t),Q(t))=\sigma\frac{P(t)^2+Q(t)^2}{2}$$

are preserved under the phase flow of the systems. This means that the phase trajectory of (6.9) lies on the circle with the center at the origin and the radius $\sqrt{p^2+q^2}$.

Here, we consider the explicit Milstein first order scheme given in (4.24), and five stochastic symplectic schemes: the mean square 0.5, first and second order schemes based on S_ω^1, and the mean square first- and second-order schemes based on S_ω^3. The coefficients G_α^1 of S_ω^1 for the system (6.9) are given by:

$$G^1_{(0)}=\frac{a}{2}(p^2+q^2), \quad G^1_{(1)}=\frac{\sigma}{2}(p^2+q^2), \quad G^1_{(0,0)}=a^2pq, \quad G^1_{(1,1)}=\sigma^2pq, \tag{6.10a}$$
$$G^1_{(1,0)}=G^1_{(0,1)}=a\sigma pq, \quad G^1_{(0,0,0)}=a^3(p^2+q^2), \quad G^1_{(1,1,1)}=\sigma^3(p^2+q^2), \tag{6.10b}$$
$$G^1_{(1,1,0)}=G^1_{(1,0,1)}=G^1_{(0,1,1)}=a\sigma^2(p^2+q^2), \quad G^1_{(1,1,1,1)}=5\sigma^4pq. \tag{6.10c}$$

The various order symplectic schemes are obtained by truncating the generating function S_ω^1 appropriately (see (5.14)) for the second order scheme).

For S_ω^3, G_α^3 for SHSs preserving Hamiltonian functions is zero when $l(\alpha)=2,4$. Thus

$$G^3_{(0)}=\frac{a}{2}(p^2+q^2), \quad G^3_{(1)}=\frac{\sigma}{2}(p^2+q^2), \quad G^3_{(0,0,0)}=\frac{a^3}{4}(p^2+q^2), \tag{6.11a}$$
$$G^3_{(1,1,1)}=\frac{\sigma^3}{4}(p^2+q^2), \quad G^3_{(1,1,0)}=G^3_{(1,0,1)}=G^3_{(0,1,1)}\frac{a\sigma^2}{4}(p^2+q^2). \tag{6.11b}$$

The first-order midpoint scheme was already applied in [8] for the system (6.9) to illustrate the superior performance on the long time intervals compared to the non-symplectic schemes. The second-order midpoint scheme is given in (5.14).

Sample phase trajectories of (6.9) from various numerical scheme are presented in Fig. 3. It can be seen that the phase trajectory of the non-symplectic scheme deviated greatly away from the exact $P(t)^2+Q(t)^2=1$. However, the symplectic schemes produce accurate numerical solutions.

Figs. 4 and 5 confirm that the symplectic schemes have the expected convergence rate. Hence, the high order symplectic schemes have a more accuracy estimation than the low order symplectic scheme with the same time step.

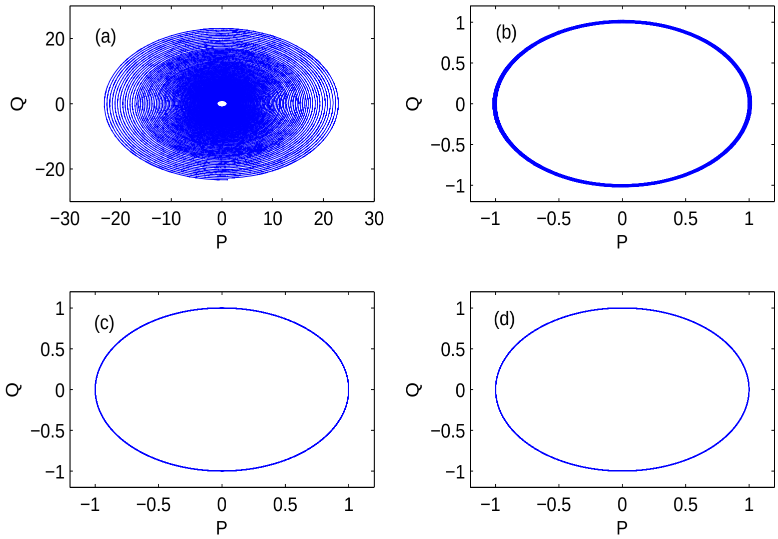

Figure 3: A sample phase trajectory of (6.9) with $a=2$, $\sigma=0.3$, $p=1$ and $q=0$: The Milstein scheme (a); S^1_ω first-order scheme (b); S^1_ω second-order scheme (c); S^3_ω second-order scheme (d) with time step $h=2^{-8}$ on the time interval $T\leq 200$.

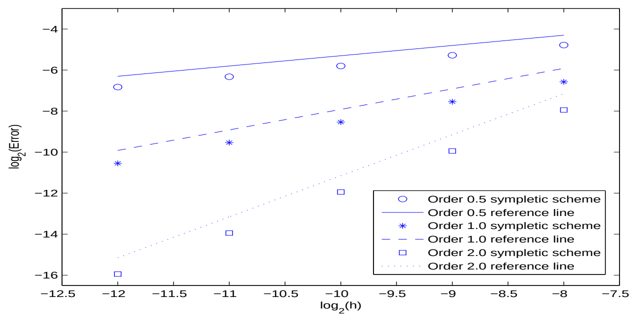

Figure 4: Convergence rate of different order S^1_ω symplectic scheme for (6.9), where error is the maximum error of (P,Q) at $T=100$.

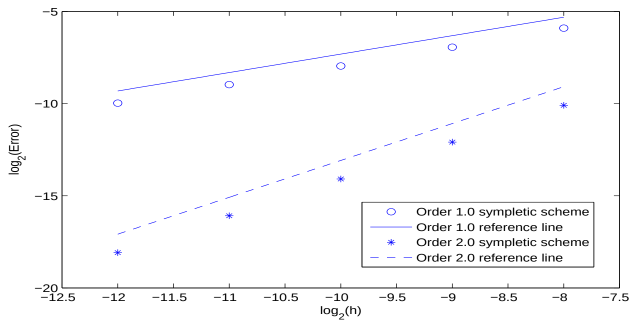

Figure 5: Convergence rate of different order S^3_ω symplectic scheme for (6.9), where error is the maximum error of (P,Q) at $T=100$.

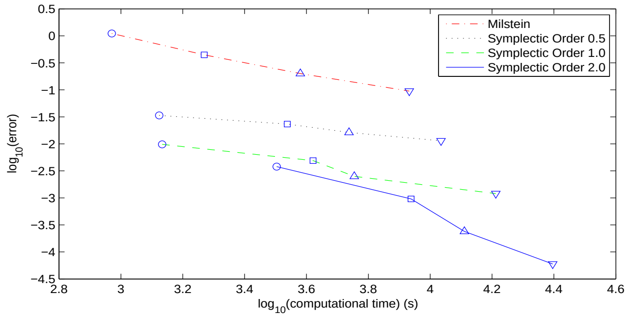

Figure 6: Computing time v.s. error for Milstein and different types of symplectic strong S_ω^1 scheme with various time step for $T=100$ with 10^5 samples, ◯: $h=0.004$; □: $h=0.002$; △: $h=0.001$, ▽: $h=0.0005$.

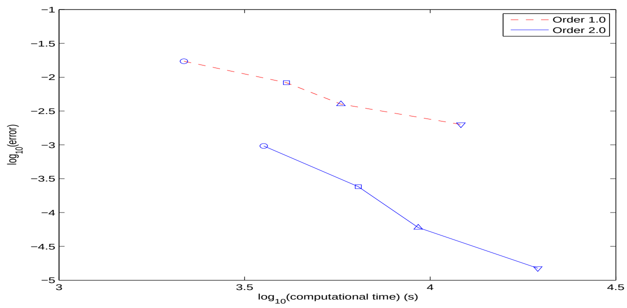

Figure 7: Computing time v.s. error for different types of symplectic strong S_ω^3 scheme with various time step for $T=100$ with 10^5 samples, ◯: $h=0.004$; □: $h=0.002$; △: $h=0.001$, ▽: $h=0.0005$.

Fig. 6 and Fig. 7 show that the higher order strong schemes are more efficient than the lower order ones. It takes about 4180 seconds to complete the first order S_ω^1 schemes simulation for $h=0.002$. The computing time for the second order S_ω^1 schemes with time step $h=0.004$ is about 3200 seconds. However, the error of second order S_ω^1 schemes is 0.0038, compared to 0.0049, the error of first order S_ω^1 scheme.

The numerical implementation for symplectic schemes of mean square order two or more is usually difficult because it requires the simulation of many multiple stochastic integrals. However, for SHSs preserving the Hamiltonian functions, the proposed higher order symplectic schemes have a simpler form because of the invariance of the coefficients under permutations.

6.3 Synchrotron oscillations

The mathematical model for oscillations of particles in storage rings is given by:

$$dP=-\beta^2\sin Q dt-\sigma_1\cos Q\circ dw_t^1-\sigma_2\sin Q\circ dw_t^2, \tag{6.12a}$$

$$dQ=Pdt. \tag{6.12b}$$

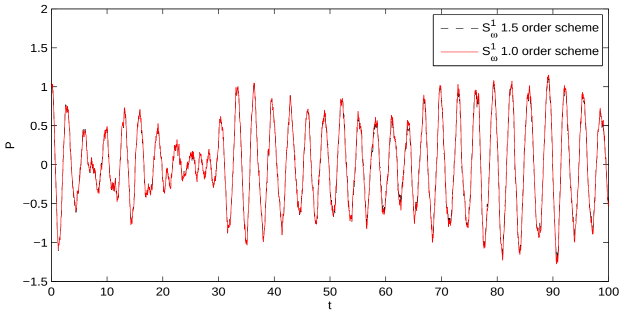

Figure 8: A sample trajectory of (6.12) for $\omega=2$, $\sigma_1=0.2$, $\sigma_2=0.1$, and time step $h=2^{-5}$.

We obtain the following formulas for the coefficients G^1_α of S^1_ω

$$G^1_{(0)}=\frac{p^2}{2}-\omega^2\cos q,\quad G^1_{(1)}=\sigma_1\sin q,\quad G^1_{(2)}=-\sigma_2\cos q,\quad G^1_{(0,0)}=\omega^2 p\sin q, \tag{6.13a}$$

$$G^1_{(0,1)}=\sigma_1 p\cos q,\quad G^1_{(0,2)}=\sigma_2 p\sin q,\quad G^1_{(0,1,1)}=\sigma_1^2\cos^2 q,\quad G^1_{(0,2,2)}=\sigma_2^2\sin^2 q. \tag{6.13b}$$

All other G^1_α in (3.24) are zero.

Since the exact solution of the nonlinear SHS (6.12) is not known, it is hard to verify the order of various symplectic schemes. However, using a very fine time step $h=2^{-8}$, we confirm that the sample trajectories from S^1_ω with first and 1.5 order are almost identical, and this is shown in Fig. 8. Moreover, the results also show that numerical schemes based on various order of S^1_ω are reliable for long time computation.

7 Conclusions

We present an approach to construct high-order symplectic schemes based on generating functions for stochastic Hamiltonian systems. The theoretical convergence analysis and numerical tests are provided for the proposed numerical methods. In general these symplectic schemes are implicit, and computationally expensive for mean square orders higher than two because they require generating approximations for multiple stochastic integrals of high order. It is also interesting to note that for stochastic Hamiltonian systems preserving Hamiltonian functions, the high order symplectic schemes have simpler forms and include less multiple stochastic integrals than the explicit Taylor expansion schemes.

Acknowledgments

We would like to thank Professor Tao Tang and the referees for their valuable comments and suggestions. This research was supported by the Natural Sciences and Engineering Research Council of Canada.

References

[1] R. de Vogelaere, Methods of integration which preserve the contact transformation property of the hamiltonian equations, Tech. Rep., Report No. 4, Dept. Math., Univ. of Notre Dame, Notre Dame, Ind., (1956).

[2] R. Ruth, A canonical integration technique, IEEE Trans. Nuclear Sci., 30 (1983), 2669–2671.

[3] K. Feng, On difference schemes and symplectic geometry, in: Proceedings of the 5-th Intern, Symposium on Differential Geometry & Differential Equations, Beijing, China, 1985, 42–58.

[4] X. Liu, Y. Qi, J. He and P. Ding, Recent progress in symplectic algorithms for use in quantum systems, Commun. Comput. Phys., 2(1) (2007), 1–53.

[5] J. Hong, S. Jiang, C. Li and H. Liu, Explicit multi-symplectic methods for Hamiltonian wave equations, Commun. Comput. Phys., 2 (2007), 662–683.

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

[7] H. Kunita, Stochastic Flows and Stochastic Differential Equations, Cambridge University Press, 1990.

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

[9] J. Hong, R. Scherer and L. Wang, Predictorcorrector methods for a linear stochastic oscillator with additive noise, Math. Comput. Model., 46(5-6) (2007), 738–764.

[10] A. Melbo and D. Higham, Numerical simulation of a linear stochastic oscillator with additive noise, Appl. Numer. Math., 51(1) (2004), 89–99.

[11] L. Wang, J. Hong, R. Scherer and F. Bai, Dynamics and variational integrators of stochastic Hamiltonian systems, Int. J. Numer. Anal. Model., 6(4) (2009), 586–602.

[12] J. Hong, L. Wang and R. Scherer, Simulation of stochastic Hamiltonian systems via generating functions, in: Proceedings IEEE 2011 4th ICCSIT, 2011.

[13] J. Hong, L. Wang and R. Scherer, Symplectic numerical methods for a linear stochastic oscillator with two additive noises, in: Proceedings of the World Congress on Engineering, Vol. I, London, U.K., 2011.

[14] L. Wang. Variational integrators and generating functions for stochastic Hamiltonian systems, KIT Scientific Publishing, University of Karlsruhe, Germany, 2007, Dissertation, `www.ksp.kit.edu`.

[15] J.-M. Bismut and Mécanique Aléatoire, Lecture Notes in Mathematics, Vol. 866, Springer-Verlag, Berlin-Heidelberg-New York, 1981.

[16] E. Hairer, Geometric Numerical Integration: Structure-Preserving Algorithms for Ordinary Differential Equations, Springer, Berlin, New York, 2006.

[17] K. Feng, H. Wu, M. Qin and D. Wang, Construction of canonical difference schemes for Hamiltonian formalism via generating functions, J. Comput. Math., 7 (1989), 71–96.

[18] C. Anton, J. Deng and Y. S. Wong, Weak symplectic schemes for stochastic Hamiltonian equations, Electronic Transactions on Numerical Analysis, to appear.

[19] P. Kloeden and E. Platen, Numerical Solutions of Stochastic Differential Equations, Springer-Verlag, Berlin, 1992.

[20] G. N. Milstein, Numerical Integration of Stochastic Differential Equations, Kluwer Academic Publishers, Dordrecht, Boston, 1995.

[21] C. Anton, Y. S. Wong and J. Deng, Symplectic schemes for stochastic Hamiltonian systems preserving Hamiltonian functions, Int. J. Numer. Anal. Model., 11(3) (2014), 427–451.