SIAM J. NUMER. ANAL.
Vol. 55, No. 6, pp. 3006–3029

HIGH ORDER CONFORMAL SYMPLECTIC AND ERGODIC SCHEMES FOR THE STOCHASTIC LANGEVIN EQUATION VIA GENERATING FUNCTIONS*

JIALIN HONG†, LIYING SUN‡, AND XU WANG†

Abstract. In this paper, we consider the stochastic Langevin equation with additive noises, which possesses both conformal symplectic geometric structure and ergodicity. We propose a methodology of constructing high weak order conformal symplectic schemes by converting the equation into an equivalent autonomous stochastic Hamiltonian system and modifying the associated generating function. To illustrate this approach, we construct a specific second order numerical scheme and prove that its symplectic form dissipates exponentially. Moreover, for the linear case, the proposed scheme is also shown to inherit the ergodicity of the original system, and the temporal average of the numerical solution is a proper approximation of the ergodic limit over long time. Numerical experiments are given to verify these theoretical results.

Key words. stochastic Langevin equation, conformal symplectic scheme, generating function, ergodicity, weak convergence

AMS subject classifications. 60H35, 65C30, 65P10

DOI. 10.1137/17M111691X

1. Introduction.

A common way to describe dissipative systems which interact with their environment, especially in the fields of molecular simulations, quantum systems, cell migrations, chemical interactions, electrical engineering, and finance (see [8, 10, 20] and references therein), is by means of the stochastic Langevin equation. The stochastic Langevin equation considered in this paper is a dissipative Hamiltonian system, whose phase flow preserves conformal symplectic geometric structure [4] as an extension of the deterministic case. Namely, its symplectic form dissipates exponentially. One can also show that the considered stochastic Langevin equation is ergodic [13, 14, 21] with a unique invariant measure, i.e., the Boltzmann–Gibbs measure [4, 6]. This dynamical behavior implies that the temporal average of the solution will converge to its spatial average, which is also known as the ergodic limit, with respect to the invariant measure over long time.

This work proposes an approach for constructing high weak order conformal symplectic schemes that accurately approximates the exact solution, while preserving both the geometric structure and the dynamical behavior of the system. We illustrate this approach by a specific case and show that the proposed scheme for this particular case inherits the ergodicity of the original system with a unique invariant measure. The weak convergence error, as well as the approximate error of the ergodic limit, is proved to be of order two.

*Received by the editors February 16, 2017; accepted for publication (in revised form) September 5, 2017; published electronically November 28, 2017.
http://www.siam.org/journals/sinum/55-6/M111691.html
Funding: The work of the authors was funded by the National Natural Science Foundation of China (91530118, 91130003, 11021101, 11290142, and 91630312).
†Institute of Computational Mathematics and Scientific/Engineering Computing, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, People's Republic of China (hjl@lsec.cc.ac.cn, wangxu@lsec.cc.ac.cn).
‡Corresponding author. Institute of Computational Mathematics and Scientific/Engineering Computing, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, People's Republic of China (liyingsun@lsec.cc.ac.cn).

There have been several works concentrating on the construction of numerical schemes for the stochastic Langevin equation, mainly based on the splitting technique. For instance, [4] constructs a class of the conformal symplectic integrators to preserve the conformal symplectic structure, and [18, 19] propose quasi-symplectic methods which can degenerate into symplectic ones when the system degenerates into a stochastic Hamiltonian system. The convergence rates of these schemes depend heavily on the splitting forms. As for the ergodicity, to the best of our knowledge its numerical analysis in general contains two aspects. The first is to construct numerical schemes that inherit the ergodicity (see, e.g., [13, 21]) and to give the error between the numerical invariant measure and the original one (see, e.g., [5, 7]). The other aspect is to approximate the ergodic limit with respect to the original invariant measure via the numerical temporal averages for some empirical test functions (see, e.g., [12, 14, 19]). In the latter case, the numerical solutions may not be ergodic.

In this paper, for the considered stochastic Langevin equation, we aim to construct numerical schemes which are of high weak order and are conformally symplectic. To achieve these aims without incurring the complexity of the high order splitting technique, we introduce a transformation from the stochastic Langevin equation to an autonomous stochastic Hamiltonian system. It then suffices to construct high order symplectic schemes for the autonomous Hamiltonian system, which turn out to be conformal symplectic schemes of the original system based on the inverse transformation of the phase spaces. The discretization of the modified equations, which are constructed by modifying the drift and diffusion functions as polynomials with respect to some time step, represents a powerful tool for obtaining high weak order schemes. For example, [1] constructs high order stochastic numerical integrators for general stochastic differential equations (SDEs), but these schemes may not be symplectic when applied to the Hamiltonian systems. Based on the internal properties of the Hamiltonian systems, [2] proposes a method for constructing high weak order stochastic symplectic schemes with multiple stochastic Itô integrals, using truncated generating functions. Based on these schemes, [24] gives their associated modified equations via generating functions. To reduce the simulation cost and still get high weak order symplectic schemes, inspired by [1, 2, 24], we modify the generating function for the equivalent stochastic Hamiltonian system and derive associated symplectic numerical methods by truncating modified generating functions. We would like to mention that this class of methods reduces the simulation of multiple stochastic Itô integrals by simulating products of increments of Wiener processes instead. We illustrate this approach with the construction of a stochastic numerical scheme that has weak order two. For the proposed numerical scheme, both the discretized phase volume and symplectic form dissipate exponentially, which coincides with the behavior of their exact counterparts in the original stochastic Langevin equation. Furthermore, the proposed scheme, similar to the original system, is proved to possess a numerical invariant measure that is unique for the linear case, which implies the ergodicity of the numerical solution. Finally, we verify that both the weak convergence error of the numerical scheme and the error of ergodic limit are of order two.

An outline of this paper is as follows. Section 2 gives a review of some basic properties of the stochastic Langevin equation, as well as the generating function of the stochastic Hamiltonian system, and also the transformation between the stochastic Langevin equation and an autonomous stochastic Hamiltonian system. In section 3, a weakly convergent conformal numerical scheme, which possesses an invariant measure, is proposed by means of modified generating functions and the transformation of phase space. In section 4, we show that both the weak convergence rate of the proposed scheme and the approximation error of the ergodic limit are of order two, based on the uniform estimate of the numerical solutions. Finally, we give some numerical tests to verify the theoretical results in section 5.

2. Stochastic Langevin equations. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, $\mathcal{F}_t$ be the filtration for $t \geq 0$, and $W(t) = \big(W_1(t), \ldots, W_m(t)\big)^\top$ be an m-dimensional standard Wiener process associated to $\{\mathcal{F}_t\}_{t\geq 0}$. Denote the 2-norm for both matrices and vectors by $\|\cdot\|$ and the determinant of matrices by $|\cdot|$, and let C be a generic constant, independent of h, that may differ from line to line.

2.1. Stochastic conformal symplectic structure and ergodicity. In this section, we focus on the stochastic Langevin equation driven by additive noises with deterministic initial values $P(0) = p \in \mathbb{R}^d$ and $Q(0) = q \in \mathbb{R}^d$, of the following form:

$$
\begin{aligned}
dP &= -f(Q)dt - vPdt - \sum_{r=1}^{m} \sigma_r dW_r(t), \\
dQ &= MPdt, \quad t \in [0, T],
\end{aligned}
\tag{1}
$$

where $f \in C^\infty(\mathbb{R}^d, \mathbb{R}^d)$, $M \in \mathbb{R}^{d\times d}$ is a positive definite symmetric matrix, $v > 0$ is the absorption coefficient, and $\sigma_r \in \mathbb{R}^d$ with $r \in \{1, \ldots, m\}$, $m \geq d$, and $\operatorname{rank}\{\sigma_1, \ldots, \sigma_m\} = d$. In addition, assume that there exists a scalar function $F \in C^\infty(\mathbb{R}^d, \mathbb{R})$ satisfying

$$
f_i(Q) = \frac{\partial F(Q)}{\partial Q_i}, \quad i = 1, \ldots, d.
$$

To simplify the notation, we will remove any mention of the dependence on $\omega \in \Omega$ unless it is absolutely necessary to avoid confusion. Note that (1), as well as all the other SDEs in what follows, holds almost surely with respect to $\mathbb{P}$. It is well known that if $v = 0$, (1) turns out to be a separable stochastic Hamiltonian system (SHS) which possesses stochastic symplectic structure and phase volume preservation [17]. However, when $v > 0$, the symplectic form of (1) dissipates exponentially, i.e.,

$$
dP(t) \wedge dQ(t) = e^{-vt} dp \wedge dq \quad \forall\, t \geq 0,
$$

which characterizes the long-time tracking of the solutions to (1), as well as the phase volume $\operatorname{Vol}(t)$. Namely, denoting by $D_t = D_t(\omega) \subset \mathbb{R}^{2d}$ a random domain which has finite volume and is independent of Wiener processes $W(t)$ with respect to the system (1), one can obtain

$$
\begin{aligned}
\operatorname{Vol}(t) &= \int_{D_t} dP^1 \cdots dP^d dQ^1 \cdots dQ^d \\
&= \int_{D_0} \left| \frac{D(P^1, \ldots, P^d, Q^1, \ldots, Q^d)}{D(p^1, \ldots, p^d, q^1, \ldots, q^d)} \right| dp^1 \cdots dp^d dq^1 \cdots dq^d,
\end{aligned}
$$

where the determinant of Jacobian matrix $\left| \frac{D(P^1,\ldots,P^d,Q^1,\ldots,Q^d)}{D(p^1,\ldots,p^d,q^1,\ldots,q^d)} \right| = e^{-vtd}$ with d being the dimension [16, 17].

As another well-known long-time behavior, the ergodicity of (1) is shown in [13] by proving that (1) possesses a unique invariant measure μ. Noticing that (1) satisfies the hypoelliptic setting

$$
\operatorname{span}\{U_i, [U_0, U_j],\, i = 0, \ldots, m,\, j = 1, \ldots, m\} = \mathbb{R}^{2d}
\tag{2}
$$

with vector fields $U_0 = ((-f(Q) - vP)^\top, (MP)^\top)^\top$ and $U_j = (\sigma_j^\top, 0)^\top$, $j = 1, \ldots, m$, which together with the following assumption yields the ergodicity of (1).

ASSUMPTION 2.1 (see [13]). *Let $F \in C^\infty(\mathbb{R}^d, \mathbb{R})$ satisfy that*

(i) *$F(u) \geq 0$ for all $u \in \mathbb{R}^d$;*

(ii) *there exist $\alpha > 0$ and $\beta \in (0, 1)$ such that for all $u \in \mathbb{R}^d$, it holds*

$$\frac{1}{2} u^\top f(u) \geq \beta F(u) + v^2 \frac{\beta(2-\beta)}{8(1-\beta)} \|u\|^2 - \alpha.$$

Intuitively speaking, the ergodicity of (1) reads that the temporal averages of $P(t)$ and $Q(t)$ starting from different initial values will converge almost everywhere to its spatial average with respect to the invariant measure μ. More precisely,

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T \mathbf{E}^{(p,q)} \left[\psi(P(t), Q(t))\right] dt = \int_{\mathbb{R}^{2d}} \psi d\mu \quad \forall\, \psi \in C_b(\mathbb{R}^{2d}, \mathbb{R}) \tag{3}$$

in $L^2(\mathbb{R}^{2d}, \mu)$, where $\mathbf{E}^{(p,q)}[\cdot]$ denotes the expectation starting from $P(0) = p$ and $Q(0) = q$. In the following, we use the notation $\mathbf{E}$ instead of $\mathbf{E}^{(p,q)}$ to simplify the notation.

Next, we aim to convert (1) into an equivalent homogenous SHS via a transformation of phase space, such that one can construct conformal symplectic schemes for (1) based on symplectic schemes of the homogenous SHS. To this end, denoting $X_i(t) = e^{vt} P_i(t)$ and $Y_i(t) = Q_i(t)$ and applying Itô's formula to $X_i(t)$ and $Y_i(t)$ for $i = 1, \ldots, d$, one can rewrite (1) as

$$dX_i = -e^{vt} f_i(Y_1, \ldots, Y_d)dt - e^{vt} \sum_{r=1}^m \sigma_r dW_r(t), \quad dY_i = e^{-vt} \sum_{j=1}^d M_{ij} X_j dt \tag{4}$$

with $X_i(0) = p_i$ and $Y_i(0) = q_i$. It is obvious that (4) is a nonautonomous SHS with time-dependent Hamiltonian functions

$$\tilde{H}_0 = e^{vt} F(Y_1, \ldots, Y_d) + \frac{1}{2} e^{-vt} \sum_{i,j=1}^d X_i M_{ij} X_j, \quad \tilde{H}_r = e^{vt} \sum_{i=1}^d \sigma_r^i Y_i.$$

To obtain an autonomous SHS we introduce two new variables $X_{d+1} \in \mathbb{R}$ and $Y_{d+1} \in \mathbb{R}$ as the $(d+1)$th components of X and Y, respectively, satisfying

$$dY_{d+1} = dt, \quad dX_{d+1} = -\frac{\partial \tilde{H}_0}{\partial t} dt - \sum_{r=1}^m \frac{\partial \tilde{H}_r}{\partial t} \circ dW_r(t)$$

with $Y_{d+1}(0) = 0$ and $X_{d+1}(0) = F(q_1, \ldots, q_d) + \frac{1}{2} \sum_{i,j=1}^d p_i M_{ij} p_j + \sum_{r=1}^m \sum_{i=1}^d \sigma_r^i q_i$. Here the notation "$\circ$" means that the equation holds in the Stratonovich integral sense. Then (4) becomes the $(2d+2)$-dimensional autonomous SHS

$$dX = -\frac{\partial H_0}{\partial Y} dt - \sum_{r=1}^m \frac{\partial H_r}{\partial Y} \circ dW_r(t), \quad dY = \frac{\partial H_0}{\partial X} dt + \sum_{r=1}^m \frac{\partial H_r}{\partial X} \circ dW_r(t), \tag{5}$$

with $X(0) = (X_1(0), \ldots, X_{d+1}(0)) \in \mathbb{R}^{d+1}$, $Y(0) = (Y_1(0), \ldots, Y_{d+1}(0)) \in \mathbb{R}^{d+1}$, and new Hamiltonian functions

$$
\begin{aligned}
H_0(X,Y) &= e^{vY_{d+1}}F(Y_1,\dots,Y_d) + \frac{1}{2}e^{-vY_{d+1}}\sum_{i,j=1}^{d} X_i M_{ij} X_j + X_{d+1},\\
H_r(X,Y) &= e^{vY_{d+1}}\sum_{i=1}^{d}\sigma_r^i Y_i.
\end{aligned}
$$

Here, (5) is called the associated autonomous SHS of (1), and its phase flow preserves the stochastic symplectic structure. Notice that the motion of the system can be described by different kinds of generating functions (see [2, 23] and references therein). We consider only the first kind of generating function S in this article.

2.2. Generating functions. For convenience, we denote $X(0)=x$ and $Y(0)=y$. It is revealed in [22] that the generating function $S(X,y,t)$ related to (5) is the solution of the following stochastic Hamilton–Jacobi partial differential equation:

$$
d_t S(X,y,t) = H_0\left(X, y+\frac{\partial S}{\partial X}\right)dt + \sum_{r=1}^{m} H_r\left(X, y+\frac{\partial S}{\partial X}\right)\circ dW_r(t). \tag{6}
$$

Moreover, the mapping $(x,y)\mapsto (X(t),Y(t))$ defined by

$$
X(t) = x - \frac{\partial S(X(t),y,t)}{\partial y},\quad Y(t) = y + \frac{\partial S(X(t),y,t)}{\partial X} \tag{7}
$$

is the stochastic flow of (5). Based on the Itô representation theorem and stochastic Taylor–Stratonovich expansion, $S(X,y,t)$ has a series expansion (see, e.g., [2, 3])

$$
S(X,y,t) = \sum_{\boldsymbol{\alpha}} G_{\boldsymbol{\alpha}}(X,y) J_{\boldsymbol{\alpha}}^t, \tag{8}
$$

where

$$
J_{\boldsymbol{\alpha}}^t = \int_0^t \int_0^{s_l}\cdots\int_0^{s_2} \circ dW_{j_1}(s_1)\circ dW_{j_2}(s_2)\circ\cdots\circ dW_{j_l}(s_l)
$$

with multi-index $\boldsymbol{\alpha} = (j_1, j_2, \dots, j_l) \in \{0,1,\dots,m\}^{\otimes l}$, $l \geq 1$, and $dW_0(s) := ds$. Before calculating coefficients $G_{\boldsymbol{\alpha}}(X,y)$ in (8), we first specify some notation. Let $l(\boldsymbol{\alpha})$ denote the length of $\boldsymbol{\alpha}$, and let $\boldsymbol{\alpha}-$ be the multi-index resulting from discarding the last index of $\boldsymbol{\alpha}$. Define $\boldsymbol{\alpha}*\boldsymbol{\alpha}' = (j_1,\dots,j_l,j_1',\dots,j_{l'}')$, where $\boldsymbol{\alpha}=(j_1,\dots,j_l)$ and $\boldsymbol{\alpha}'=(j_1',\dots,j_{l'}')$. The concatenation "$*$" between a set of multi-indices Λ and $\boldsymbol{\alpha}$ is $\Lambda*\boldsymbol{\alpha} = \{\boldsymbol{\beta}*\boldsymbol{\alpha} | \boldsymbol{\beta}\in\Lambda\}$. Furthermore, define

$$
\Lambda_{\boldsymbol{\alpha},\boldsymbol{\alpha}'} = \begin{cases}
\{(j_1,j_1'),(j_1',j_1)\} & \text{if} \quad l=l'=1,\\
\{\Lambda_{(j_1),\boldsymbol{\alpha}'-}*(j_{l'}'), \boldsymbol{\alpha}'*(j_1)\} & \text{if} \quad l=1, l'\neq 1,\\
\{\Lambda_{\boldsymbol{\alpha}-,(j_1')}*(j_l), \boldsymbol{\alpha}*(j_1')\} & \text{if} \quad l\neq 1, l'=1,\\
\{\Lambda_{\boldsymbol{\alpha}-,\boldsymbol{\alpha}'}*(j_l), \Lambda_{\boldsymbol{\alpha},\boldsymbol{\alpha}'-}*(j_{l'}')\} & \text{if} \quad l\neq 1, l'\neq 1.
\end{cases}
$$

For $k>2$, let $\Lambda_{\boldsymbol{\alpha}_1,\dots,\boldsymbol{\alpha}_k} = \{\Lambda_{\boldsymbol{\beta},\boldsymbol{\alpha}_k} | \boldsymbol{\beta}\in\Lambda_{\boldsymbol{\alpha}_1,\dots,\boldsymbol{\alpha}_{k-1}}\}$. We refer the reader to [2] for more details about this notation. Substituting (8) into (6) and applying Taylor expansions to H_r $(r=0,1,\dots,m)$ at (X,y), we obtain $G_{\boldsymbol{\alpha}} = H_r$ with $\boldsymbol{\alpha}=(r)$ being a single index and

$$G_{\boldsymbol{\alpha}} = \sum_{i=1}^{l(\boldsymbol{\alpha})-1} \frac{1}{i!} \sum_{k_1,\ldots,k_i=1}^{d+1} \frac{\partial^i H_{j_l}(X,y)}{\partial y_{k_1}\cdots\partial y_{k_i}} \sum_{\substack{l(\boldsymbol{\alpha}_1)+\cdots+l(\boldsymbol{\alpha}_i)=l(\boldsymbol{\alpha})-1 \\ \boldsymbol{\alpha}-\in\Lambda_{\boldsymbol{\alpha}_1,\ldots,\boldsymbol{\alpha}_i}}} \frac{\partial G_{\boldsymbol{\alpha}_1}}{\partial X_{k_1}}\cdots\frac{\partial G_{\boldsymbol{\alpha}_i}}{\partial X_{k_i}}$$

for any $\boldsymbol{\alpha} = (j_1, j_2, \ldots, j_l)$ with $l \geq 2$ (see, e.g., [2, 3]). To make it clear, the simplified expressions of $G_{\boldsymbol{\alpha}}$ are given when $l = 2$ or 3: $G_{(j_1,j_2)} = \sum_{i=1}^{d+1} \frac{\partial H_{j_2}}{\partial y_i}\frac{\partial H_{j_1}}{\partial X_i}$ and

$$G_{(j_1,j_2,j_3)} = \sum_{i=1}^{d+1} \frac{\partial H_{j_3}}{\partial y_i}\frac{\partial G_{(j_1,j_2)}}{\partial X_i} + \frac{1}{2}\sum_{i,j=1}^{d+1}\frac{\partial^2 H_{j_3}}{\partial y_i\partial y_j}\left(\frac{\partial H_{j_1}}{\partial X_i}\frac{\partial H_{j_2}}{\partial X_j} + \frac{\partial H_{j_2}}{\partial X_i}\frac{\partial H_{j_1}}{\partial X_j}\right).$$

Let $C_1 := e^{vy_{d+1}}$ and $C_2 := e^{-vy_{d+1}}$. Here y_{d+1} denotes the $(d+1)$th component of y. Note that y is the initial point of the considered interval; that is, if we consider the problem on the interval $[s,t]$, then $y = Y(s)$. For $r_1, r_2, r_3 \in \{1, \ldots, m\}$, we have

$$\begin{aligned}
&G_{(r_1,r_2)} = G_{(r_1,0)} = G_{(r_1,r_2,r_3)} = G_{(r_1,r_2,0)} = G_{(r_1,0,r_2)} = 0,\\
&G_{(0,r_1)} = \sum_{i,j=1}^{d} \sigma_{r_1}^i M_{ij}X_j + vC_1\sum_{i=1}^{d}\sigma_{r_1}^i q_i, \quad G_{(0,r_1,r_2)} = C_1\sigma_{r_1}^\top M\sigma_{r_2},\\
&G_{(0,0)} = \sum_{i,j=1}^{d} f_i(y)M_{ij}X_j + vC_1F(y) - \frac{1}{2}vC_2\sum_{i,j=1}^{d}X_iM_{ij}X_j.
\end{aligned}$$

For a fixed small time step h, using (8) and applying Taylor expansion to $\frac{\partial S}{\partial y_i} := \frac{\partial S}{\partial y_i}(X,y,h)$ and $\frac{\partial S}{\partial X_i} := \frac{\partial S}{\partial X_i}(X,y,h)$ at point (x,y,h) for $i = 1, \ldots, d$, we obtain

$$\begin{aligned}
\frac{\partial S}{\partial y_i} &= C_1\left[\sum_{r=1}^{m}\sigma_r^i(J_{(r)}^h + vJ_{(0,r)}^h) + f_i(y)\left(h + \frac{vh^2}{2}\right)\right] + \frac{h^2}{2}\sum_{j,k=1}^{d}\frac{\partial^2 F(y)}{\partial y_i\partial y_j}M_{jk}x_k + R_1,\\
\frac{\partial S}{\partial X_i} &= C_2\sum_{j=1}^{d}M_{ij}x_j\left(h - \frac{vh^2}{2}\right) - \sum_{j=1}^{d}\sum_{r=1}^{m}M_{ij}\sigma_r^jJ_{(r,0)}^h - \frac{h^2}{2}\sum_{j=1}^{d}M_{ij}f_j(y) + R_2,
\end{aligned}$$

where every term in R_1 and R_2 contains the product of multiply stochastic integrals whose lowest order is at least $\frac{5}{2}$, as do the remainder terms R_l with $l = 3, \ldots, 7$ in what follows. Furthermore, $\frac{\partial S}{\partial X_{d+1}}(X,y,h) = h$ and

$$\begin{aligned}
\frac{\partial S}{\partial y_{d+1}} = vh&\left(C_1F(y) - \frac{C_2}{2}\sum_{i,j=1}^{d}x_iM_{ij}x_j\right)\left(1 + \frac{vh}{2}\right) + vC_1\sum_{r=1}^{m}\sum_{i=1}^{d}\sigma_r^iy_i(J_{(r)}^h + vJ_{(0,r)}^h)\\
&+ \sum_{i,j=1}^{d}\sum_{r=1}^{m}v\sigma_r^iM_{ij}x_jhJ_{(r)}^h + vC_1\sum_{r_1,r_2=1}^{m}\sigma_{r_1}^\top M\sigma_{r_2}J_{(0,r_1,r_2)}^h\\
&+ v\sum_{i,j=1}^{d}\left(C_2\frac{\partial F(y)}{\partial y_i}M_{ij}x_jh^2 - \frac{1}{2}C_1\sum_{r_1,r_2=1}^{m}\sigma_{r_1}^iM_{ij}\sigma_{r_2}^jhJ_{(r_1)}^hJ_{(r_2)}^h\right) + R_3,
\end{aligned}$$

where $\frac{\partial S}{\partial y_{d+1}}$ takes the value at (X,y,h).

By truncating the generating function, the weakly convergent stochastic symplectic numerical schemes have been proposed by several authors (see, e.g., [2, 17, 22]). In these approaches, some techniques are applied to simulate the multiple integrals in the truncated generating functions and obtain high weak order schemes. To reduce the simulation of multiple integrals, we introduce a modified generating function to construct more concise symplectic schemes in section 3, from which conformal symplectic and ergodic schemes for stochastic dynamical systems (1) are deduced by using the transformation of the phase space.

3. High order conformal symplectic and ergodic schemes.

To construct high order symplectic numerical integrators for (5), we modify the stochastic Hamiltonian functions first. Namely, we consider the following $(2d+2)$-dimensional stochastic Hamiltonian system:

$$\begin{aligned} dX^M &= -\frac{\partial H_0^M(X^M,Y^M)}{\partial Y^M}dt - \sum_{r=1}^{m}\frac{\partial H_r^M(X^M,Y^M)}{\partial Y^M}\circ dW_r(t), \quad X^M(0)=x,\\ dY^M &= \frac{\partial H_0^M(X^M,Y^M)}{\partial X^M}dt + \sum_{r=1}^{m}\frac{\partial H_r^M(X^M,Y^M)}{\partial X^M}\circ dW_r(t), \quad Y^M(0)=y, \end{aligned} \tag{9}$$

where

$$\begin{aligned} H_0^M(X^M,Y^M) &= H_0(X^M,Y^M) + H_0^{[1]}(X^M,Y^M)h + \cdots + H_0^{[\tau]}(X^M,Y^M)h^\tau,\\ H_r^M(X^M,Y^M) &= H_r(X^M,Y^M) + H_r^{[1]}(X^M,Y^M)h + \cdots + H_r^{[\tau]}(X^M,Y^M)h^\tau \end{aligned} \tag{10}$$

with functions $H_i^{[j]}$, $i=0,\dots r$, $j=1,\dots,\tau$, $\tau\in\mathbb{N}_+$ to be determined. Meanwhile, according to the definition of $G_{\boldsymbol\alpha}$ in subsection 2.2, we get the associated generating function of (9), which is called the modified generating function of (5). Our goal is to choose undetermined functions in (10) such that the proposed scheme is of weak order $k+k'$ when approximating (5), even though it is only a kth order approximation of (9) for some positive integers k and k'. Now we first give a symplectic numerical approximation to (9) via its generating function, such that this scheme shows weak order k for (9) without specific choices of $H_i^{[j]}$ (see [2] and references therein). In detail, we replace the multiple Stratonovich integrals $J_{\boldsymbol\alpha}^t$ in the modified generating function by an equivalent linear combination of multiple Itô integrals

$$I_{\boldsymbol\beta}^t := \int_0^t\int_0^{s_l}\cdots\int_0^{s_2} dW_{i_1}(s_1)dW_{i_2}(s_2)\cdots dW_{i_l}(s_l)$$

with multi-index $\boldsymbol\beta=(i_1,i_2,\dots,i_l)\in\{0,1,\dots,m\}^{\otimes l}$, $l\ge 1$, based on the relation

$$J_{\boldsymbol\alpha}^t = \begin{cases} \sum_{\boldsymbol\beta} C_{\boldsymbol\alpha}^{\boldsymbol\beta} I_{\boldsymbol\beta}^t, & l(\boldsymbol\alpha)\ge 2,\\ I_{\boldsymbol\alpha}^t, & l(\boldsymbol\alpha)=1, \end{cases}$$

where $C_{\boldsymbol\alpha}^{\boldsymbol\beta}$ are certain constants given in [11]. Denote by

$$S^G(X^G,y,t) = \sum_{\boldsymbol\alpha} G_{\boldsymbol\alpha}^G(X^G,y)\sum_{l(\boldsymbol\beta)\le k} C_{\boldsymbol\alpha}^{\boldsymbol\beta} I_{\boldsymbol\beta}^t \tag{11}$$

the truncated modified generating function (see, e.g., [2, 3, 11]), where

$$G^G_{\boldsymbol{\alpha}} = \sum_{i=1}^{l(\boldsymbol{\alpha})-1} \frac{1}{i!} \sum_{k_1,\ldots,k_i=1}^{d+1} \frac{\partial^i H^M_{j_l}(X^G, y)}{\partial y_{k_1}\cdots\partial y_{k_i}} \sum_{\substack{l(\boldsymbol{\alpha}_1)+\cdots+l(\boldsymbol{\alpha}_i)=l(\boldsymbol{\alpha})-1 \\ \boldsymbol{\alpha}-\in\Lambda_{\boldsymbol{\alpha}_1,\ldots,\boldsymbol{\alpha}_i}}} \frac{\partial G^G_{\boldsymbol{\alpha}_1}}{\partial X^G_{k_1}} \cdots \frac{\partial G^G_{\boldsymbol{\alpha}_i}}{\partial X^G_{k_i}}$$

for $l(\boldsymbol{\alpha}) \geq 2$, and $G^G_{(r)} = H^M_r$ for $r = 0, 1, \ldots, m$. Then we get the following one-step approximation:

$$X^G = x - \frac{\partial S^G(X^G, y, h)}{\partial y}, \quad Y^G = y + \frac{\partial S^G(X^G, y, h)}{\partial X^G}, \tag{12}$$

which preserves symplectic structure and is of weak order k for (9). Notice that the truncated modified generating function contains undetermined functions $H^{[j]}_i$, $i = 0, \ldots r$, $j = 1, \ldots, \tau$ in (10). To specify high weak order symplectic schemes, we need to determine all the terms $H^{[j]}_i$ such that the numerical scheme based on (12) satisfies

$$\left|\mathbf{E}\phi(X(h), Y(h)) - \mathbf{E}\phi(X^G, Y^G)\right| = O(h^{k+k'+1}) \tag{13}$$

for all κ times continuously differentiable functions $\phi \in C^\kappa_P(\mathbb{R}^{2d+2}, \mathbb{R})$ with polynomial growth; that is, the numerical scheme based on (12) is of weak order $k + k'$ for (5). Conditions on κ will be specified later. The detailed approach of choosing the undetermined functions will be illustrated with the case $k = k' = 1$ in the next section.

3.1. Numerical schemes via modified generating function. For $k = k' = 1$, it is sufficient to consider $\tau = 1$ in (10). Based on the fact that $G^G_{(r)} = H^M_r$ for $r = 0, 1, \ldots, m$, we rewrite the truncated generating function (11) as

$$S^G(X^G, y, h) = \left(H^M_0(X^G, y) + \frac{1}{2}\sum_{r=1}^m G^G_{(r,r)}(X^G, y)\right) h + \sum_{r=1}^m H^M_r(X^G, y) I^h_{(r)}, \tag{14}$$

where

$$G^G_{(r,r)} = C_1 \sum_{i=1}^d \sigma^i_r \left(\frac{\partial H^{[1]}_r}{\partial X^G_i} + v y_i \frac{\partial H^{[1]}_r}{\partial X^G_{d+1}}\right) h + \sum_{i=1}^{d+1} \frac{\partial H^{[1]}_r}{\partial y_i} \frac{\partial H^{[1]}_r}{\partial X^G_i} h^2.$$

According to (14), the one-step approximation (12) turns out to be

$$\begin{aligned} X^G &= x - \left(\frac{\partial H^M_0(X^G, y)}{\partial y} + \frac{1}{2}\sum_{r=1}^m \frac{\partial G^G_{(r,r)}(X^G, y)}{\partial y}\right) h - \sum_{r=1}^m \frac{\partial H^M_r(X^G, y)}{\partial y} J^h_{(r)}, \\ Y^G &= y + \left(\frac{\partial H^M_0(X^G, y)}{\partial X^G} + \frac{1}{2}\sum_{r=1}^m \frac{\partial G^G_{(r,r)}(X^G, y)}{\partial X^G}\right) h + \sum_{r=1}^m \frac{\partial H^M_r(X^G, y)}{\partial X^G} J^h_{(r)}. \end{aligned} \tag{15}$$

In what follows, let $\frac{\partial S^G}{\partial y_j} := \frac{\partial S^G}{\partial y_j}(X^G, y, h)$, $\frac{\partial S^G}{\partial X^G_j} := \frac{\partial S^G}{\partial X^G_j}(X^G, y, h)$, $\frac{\partial H^{[1]}_r}{\partial y_j} := \frac{\partial H^{[1]}_r}{\partial y_j}(x, y)$, and $\frac{\partial H^{[1]}_r}{\partial x_j} := \frac{\partial H^{[1]}_r}{\partial x_j}(x, y)$ for $j = 1, \ldots, d+1$ and $r = 0, 1, \ldots, m$. Applying Taylor expansion to $\frac{\partial S^G}{\partial y_i}$ and $\frac{\partial S^G}{\partial X_i^G}$ at (x,y,h), for $i=1,\ldots,d$, we obtain

$$\begin{aligned}\frac{\partial S^G}{\partial X_i^G} &= C_2\sum_{j=1}^{d} M_{ij}x_j h + \sum_{r=1}^{m}\left(\frac{\partial H_r^{[1]}}{\partial x_i} - \sum_{j=1}^{d} M_{ij}\sigma_r^j\right) I_{(r)}^h h - \sum_{j=1}^{d} M_{ij} f_j(y) h^2 + \frac{\partial H_0^{[1]}}{\partial x_i} h^2\\ &\quad + \sum_{r=1}^{m} \frac{\partial^2 H_r^{[1]}}{\partial x_i \partial x_{d+1}} (X_{d+1}^G - x_{d+1}) I_{(r)}^h h - C_1 \sum_{r_1,r_2=1}^{m}\sum_{j=1}^{d} \frac{\partial^2 H_{r_1}^{[1]}}{\partial x_i \partial x_j}\sigma_{r_2}^j I_{(r_1)}^h I_{(r_2)}^h h\\ &\quad + \frac{1}{2}C_1 \sum_{j=1}^{d}\sum_{r=1}^{m}\sigma_r^j\left(\frac{\partial^2 H_r^{[1]}}{\partial x_i\partial x_j} + v y_i \frac{\partial^2 H_r^{[1]}}{\partial x_i \partial x_{d+1}}\right) h^2 + R_4\end{aligned}$$

and

$$\begin{aligned}\frac{\partial S^G}{\partial y_i} &= C_1\sum_{r=1}^{m}\left(\sigma_r^i I_{(r)}^h + f_i(y) h\right) + \sum_{r=1}^{m}\frac{\partial H_r^{[1]}}{\partial y_i} I_{(r)}^h h + \sum_{r=1}^{m}\sum_{j=1}^{d+1}\frac{\partial^2 H_r^{[1]}}{\partial y_i \partial x_j}(X_j^G - x_j) I_{(r)}^h h\\ &\quad + \left(\frac{\partial H_0^{[1]}}{\partial y_i} + \frac{C_1}{2}\sum_{r=1}^{m}\left[\sum_{j=1}^{d}\sigma_r^j \frac{\partial^2 H_r^{[1]}}{\partial y_i \partial x_j} + v\sigma_r^i\left(\frac{\partial H_r^{[1]}}{\partial x_{d+1}} + y_i \frac{\partial^2 H_r^{[1]}}{\partial x_{d+1}\partial y_i}\right)\right]\right) h^2 + R_5.\end{aligned}$$

Similarly,

$$\begin{aligned}\frac{\partial S^G}{\partial X_{d+1}^G} &= h + \sum_{r=1}^{m}\frac{\partial H_r^{[1]}}{\partial x_{d+1}} I_{(r)}^h h + \sum_{j=1}^{d+1}\frac{\partial^2 H_r^{[1]}}{\partial x_{d+1}\partial x_j}\left(X_j^G - x_j\right) I_{(r)}^h h + \frac{\partial H_0^{[1]}}{\partial x_{d+1}} h^2\\ &\quad + C_1\sum_{i=1}^{d}\sigma_r^i \frac{\partial^2 H_r^{[1]}}{\partial x_i \partial x_{d+1}} h^2 + C_1 \sum_{i=1}^{d} v\sigma_r^i y_i \frac{\partial^2 H_r^{[1]}}{\partial x_{d+1}^2} h^2 + R_6,\end{aligned}$$

and

$$\begin{aligned}\frac{\partial S^G}{\partial y_{d+1}} &= v\left(C_1 F(y) - \frac{1}{2}C_2\sum_{i,j=1}^{d} x_i M_{ij} x_j\right) h + vC_1\sum_{r=1}^{m}\sum_{i=1}^{d}\sigma_r^i y_i I_{(r)}^h + \sum_{r=1}^{m}\frac{\partial H_r^{[1]}}{\partial y_{d+1}} h I_{(r)}^h\\ &\quad + \sum_{i,j=1}^{d}\sum_{r=1}^{m} v\sigma_r^i M_{ij} x_j h I_{(r)}^h + \sum_{r=1}^{m}\sum_{i=1}^{d+1}\frac{\partial^2 H_r^{[1]}}{\partial y_{d+1}\partial x_i}(X_i^G - x_i) h I_{(r)}^h + \frac{\partial H_0^{[1]}}{\partial y_{d+1}} h^2\\ &\quad + \frac{C_1}{2}\sum_{i=1}^{d}\sum_{r=1}^{m}\sigma_r^i\left(v\frac{\partial H_r^{[1]}}{\partial x_i} + v^2 y_i \frac{\partial H_r^{[1]}}{\partial x_{d+1}} + \frac{\partial^2 H_r^{[1]}}{\partial x_i \partial y_{d+1}} + v y_i \frac{\partial^2 H_r^{[1]}}{\partial x_{d+1}\partial y_{d+1}}\right) h^2\\ &\quad + v\sum_{i,j=1}^{d}\left(C_2 \frac{\partial F(y)}{\partial y_i} M_{ij} x_j h^2 - \frac{C_1}{2}\sum_{r_1,r_2=1}^{m}\sigma_{r_1}^i M_{ij}\sigma_{r_2}^j h I_{(r_1)}^h I_{(r_2)}^h\right) + R_7.\end{aligned}$$

Applying Taylor expansion to $\phi(X(h),Y(h))$ and $\phi(X^G,Y^G)$ at (x,y) and taking expectations, we have

$$
\begin{aligned}
&\mathbf{E}\phi(X(h),Y(h)) - \mathbf{E}\phi(X^G,Y^G)\\
=&\sum_{i=1}^{d+1}\frac{\partial\phi(x,y)}{\partial x_i}\mathbf{E}\left(\frac{\partial S^G}{\partial y_i}-\frac{\partial S}{\partial y_i}\right)+\sum_{i=1}^{d+1}\frac{\partial\phi(x,y)}{\partial y_i}\mathbf{E}\left(\frac{\partial S}{\partial X_i}-\frac{\partial S^G}{\partial X_i^G}\right)\\
&+\frac{1}{2}\sum_{i,j=1}^{d+1}\frac{\partial^2\phi(x,y)}{\partial x_i\partial x_j}\mathbf{E}\left(\frac{\partial S}{\partial y_i}\frac{\partial S}{\partial y_j}-\frac{\partial S^G}{\partial y_i}\frac{\partial S^G}{\partial y_j}\right)\\
&+\sum_{i,j=1}^{d+1}\frac{\partial^2\phi(x,y)}{\partial y_i\partial x_j}\mathbf{E}\left(\frac{\partial S^G}{\partial X_i^G}\frac{\partial S^G}{\partial y_j}-\frac{\partial S}{\partial X_i}\frac{\partial S}{\partial y_j}\right)\\
&+\frac{1}{2}\sum_{i,j=1}^{d+1}\frac{\partial^2\phi(x,y)}{\partial y_i\partial y_j}\mathbf{E}\left(\frac{\partial S}{\partial X_i}\frac{\partial S}{\partial X_j}-\frac{\partial S^G}{\partial X_i^G}\frac{\partial S^G}{\partial X_j^G}\right)+\cdots.
\end{aligned}
\tag{16}
$$

To make the symplectic numerical approximation be of higher weak order, we choose $H_i^{[j]}$, $i=0,\dots,r$, $j=1,\dots,\tau$, such that the terms containing h and h^2 in the right-hand side of (16) vanish. Note that the coefficients of $J_{(r)}^h$ and h in $\frac{\partial S^G}{\partial X_i^G}$ and $\frac{\partial S^G}{\partial y_i}$ are the same as those in $\frac{\partial S}{\partial X_i}$ and $\frac{\partial S}{\partial y_i}$ with $i=1,\dots,d+1$, respectively. Then we get

$$
\mathbf{E}\left(\frac{\partial S^G}{\partial X_{d+1}^G}\frac{\partial S^G}{\partial y_{d+1}}-\frac{\partial S}{\partial X_{d+1}}\frac{\partial S}{\partial y_{d+1}}\right)=\sum_{r=1}^{m}\sum_{i=1}^{d}vC_1\sigma_r^i y_i\frac{\partial H_r^{[1]}}{\partial x_{d+1}}h^2+h^3e_1(x,y),
$$

where $e_1(x,y)$ denotes the coefficient of the term containing h^3 and can be calculated based on the expression of the partial derivatives of S^G and S, as do the other remainder terms e_l, $l=2,\dots,7$, in what follows. Thus, we choose $\frac{\partial H_r^{[1]}}{\partial x_{d+1}}=0$ for $r=1,\dots,m$. Substituting $\frac{\partial H_r^{[1]}}{\partial x_{d+1}}=0$ into $\frac{\partial S^G}{\partial X_{d+1}^G}$, we have

$$
\mathbf{E}\left(\frac{\partial S^G}{\partial X_{d+1}^G}-\frac{\partial S}{\partial X_{d+1}}\right)=\frac{\partial H_0^{[1]}}{\partial x_{d+1}}h^2+\mathbf{E}(R_6)=\frac{\partial H_0^{[1]}}{\partial x_{d+1}}h^2+h^3e_2(x,y),
$$

which leads us to make $\frac{\partial H_0^{[1]}}{\partial x_{d+1}}=0$. In the same way, using $\frac{\partial H_r^{[1]}}{\partial x_{d+1}}=0$ for $r=0,1,\dots,m$, we derive

$$
\mathbf{E}\left(\frac{\partial S}{\partial y_i}\frac{\partial S}{\partial y_j}-\frac{\partial S^G}{\partial y_i}\frac{\partial S^G}{\partial y_j}\right)=C_1\sum_{r=1}^{m}\left(vC_1\sigma_r^i\sigma_r^j-\sigma_r^i\frac{\partial H_r^{[1]}}{\partial y_j}-\sigma_r^j\frac{\partial H_r^{[1]}}{\partial y_i}\right)h^2+h^3e_3(x,y)
$$

and

$$
\mathbf{E}\left(\frac{\partial S}{\partial y_i}\frac{\partial S}{\partial X_j}-\frac{\partial S^G}{\partial y_i}\frac{\partial S^G}{\partial X_j^G}\right)=C_1\sum_{r=1}^{m}\sigma_r^i\left(\frac{1}{2}\sum_{k=1}^{d}M_{jk}\sigma_r^k-\frac{\partial H_r^{[1]}}{\partial x_j}\right)h^2+h^3e_4(x,y)
$$

with $i,j=1,\dots,d$, and hence choose

$$
\frac{\partial H_r^{[1]}}{\partial y_i}=\frac{1}{2}vC_1\sigma_r^i,\quad \frac{\partial H_r^{[1]}}{\partial x_i}=\frac{1}{2}\sum_{j=1}^{d}M_{ij}\sigma_r^j,\quad r=1,\dots,m.
$$

The last term in (16) is of order 3 due to the following estimate:

$$\mathbf{E}\left(\frac{\partial S}{\partial X_i}\frac{\partial S}{\partial X_j}-\frac{\partial S^G}{\partial X_i^G}\frac{\partial S^G}{\partial X_j^G}\right)=h^3e_5(x,y),\quad i,j=1,\ldots,d+1.$$

Since both $\frac{\partial H_r^{[1]}}{\partial y_i}$ and $\frac{\partial H_r^{[1]}}{\partial x_i}$, with $r=0,1,\ldots,m$, are independent of x_i and y_i, we have

$$\mathbf{E}\left(\frac{\partial S}{\partial y_i}-\frac{\partial S^G}{\partial y_i}\right)=\left(\frac{1}{2}\sum_{j,k=1}^{d}\frac{\partial^2F(y)}{\partial y_i\partial y_j}M_{jk}x_k+\frac{1}{2}vC_1f_i(y)-\frac{\partial H_0^{[1]}}{\partial y_i}\right)h^2+h^3e_6(x,y),$$

$$\mathbf{E}\left(\frac{\partial S}{\partial X_i}-\frac{\partial S^G}{\partial X_i^G}\right)=\left(\frac{1}{2}\sum_{j=1}^{d}M_{ij}f_j(y)-\frac{1}{2}\sum_{j=1}^{d}vC_2M_{ij}x_j-\frac{\partial H_0^{[1]}}{\partial x_i}\right)h^2+h^3e_7(x,y)$$

for $i=1,\ldots,d$. We choose $H_0^{[1]}$ such that the above terms containing h^2 vanish, i.e.,

$$\frac{\partial H_0^{[1]}}{\partial y_i}=\frac{1}{2}\sum_{j,k=1}^{d}\frac{\partial^2F(y)}{\partial y_i\partial y_j}M_{jk}x_k+\frac{1}{2}vC_1f_i(y),$$

$$\frac{\partial H_0^{[1]}}{\partial x_i}=\frac{1}{2}\sum_{j=1}^{d}M_{ij}\left(f_j(y)-vC_2x_j\right).$$

Substituting the above results on the partial derivatives of $H_r^{[1]}$, $r=0,1,\ldots,m$, into (15), we have the following scheme of (9):

$$\begin{aligned}X_i^G&=x_i-\sum_{r=1}^{m}e^{vt_n}\sigma_r^iI_{(r)}^h-e^{vt_n}f_i(y)h-\frac{1}{2}\sum_{r=1}^{m}ve^{vt_n}\sigma_r^ihI_{(r)}^h\\&\quad-\frac{1}{2}\sum_{j,k=1}^{d}\frac{\partial^2F(y)}{\partial y_i\partial y_j}M_{jk}X_k^Gh^2-\frac{1}{2}ve^{vt_n}f_i(y)h^2,\\Y_i^G&=y_i+\sum_{j=1}^{d}e^{-vt_n}M_{ij}X_j^Gh+\frac{1}{2}\sum_{r=1}^{m}\sum_{j=1}^{d}M_{ij}\sigma_r^jI_{(r)}^hh\\&\quad+\frac{1}{2}\sum_{j=1}^{d}M_{ij}\left(f_j(y)-ve^{-vt_n}X_j^G\right)h^2,\end{aligned}\tag{17}$$

which is started at time $t_n=nh$ for $n=1,\ldots,N=T/h$. That is, $x_i=X_i(t_n)$, $y_i=Y_i(t_n)$ for $i=1,\ldots,d$, and $y_{d+1}=t_n$.

To transform scheme (17) into an equivalent scheme of (1), we denote $P_i^h[n]:=e^{-vt_n}x_i$, $Q_i^h[n]:=y_i$, $P_i^h[n+1]:=e^{-vt_{n+1}}X_i^G$, and $Q_i^h[n+1]:=Y_i^G$ for $i=1,\ldots,d$. Based on the transformation between two phase spaces of (1) and (5), we get

$$\begin{aligned}P^h[n+1]&=e^{-vh}P^h[n]-\frac{h^2}{2}\nabla^2F(Q^h[n])MP^h[n+1]-h\Big(1+\frac{vh}{2}\Big)e^{-vh}f(Q^h[n])\\&\quad-\Big(1+\frac{vh}{2}\Big)e^{-vh}\sigma\Delta_{n+1}W,\\Q^h[n+1]&=Q^h[n]+h\left(1-\frac{vh}{2}\right)e^{vh}MP^h[n+1]+\frac{h^2}{2}Mf(Q^h[n])+\frac{h}{2}M\sigma\Delta_{n+1}W,\end{aligned}\tag{18}$$

where $\sigma = (\sigma_1, \ldots, \sigma_r)$ and $\Delta_{n+1}W = W(t_{n+1}) - W(t_n)$. Notice that $\Delta_n W$ can be simulated by $\xi^n\sqrt{h}$ with $\xi^n = (\xi_1^n, \ldots, \xi_d^n)^\top$ being an $\mathcal{F}_{t_n}$-adapted d-dimensional normal distributed random vector.

Remark 3.1. The proposed scheme (18) also has exponentially dissipative phase volume. More precisely, denoting $D(q) = (I_d + \frac{h^2}{2}\nabla^2 F(q)M)^{-1}$, the determinant of Jacobian matrix

$$\begin{vmatrix} \frac{\partial P^h[1]}{\partial p} & \frac{\partial P^h[1]}{\partial q} \\ \frac{\partial Q^h[1]}{\partial p} & \frac{\partial Q^h[1]}{\partial q} \end{vmatrix} = \begin{vmatrix} e^{-vh}D(q) & \frac{\partial P^h[1]}{\partial q} \\ h(1-\frac{vh}{2})MD(q) & D(q)^{-\top} + h(1-\frac{vh}{2})e^{vh}M\frac{\partial P^h[1]}{\partial q} \end{vmatrix}$$
$$= |e^{-vh}I_d||D(q)||D(q)^{-\top}| = e^{-vhd}.$$

Furthermore,

$$\begin{vmatrix} \frac{\partial P^h[n]}{\partial p} & \frac{\partial P^h[n]}{\partial q} \\ \frac{\partial Q^h[n]}{\partial p} & \frac{\partial Q^h[n]}{\partial q} \end{vmatrix} = e^{-vt_n d}.$$

3.2. Conformal symplectic structure and ergodicity. In this subsection, we prove the conformal symplecticity of the proposed scheme (18) as well as its ergodicity.

THEOREM 3.2. *The proposed scheme* (18) *preserves conformal symplectic structure, i.e,*

$$dP^h[n+1] \wedge dQ^h[n+1] = e^{-vh} dP^h[n] \wedge dQ^h[n].$$

Proof. Based on (18), we obtain

$$\begin{aligned} &dP^h[n+1] \wedge dQ^h[n+1] \\ &= dP^h[n+1] \wedge dQ^h[n] + \frac{1}{2}h^2 dP^h[n+1] \wedge M\nabla^2 F dQ^h[n] \\ &= e^{-vh} dP^h[n] \wedge dQ^h[n] - \frac{h^2}{2} d\left[\nabla^2 F(Q^h[n])MP^h[n+1]\right] \wedge dQ^h[n] \\ &\quad + \frac{h^2}{2} dP^h[n+1] \wedge M\nabla^2 F(Q^h[n]) dQ^h[n]. \end{aligned}$$

Denote $\tilde{P}^h := MP^h[n+1]$; then the second term becomes

$$\begin{aligned} &\frac{h^2}{2} d\left[\nabla^2 F(Q^h[n])\tilde{P}^h\right] \wedge dQ^h[n] \\ =&\frac{h^2}{2}\sum_{i,j,l=1}^{d} \frac{\partial^3 F}{\partial q_i \partial q_j \partial q_l}\tilde{P}_j^h dQ_l^h[n] \wedge dQ_i^h[n] - \frac{h^2}{2}\nabla^2 F(Q^h[n])MdP^h[n+1] \wedge dQ^h[n]. \end{aligned}$$

Since matrix M is symmetric and the first term in the right-hand side of the above equation vanishes, we finally get

$$dP^h[n+1] \wedge dQ^h[n+1] = e^{-vh} dP^h[n] \wedge dQ^h[n]. \qquad \square$$

To show the ergodicity of (18), we first introduce the following conditions which are sufficient to ensure the existence and uniqueness of the invariant measure (see [13] and references therein). Then we will show that these conditions are exactly satisfied by the proposed scheme.

Condition 3.3. *The Markov chain* $Z_n := (P^h[n]^\top, Q^h[n]^\top)^\top$ *with* $Z_0 = z$ *satisfies*

(i) *for any* $\gamma \geq 1$*, there exists* $C_2 = C(\gamma) > 0$ *which is independent of* h*, such that* $\mathbf{E}\|Z_1\|^\gamma \leq C_2(1 + \|z\|^\gamma)$ *for all* $z \in \mathbb{R}^{2d}$*;*

(ii) *there exist* $C_1 > 0$ *and* $\epsilon > 0$ *which are independent of* h*, such that* $\mathbf{E}\|Z(h) - Z_1\|^2 \leq C_1(1 + \|z\|^2)h^{\epsilon+2}$ *for all* $z \in \mathbb{R}^{2d}$*, where* $Z(h) = (P(h)^\top, Q(h)^\top)^\top$*.*

Condition 3.4. *For some fixed compact set* $G \in \mathcal{B}(\mathbb{R}^{2d})$ *with* $\mathcal{B}(\mathbb{R}^{2d})$ *denoting the Borel* σ*-algebra on* $\mathbb{R}^{2d}$*, the Markov chain* $Z_n := (P^h[n]^\top, Q^h[n]^\top)^\top \in \mathcal{F}_{t_n}$ *with transition kernel* $\mathcal{P}_n(z, A)$ *satisfies*

(i) *for some* $z^* \in \mathrm{int}(G)$ *and for any* $\delta > 0$*, there exists a positive integer* n *such that*

$$\mathcal{P}_n(z, B_\delta(z^*)) > 0 \quad \forall\, y \in G,$$

where $B_\delta(z^*)$ *denotes the open ball of radius* δ *centered at* z^**;*

(ii) *for any* $n \in \mathbb{N}$*, the transition kernel* $\mathcal{P}_n(z, A)$ *possesses a density* $\rho_n(z, w)$ *which is jointly continuous in* $(z, w) \in G \times G$*.*

Theorem 3.5 (see [13, Theorem 7.3]). *For some* $K \in \mathbb{N}$*, if Conditions* 3.3 *and* 3.4 *are satisfied by a Markov chain* Z_n *when sampled at rate* K*, that is, these conditions hold for the chain* $\tilde{Z}_n := Z_{nK}$*, then* Z_n *has a unique invariant measure.*

Theorem 3.6. *Assume that the vector field* f *is globally Lipschitz. The solution* $(P^h[n], Q^h[n])$ *of* (18)*, which is an* $\mathcal{F}_{t_n}$*-adapted Markov chain, satisfies Condition* 3.3 *and hence admits an invariant measure* μ_h *on* $\mathbb{R}^{2d}$*. In addition, if* f *is a linear function, then Condition* 3.4 *is also satisfied and the invariant measure is unique, that is,* (18) *is ergodic.*

Proof. Step 1. We first show that scheme (18) satisfies Condition 3.3. Denote $Z(t) = (P(t)^\top, Q(t)^\top)^\top \in \mathbb{R}^{2d}$, $Z_n = (P^h[n]^\top, Q^h[n]^\top)^\top \in \mathbb{R}^{2d}$, $\sigma = (\sigma_1, \ldots, \sigma_r) \in \mathbb{R}^{d\times r}$, $W = (W_1, \ldots, W_r)^\top \in \mathbb{R}^r$, and $D(q) = (I_d + \frac{h^2}{2}\nabla^2 F(q)M)^{-1}$. We rewrite (18) as

$$\begin{aligned} P^h[1] &= D(q)\left(e^{-vh}p - \Big(1 + \frac{vh}{2}\Big)e^{-vh}\sigma\Delta_1 W - h\Big(1 + \frac{vh}{2}\Big)e^{-vh}f(q)\right), \\ Q^h[1] &= q + h\Big(1 - \frac{vh}{2}\Big)e^{vh}MP^h[1] + \frac{h^2}{2}Mf(q) + \frac{h}{2}M\sigma\Delta_1 W \end{aligned} \tag{19}$$

with $z := (P_0^\top, Q_0^\top)^\top = (p^\top, q^\top)^\top$, which yields

$$\begin{aligned} \mathbf{E}\|P^h[1]\|^\gamma + \mathbf{E}\|Q^h[1]\|^\gamma &\leq C(1 + \|p\|^\gamma + \|q\|^\gamma) + C(1 + \|q\|^\gamma + \mathbf{E}\|P^h[1]\|^\gamma) \\ &\leq C(1 + \|p\|^\gamma + \|q\|^\gamma) \end{aligned} \tag{20}$$

based on the fact that vector field f is globally Lipschitz, the matrix $I + \frac{h^2}{2}\nabla^2 F(q)M$ is positive definite, and $\|D(q)\| \leq 1$ for any $q \in \mathbb{R}^d$ and $h \in (0, 1)$. As the norm $\|Z_1\| = (\|P^h[1]\|^2 + \|Q^h[1]\|^2)^{\frac{1}{2}}$ is equivalent to the norm $(\|P^h[1]\|^\gamma + \|Q^h[1]\|^\gamma)^{\frac{1}{\gamma}}$, Condition 3.3(i) holds.

Rewrite (1) into the following mild solution form:

$$\begin{aligned} P(h) &= p - \int_0^h e^{-v(h-s)}f(Q(s))ds - \int_0^h e^{-v(h-s)}\sigma dW(s), \\ Q(h) &= q + \int_0^h MP(s)ds \end{aligned}$$

with $P(0) = p$ and $Q(0) = q$. Based on (18), we have

$$\begin{aligned}P(h) - P^h[1] =& \left[h\Big(1 + \frac{vh}{2}\Big)e^{-vh}f(q) + \frac{h^2}{2}\nabla^2 F(q)MP^h[1] - \int_0^h e^{-v(h-s)}f(Q(s))ds\right]\\&+ \left[\Big(1 + \frac{vh}{2}\Big)e^{-vh}\sigma\Delta_1 W - \int_0^h e^{-v(h-s)}\sigma dW(s)\right]\\=:& \, I + II,\\Q(h) - Q^h[1] =& \left[\int_0^h MP(s)ds - h\Big(1 - \frac{vh}{2}\Big)e^{vh}MP^h[1]\right] - \left[\frac{h}{2}M\sigma\Delta_1 W + \frac{h^2}{2}Mf(q)\right]\\=:& \, III + IV.\end{aligned}$$

Now we estimate terms I, II, III, and IV, respectively:

$$\begin{aligned}\mathbf{E}\|I\|^2 \leq& C\mathbf{E}\left\|\frac{h^2}{2}\nabla^2 F(q)P^h[1]\right\|^2 + C\mathbf{E}\left\|\int_0^h e^{-v(h-s)}\left(f(Q(s)) - f(q)\right)ds\right\|^2\\&+ C\left\|\int_0^h e^{-v(h-s)}dsf(q) - h\Big(1 + \frac{vh}{2}\Big)e^{-vh}f(q)\right\|^2\\\leq& Ch^4(1 + \|z\|^2) + C\int_0^h e^{-2v(h-s)}ds\int_0^h \left(\|Q(s) - Q^h[1]\|^2 + \|Q^h[1] - q\|^2\right)ds\\&+ C\left(\frac{1 - e^{-vh}}{v} - h\Big(1 + \frac{vh}{2}\Big)e^{-vh}\right)^2(1 + \|q\|^2)\\\leq& Ch^3(1 + \|z\|^2) + C\int_0^h \|Q(s) - Q^h[1]\|^2 ds, \qquad (21)\end{aligned}$$

where in the last step we have used (20). For the term II, based on the Itô isometry,

$$\mathbf{E}\|II\|^2 \leq \int_0^h \left(\Big(1 + \frac{vh}{2}\Big)e^{-vh} - e^{-v(h-s)}\right)^2 ds\mathrm{Tr}\left(\sigma\sigma^\top\right) \leq Ch^3. \qquad (22)$$

Similarly, we have

(23)

$$\begin{aligned}\mathbf{E}\|III\|^2 \leq& C\mathbf{E}\left\|\int_0^h M\left(P(s) - P^h[1]\right)ds\right\|^2 + C\mathbf{E}\left\|h\left(1 - \Big(1 - \frac{vh}{2}\Big)e^{vh}\right)MP^h[1]\right\|^2\\\leq& C\int_0^h \|P(s) - P^h[1]\|^2 ds + Ch^4(1 + \|z\|^2)\end{aligned}$$

and

$$\mathbf{E}\|IV\|^2 \leq Ch^3(1 + \|q\|^2). \qquad (24)$$

From (21)–(24), we conclude

$$\mathbf{E}\|Z(h) - Z_1\|^2 \leq C\int_0^h \mathbf{E}\|Z(s) - Z_1\|^2 ds + Ch^3(1 + \|z\|^2),$$

which together with Gronwall's inequality yields Condition 3.3(ii) with $\epsilon = 1$. In this case, there exist real numbers $\tilde{\alpha} \in (0,1)$ and $\tilde{\beta} \in [0,\infty)$ such that $\mathbf{E}[V(Z_{n+1})|\mathcal{F}_{t_n}] \leq \tilde{\alpha} V(Z_n) + \tilde{\beta}$ for $V(z) = \frac{1}{2}\|p\|^2 + F(q) + \frac{v}{2}p^\top q + \frac{v^2}{4}\|q\|^2 + 1$ with $z = (p^\top, q^\top)^\top$ (see Theorem 7.2 in [13]). Hence,

$$\mathbf{E}[V(Z_{n+1})] \leq \tilde{\alpha}\mathbf{E}[V(Z_n)] + \tilde{\beta} \leq \tilde{\alpha}^{n+1}\mathbf{E}[V(Z_0)] + \tilde{\beta}\frac{1-\tilde{\alpha}^n}{1-\tilde{\alpha}} \leq C(Z_0),$$

which induces the existence of invariant measures (see Proposition 7.10 in [9]).

Step 2. We now consider the chain Z_{2n} sampled at rate $K = 2$ and verify Condition 3.4 when f is linear with a constant $C_f := \nabla f = \nabla^2 F$. Let $G := \{(P^\top, Q^\top)^\top \in \mathbb{R}^{2d} : Q = 0, \|P\| \leq 1\}$, which is a compact set. For any $z = (p^\top, 0)^\top \in G$ and $w = (w_1^\top, w_2^\top)^\top \in B$ with $B \in \mathcal{B}(\mathbb{R}^{2d})$, we aim to show that $\Delta_1 W$ and $\Delta_2 W$ can be properly chosen to ensure that $P^h[2] = w_1$ and $Q^h[2] = w_2$ starting from $(P_0^\top, Q_0^\top)^\top = z$. Denoting $L_h = h(1-\frac{vh}{2})e^{vh}M$, from (18), we have

(25)
$$w_1 = e^{-vh}P^h[1] - \frac{h^2}{2}C_f M w_1 - h\Big(1+\frac{vh}{2}\Big)e^{-vh}f(Q^h[1]) - \Big(1+\frac{vh}{2}\Big)e^{-vh}\sigma\Delta_2 W,$$

(26)
$$\begin{aligned} w_2 &= Q^h[1] + L_h w_1 + \frac{h^2}{2}Mf(Q^h[1]) + \frac{h}{2}M\sigma\Delta_2 W \\ &= Q^h[1] + L_h w_1 + \frac{h}{2}\Big(1+\frac{vh}{2}\Big)^{-1}e^{vh}M\left(e^{-vh}P^h[1] - w_1 - \frac{h^2}{2}C_f M w_1\right), \end{aligned}$$

(27)
$$P^h[1] = e^{-vh}p - \frac{h^2}{2}C_f M P^h[1] - h\Big(1+\frac{vh}{2}\Big)e^{-vh}f(0) - \Big(1+\frac{vh}{2}\Big)e^{-vh}\sigma\Delta_1 W,$$

(28)
$$\begin{aligned} Q^h[1] &= L_h P^h[1] + \frac{h^2}{2}Mf(0) + \frac{h}{2}M\sigma\Delta_1 W \\ &= L_h P^h[1] + \frac{h}{2}\Big(1+\frac{vh}{2}\Big)^{-1}e^{vh}M\left(e^{-vh}p - P^h[1] - \frac{h^2}{2}C_f M P^h[1]\right). \end{aligned}$$

Notice that (26) and (28) form a linear system, from which we can get the solution $P^h[1]$ and $Q^h[1]$ based on the positive definite coefficient matrix. Then $\Delta_2 W$ and $\Delta_1 W$ can be uniquely determined by (25) and (27), respectively. Condition 3.4(i) is then ensured according to the property that Brownian motions hit a cylinder set with positive probability. For Condition 3.4(ii), from (19), we can find out that $P^h[1]$ has a C^∞ density based on the facts that $\Delta_1 W$ has a C^∞ density, σ is full rank, and $D(q)$ is positive definite for any $q \in \mathbb{R}^d$. Thus, $Q^h[1]$ also has a C^∞ density, and Theorem 3.5 is applied to complete the proof. ∎

Remark 3.7. For the nonlinear case, the uniqueness of the invariant measure is unsolved since both equations in (18) contain the same noise, which is totally different from the continuous case and brings essential difficulties when showing the irreducible property. For higher k and k', following the same procedure as for the case $k = k' = 1$ (see also [1]), choosing undetermined functions such that the error in (13) is of higher order, we can also get higher weak order symplectic schemes for (5), which turn out to be high weak order conformal symplectic schemes for the original system (1) based on the inverse transformation $(X, Y) \mapsto (P, Q)$. It is worth mentioning that the solvability of undetermined functions, as well as the ergodicity of the schemes, is unknown for high order cases, as far as we know.

4. Approximation error. In this section, we consider the weak convergence order of (18) by investigating the local convergence error first. Furthermore, based on the local convergence error and the hypoelliptic setting (2), we can also get the approximation error of the ergodic limit. Denote the exact solution of (1) and the numerical solution by $Z(t) = (P(t)^\top, Q(t)^\top)^\top$ and $Z_n = (P^h[n]^\top, Q^h[n]^\top)^\top$, respectively. The next theorem gives that the moments of (1) are uniformly bounded, and its proof follows the same procedure as that of Lemma 3.3 in [13].

THEOREM 4.1. *Let Assumption* 2.1 *hold. Then for any $k \in \mathbb{N}_+$, the kth moments of $P(t)$ and $Q(t)$ are uniformly bounded with respect to $t \in \mathbb{R}_+$.*

Before proving the main convergence theorem, we first show the boundedness of the numerical solution to (18) in the following theorem.

THEOREM 4.2. *Assume that the coefficient f of* (1) *is globally Lipschitz and satisfies the linear growth condition, i.e.,*

$$\|f(u) - f(w)\| \le L\|u - w\|, \quad \|f(u)\| \le C_f(1 + \|u\|) \tag{29}$$

for some constants $L > 0$ and $C_f \ge 0$, and any $u, w \in \mathbb{R}^d$. Then there exists a positive constant h_0 such that for any $h \le h_0$, it holds that

$$\sup_{n\in\{1,\dots,N\}} \mathbf{E}\left[\|P^h[n]\|^k + \|Q^h[n]\|^k\right] < \infty.$$

Proof. For any fixed initial value $z = (p^\top, q^\top)^\top$, random variable $\xi := \xi^1$, and h, we have based on (18) that

$$\|P^h[1] - p\| \le |e^{-vh} - 1|\|p\| + h\Big(1 + \frac{vh}{2}\Big)\|f(q)\| + \sqrt{h}\Big(1 + \frac{vh}{2}\Big)\|\sigma\xi\| \\ + \frac{h^2}{2}\|\nabla^2 F(q)\|\|M\|\|p\| + \frac{h^2}{2}\|\nabla^2 F(q)\|\|M\|\|P^h[1] - p\|.$$

Denote $C_v := 1 + \frac{vh}{2}$. Using the global Lipschitz condition and mean value theorem, there exists some $\theta \in (0, 1)$ such that

$$\begin{aligned}\|P^h[1] - p\| &\le | - vhe^{-v\theta h}|\|p\| + hC_f(1 + \|z\|) + \sqrt{h}C_v\|\sigma\xi\| \\ &\quad + \frac{h^2}{2}L\|M\|\|z\| + \frac{h^2}{2}L\|M\|\|P^h[1] - p\| \\ &\le C(1 + \|z\|)(\|\xi\|\sqrt{h} + h) + L\|M\|\|P^h[1] - p\|\frac{h^2}{2}.\end{aligned}$$

It is obvious that there exists a positive constant h_0 such that for any $h \le h_0$,

$$L\|M\|\frac{h^2}{2} \le \frac{1}{2}.$$

It then yields

$$\|P^h[1] - p\| \le 2C(1 + \|z\|)(\|\xi\|\sqrt{h} + h).$$

On the other hand, for $h \le h_0$, we have

$$
\begin{aligned}
&\|\mathbf{E}(P^h[1]-p)\| \\
&\le \left\|(e^{-vh}-1)p - \frac{h^2}{2}\nabla^2 F(q)Mp - hC_v e^{-vh} f(q)\right\| + \left\|\frac{h^2}{2}\nabla^2 F(q)M\mathbf{E}(P^h[1]-p)\right\| \\
&\le vh\|p\| + hL\|M\|\|p\| + hC_f C_v(1+\|z\|) + \frac{h^2}{2}L\|M\|\|\mathbf{E}(P^h[1]-p)\|,
\end{aligned}
$$

which leads to

$$\|\mathbf{E}(P^h[1]-p)\| \le C(1+\|z\|)h.$$

Based on the estimate of $P^h[1]-p$, similarly, we have

$$\|Q^h[1]-q\| \le C(1+\|z\|)(\|\xi\|\sqrt{h}+h), \quad \|\mathbf{E}(Q^h[1]-q)\| \le C(1+\|z\|)h.$$

We can conclude that, for $Z_1 = (P^h[1]^\top, Q^h[1]^\top)^\top$,

$$\|Z_1 - z\| \le C(\|\xi\| + \sqrt{h})(1+\|z\|)\sqrt{h} \le C(\|\xi\|+1)(1+\|z\|)\sqrt{h}. \tag{30}$$

Thus, we complete the proof according to Lemma 9.1 in [15]. ∎

Based on the above preliminaries, our result concerning the weak convergence order of the proposed scheme is as follows.

THEOREM 4.3. *Under the assumptions in Theorem* 4.2, *the proposed scheme* (18) *is of weak order* 2. *More precisely,*

$$\left|\mathbf{E}\psi\left(P(T),Q(T)\right) - \mathbf{E}\psi\left(P^h[N],Q^h[N]\right)\right| = O(h^2)$$

for all $\psi \in C_P^6(\mathbb{R}^{2d},\mathbb{R})$ *and* $T = Nh$.

Proof. Without loss of generality, we consider the case of $d=1$. Based on Itô's formula and Theorems 4.1 and 4.2, we obtain

$$
\begin{aligned}
P(h) &= p - \int_0^h \left(f(Q(s)) + vP(s)\right)ds - \sum_{r=1}^m \int_0^h \sigma_r dW_r(s) \\
&= p - \int_0^h \left(f(q) + \int_0^s \nabla^2 F(Q(\theta))MP(\theta)d\theta\right)ds - \sum_{r=1}^m \int_0^h \sigma_r dW_r(s) \\
&\quad - v\int_0^h \left(p - \int_0^s f(Q(\theta))d\theta - \int_0^s vP(\theta)d\theta - \sum_{r=1}^m \sigma_r dW_r(\theta)\right)ds,
\end{aligned}
$$

which leads to

$$
\begin{aligned}
P(h) &= p - f(q)h - vph - \frac{1}{2}\nabla^2 F(q)Mph^2 - \sum_{r=1}^m \int_0^h \sigma_r dW_r(s) \\
&\quad + \frac{1}{2}vf(q)h^2 + \frac{1}{2}v^2ph^2 + v\sum_{r=1}^m \int_0^h \int_0^s \sigma_r dW_r(\theta)ds + \delta_1,
\end{aligned} \tag{31}
$$

where $\mathbf{E}\|\delta_1\| = O(h^3)$ and $\mathbf{E}\|\delta_1\|^2 = O(h^5)$. Analogously, it also holds that

$$
\begin{aligned}
Q(h) &= q + \int_0^h M\left(p - \int_0^s f(Q(\theta))d\theta - v\int_0^s P(\theta)d\theta - \sum_{r=1}^m \int_0^s \sigma_r dW_r(\theta)\right)ds \\
&= q + Mph - \frac{1}{2}f(q)h^2 - \frac{1}{2}vMph^2 - \sum_{r=1}^m M\sigma_r \int_0^h \int_0^s dW_r(\theta)ds + \delta_2
\end{aligned} \tag{32}
$$

with $\mathbf{E}\|\delta_2\| = O(h^3)$ and $\mathbf{E}\|\delta_2\|^2 = O(h^5)$. For (18), applying Taylor expansion to $P^h[1]$ and $Q^h[1]$ at (p,q), we obtain

$$\begin{aligned} P^h[1] &= p - f(q)h - vph - \frac{1}{2}\nabla^2 F(q)Mph^2 - \sum_{r=1}^m \sigma_r \Delta_1 W \\ &\quad + \frac{1}{2}vf(q)h^2 + \frac{1}{2}v^2ph^2 + \frac{1}{2}v\sum_{r=1}^m \sigma_r h\Delta_1 W + \delta_3, \end{aligned} \tag{33}$$

$$Q^h[1] = q + Mph - \frac{1}{2}f(q)h^2 - \frac{1}{2}vMph^2 - \frac{1}{2}\sum_{r=1}^m M\sigma_r h\Delta_1 W + \delta_4, \tag{34}$$

where $\mathbf{E}\|\delta_i\| = O(h^3)$ and $\mathbf{E}\|\delta_i\|^2 = O(h^5)$ with $i = 3, 4$. Due to (31) and (33), we know that

$$P(h) - P^h[1] = v\sum_{r=1}^m \sigma_r \left(\int_0^h \int_0^s dW_r(\theta)ds - \frac{1}{2}h\Delta_1 W\right) + (\delta_1 - \delta_3),$$

and thus $\|\mathbf{E}(P(h) - P^h[1])\| = O(h^3)$. Similarly, based on (32) and (34), we have $\|\mathbf{E}(Q(h) - Q^h[1])\| = O(h^3)$. For $i = 2, 3, 4, 5$, we obtain

$$\begin{aligned} \left\|\mathbf{E}\left[(P(h)-p)^i - (P^h[1]-p)^i\right]\right\| &\le Ch^3 + O(h^4), \\ \left\|\mathbf{E}\left[(Q(h)-q)^i - (Q^h[1]-q)^i\right]\right\| &\le Ch^3 + O(h^4). \end{aligned}$$

Moreover, for $i_1 + i_2 = 2, 3, 4, 5$ and $i_1 \ge 1$,

$$\left\|\mathbf{E}\left[(P(h)-p)^{i_1}(Q(h)-q)^{i_2} - (P^h[1]-p)^{i_1}(Q^h[1]-q)^{i_2}\right]\right\| \le Ch^3 + O(h^4).$$

By Taylor expansion and the mean value theorem, we obtain

$$\begin{aligned} &\left|\mathbf{E}\left[\psi(P(h),Q(h)) - \psi(P^h[1],Q^h[1])\right]\right| \\ &\le \left|\frac{\partial\psi}{\partial p}(p,q)\right| \left\|\mathbf{E}(P(h) - P^h[1])\right\| + \left|\frac{\partial\psi}{\partial q}(p,q)\right| \left\|\mathbf{E}(Q(h) - Q^h[1])\right\| \\ &\quad + \sum_{j=2}^5\sum_{i=0}^j \left|\frac{\partial^j\psi(p,q)}{\partial p^i\partial q^{j-i}}\right| \left\|\mathbf{E}[(P(h)-p)^i(Q(h)-q)^{j-i} - (P^h[1]-p)^i(Q^h[1]-q)^{j-i}]\right\| \\ &\quad + \sum_{i=0}^6 \mathbf{E}\left(\left|\frac{\partial^6\psi(p+\theta_1 P(h), q+\theta_1 Q(h))}{\partial p^i\partial q^{6-i}}\right| \left\|(P(h)-p)^i(Q(h)-q)^{6-i}\right\|\right) \\ &\quad + \sum_{i=0}^6 \mathbf{E}\left(\left|\frac{\partial^6\psi(p+\theta_2 P^h[1], q+\theta_2 Q^h[1])}{\partial p^i\partial q^{6-i}}\right| \left\|(P^h[1]-p)^i(Q^h[1]-q)^{6-i}\right\|\right) \end{aligned} \tag{35}$$

with constants $0 \le \theta_1 \le 1$ and $0 \le \theta_2 \le 1$. Here, based on (31)–(34) and Theorems 4.1 and 4.3, we derive

$$\begin{aligned} &\mathbf{E}\left(\left|\frac{\partial^6\psi(p+\theta_1 P(h), q+\theta_1 Q(h))}{\partial p^i\partial q^{6-i}}\right| \left\|(P(h)-p)^i(Q(h)-q)^{6-i}\right\|\right) \\ &\le C\left(\mathbf{E}\left\|(P(h)-p)^{2i}(Q(h)-q)^{12-2i}\right\|\right)^{\frac{1}{2}} \le Ch^{6-\frac{i}{2}}, \end{aligned}$$

where we also use the fact that $\psi \in C_P^6(\mathbb{R}^{2d}, \mathbb{R})$. Analogously,

$$\mathbf{E}\left(\left|\frac{\partial^6\psi(p+\theta_2 P^h[1], q+\theta_2 Q^h[1])}{\partial p^i \partial q^{6-i}}\right| \left\|(P^h[1]-p)^i(Q^h[1]-q)^{6-i}\right\|\right) = O(h^{6-\frac{i}{2}})$$

for $0 \le i \le 6$. Finally, we deduce

$$\left|\mathbf{E}\psi(P(h), Q(h)) - \mathbf{E}\psi(P^h[1], Q^h[1])\right| \le O(h^3), \tag{36}$$

which, together with Theorem 9.1 in [15], yields global weak order two for the proposed scheme (18). ∎

According to the above theorem and the condition (2), we can get that the temporal average of the proposed scheme (18) is a proper approximation of the ergodic limit $\int_{\mathbb{R}^{2d}} \psi d\mu$.

THEOREM 4.4. *For any $\psi \in C_b^6(\mathbb{R}^{2d}, \mathbb{R})$ and any initial values, under assumptions in Theorems* 3.6 *and* 4.3*, the scheme* (18) *satisfies*

$$\left|\frac{1}{N}\sum_{n=1}^{N} \mathbf{E}\psi(P^h[n], Q^h[n]) - \int_{\mathbb{R}^{2d}} \psi d\mu\right| \le C\left(h^2 + \frac{1}{T}\right).$$

In fact, one can check that the assumptions in Theorem 5.6 in [14] are satisfied by (18) and thus deduce this result.

5. Numerical experiments.

The first example (section 5.1) tests the numerical approximation by simulating a linear stochastic Langevin equation. In section 5.2, numerical tests of the conformal symplectic scheme for the nonlinear case are presented. In all of the experiments, the expectation is approximated by taking the average over 5000 realizations.

5.1. A linear oscillator with damping.

Consider the following two-dimensional stochastic Langevin equation:

$$\begin{aligned} dP &= -aQdt - vPdt - \sigma dW(t), \quad P(0) = p,\\ dQ &= aPdt, \quad Q(0) = q, \end{aligned} \tag{37}$$

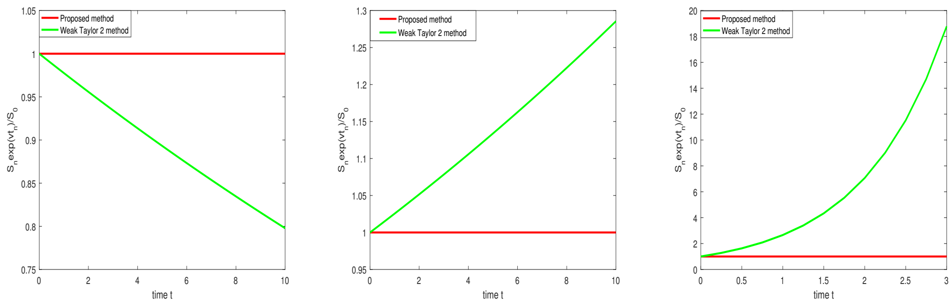
where a, $v > 0$ and $\sigma \neq 0$ are constants and $W(t)$ is a one-dimensional standard Wiener process. The solution to (37) possesses a unique invariant measure μ_1:

$$d\mu_1 = \rho_1(p,q)dpdq,$$

where $\rho_1(p,q) = \Theta \exp\left(-\frac{av(p^2+q^2)}{\sigma^2}\right)$ is known as the Boltzmann–Gibbs density and $\Theta = \left(\int_{\mathbb{R}^2} \exp\left(-\frac{av(p^2+q^2)}{\sigma^2}\right)dpdq\right)^{-1}$ is a renormalization constant. The proposed scheme applied to (37) yields

$$\begin{aligned} P_{n+1} &= e^{-vh}P_n - \frac{h^2}{2}a^2 P_{n+1} - h\Big(1+\frac{vh}{2}\Big)e^{-vh}Q_n - \Big(1+\frac{vh}{2}\Big)e^{-vh}\sigma\Delta_{n+1}W,\\ Q_{n+1} &= Q_n + h\Big(1-\frac{vh}{2}\Big)e^{vh}aP_{n+1} + \frac{h^2}{2}a^2 Q_n + \frac{h}{2}a\sigma\Delta_{n+1}W. \end{aligned} \tag{38}$$

Based on Theorems 3.2 and 3.6, scheme (38) inherits both the conformal symplecticity and ergodicity of the original system. To verify these properties numerically, we choose $p = 3$ and $q = 1$.



(a) $v = 1$. (b) $v = 2$. (c) $v = 4$.

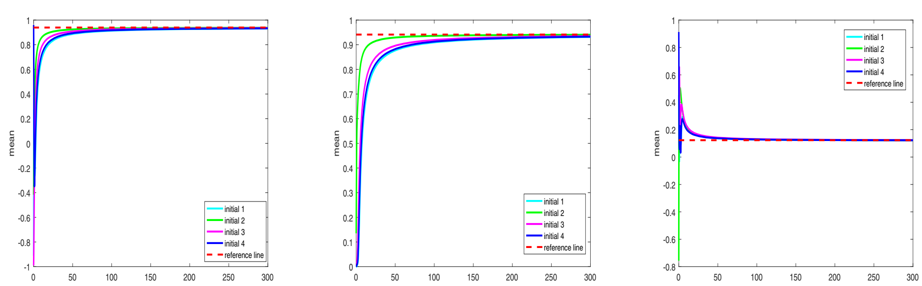
FIG. 1. *The value* $\frac{S_n\exp(vt_n)}{S_0}$ *of two numerical schemes* ($a = 1$ *and* $\sigma = 1$).

Figure 1 shows the value $\frac{S_n\exp(vt_n)}{S_0}$ of the weak Taylor 2 method and the proposed scheme, with v being different dissipative scales and S_n being the triangle square at step n. We choose the original triangle which is produced by three points $(-1, 5)^\top$, $(20, 2)^\top$, $(0, 30)^\top$. We find out that the discrete phase square of the proposed scheme exhibits exponential decay, i.e., $S_n = \exp(-vt_n)S_0$ with the same dissipative coefficient v as in the continuous case, while the weak Taylor 2 scheme does not.

For ergodicity and weak convergence of the proposed scheme, we have taken the three different kinds of test functions (a) $\psi(p, q) = \cos(p + q)$, (b) $\psi(p, q) = \exp\left(-\frac{p^2}{2} - \frac{q^2}{2}\right)$, and (c) $\psi(p, q) = \sin(p^2 + q^2)$ as the test functions. To verify that the temporal averages starting from different initial values will converge to the spatial average, i.e., the ergodic limit

$$\int_{\mathbb{R}^2} \psi(p, q)d\mu_1 = \int_{\mathbb{R}^2} \psi(p, q)\rho_1(p, q)dpdq,$$

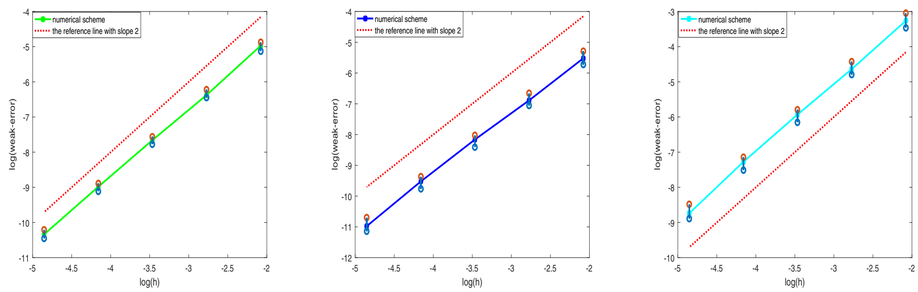
we introduce the reference value for a specific test function ψ to represent the ergodic limit: since the function ψ is uniformly bounded and the density function ρ_1 dissipates exponentially, the integrator is almost zero when $p^2 + q^2$ is sufficiently large. Thus, we choose $\int_{-10}^{10}\int_{-10}^{10} \psi(p, q)\rho_1(p, q)dpdq$ as the reference value, which appears as the dashed line in Figure 2. We can tell from Figure 2 that the tempo-



(a) $\psi(p, q) = \cos(p + q)$. (b) $\psi(p, q) = \exp(-\frac{p^2}{2} - \frac{q^2}{2})$. (c) $\psi(p, q) = \sin(p^2 + q^2)$.

FIG. 2. *The temporal averages* $\frac{1}{N}\sum_{n=1}^{N} \mathbf{E}\psi(P_n, Q_n)$ *starting from different initial values* ($a = 1$, $v = 2$, $\sigma = 0.5$, *and* $T = 300$).

ral averages $\frac{1}{N}\sum_{n=1}^{N}\mathbf{E}\psi(P_n,Q_n)$ of the proposed scheme starting from four different initial values, initial(1) $=(-10,1)^\top$, initial(2) $=(2,0)^\top$, initial(3) $=(0,3)^\top$, and initial(4) $=(4,2)^\top$, converge to the reference line with error no more than $h^2+\frac{1}{T}$, which coincides with Theorem 4.4.



(a) $\psi(p,q)=\cos(p+q)$. (b) $\psi(p,q)=\exp(-\frac{p^2}{2}-\frac{q^2}{2})$. (c) $\psi(p,q)=\sin(p^2+q^2)$.

FIG. 3. *Rate of convergence in weak sense* ($a=1$, $v=2$, *and* $\sigma=0.5$).

Figure 3 plots the value $\ln|\mathbf{E}\psi(P(T),Q(T))-\mathbf{E}\psi(P_N,Q_N)|$ against $\ln h$ for five different step sizes $h=[2^{-3},2^{-4},2^{-5},2^{-6},2^{-7}]$ at $T=1$, where $(P(T),Q(T))$ and (P_N,Q_N) represent the exact and numerical solutions at time T, respectively. It can be seen that the weak order of (38) is two, as indicated by the reference line of slope 2.

5.2. A nonlinear oscillator with linear damping. In this section, we consider the following equation:

$$\begin{aligned} dP &= -(4Q^3-6Q)dt - vPdt + \sqrt{2\beta^{-1}v}\,dW(t), \quad P(0)=p,\\ dQ &= Pdt, \quad Q(0)=q, \end{aligned} \tag{39}$$

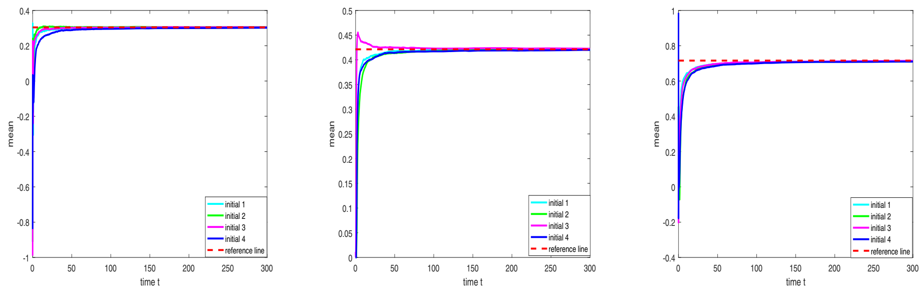
where v, $\beta>0$ are fixed constants and $W(t)$ denotes a one-dimensional standard Wiener process. Similarly to (37), [14] shows that the dynamics generated by (39) is ergodic with the invariant measure μ_2, which can be characterized by the Boltzmann–Gibbs density

$$\rho_2(p,q)=\Theta\exp\left(-\beta\left(\frac{1}{2}p^2+\left(\frac{3}{2}-q^2\right)^2\right)\right)$$

with the renormalization constant $\Theta=\left(\int_{\mathbb{R}^2}e^{-\beta(\frac{1}{2}p^2+(\frac{3}{2}-q^2)^2)}dpdq\right)^{-1}$. Based on (18), we get the associated conformal symplectic scheme

$$\begin{aligned} P_{n+1} &= e^{-vh}P_n - \frac{h^2}{2}P_{n+1}\big(12Q_n^2-4\big) - he^{-vh}\Big(1+\frac{vh}{2}\Big)\Big(4Q_n^3-6Q_n\Big)\\ &\quad + e^{-vh}\Big(1+\frac{vh}{2}\Big)\sqrt{2\beta^{-1}v}\,\Delta_{n+1}W,\\ Q_{n+1} &= Q_n + he^{vh}\Big(1-\frac{vh}{2}\Big)P_{n+1} + \frac{h^2}{2}\Big(4Q_n^3-6Q_n\Big) - \frac{h}{2}\sqrt{2\beta^{-1}v}\,\Delta_{n+1}W. \end{aligned} \tag{40}$$

Since this nonglobal Lipschitz case is not included in Theorems 3.6 and 4.3, we investigate its ergodicity and weak convergence order in view of numerical tests.



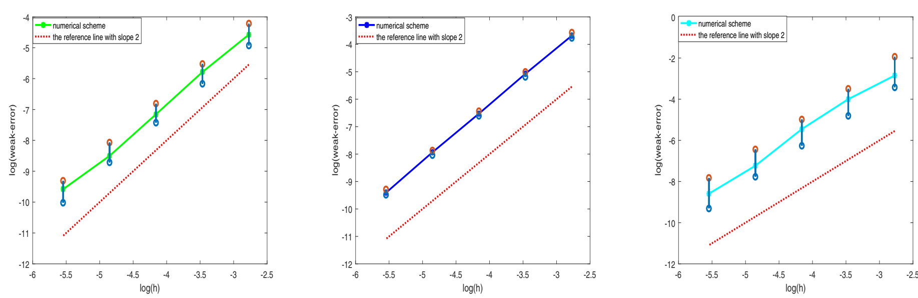
(a) $\psi(p,q)=\cos(p+q)$. (b) $\psi(p,q)=\exp(-\frac{p^2}{2}-\frac{q^2}{2})$. (c) $\psi(p,q)=\sin(p^2+q^2)$.

FIG. 4. *The temporal averages* $\frac{1}{N}\sum_{n=1}^{N}\mathbf{E}\psi(P_n,Q_n)$ *starting from different initial values with* $T=300$.

Let $v=4$, $\beta=2$, and test functions ψ be the same as those in section 5.1. Figure 4 shows the temporal averages $\frac{1}{N}\sum_{n=1}^{N}\mathbf{E}\psi(P_n,Q_n)$ of (40) starting from different initial values initial(1) $=(-10,1)^\top$, initial(2) $=(2,7)^\top$, initial(3) $=(0,3)^\top$, and initial(4) $=(4,6)^\top$. We also use $\int_{-10}^{10}\int_{-10}^{10}\psi(p,q)\rho_2(p,q)dpdq$ as an approximation of the reference value, i.e., the ergodic limit

$$\int_{\mathbb{R}^2}\psi(p,q)d\mu=\int_{\mathbb{R}^2}\psi(p,q)\rho_2(p,q)dpdq.$$

Figure 4 indicates that the proposed scheme also converges to the reference line when time goes to infinity.



(a) $\psi(p,q)=\cos(p+q)$. (b) $\psi(p,q)=\exp(-\frac{p^2}{2}-\frac{q^2}{2})$. (c) $\psi(p,q)=\sin(p^2+q^2)$.

FIG. 5. *Rate of convergence in weak sense* ($p=-2$ *and* $q=-2$).

The value $\ln|\mathbf{E}\psi(P(T),Q(T))-\mathbf{E}\psi(P_N,Q_N)|$ against $\ln h$ for five different step sizes $h=[2^{-4},2^{-5},2^{-6},2^{-7},2^{-8}]$ at $T=0.5$ is shown in Figure 5, similarly to Figure 3. Compared with the reference line of slope 2 in Figure 5, it can be seen that (40) has order two in the sense of weak approximations.

6. Conclusion. In this paper, an approach for constructing high weak order conformal symplectic schemes for stochastic Langevin equations is developed, motivated by the ideas in [1, 2, 18, 24]. The key points are that the generating function is applied to ensure that the proposed scheme preserves the geometric structure, while the modified technique is used to reduce the simulation of multiple integrations. We show that, for the case $k = k' = 1$, the proposed scheme could inherit both the conformal symplectic geometric structure (under Lipschitz assumption) and the ergodicity (under linear assumption) of the stochastic Langevin equation. Numerical experiments verify our theoretical results. In addition, the numerical tests of an oscillator with nonglobal Lipschitz coefficients indicate that the proposed scheme could also inherit the internal properties of the original system, which implies that our results may possibly be extended to the nonglobal Lipschitz case. The theoretical analysis of this extension is also ongoing.

Acknowledgment. The authors are very grateful to Prof. Michael Tretyakov for his helpful discussions and suggestions.

REFERENCES

[1] A. Abdulle, D. Cohen, G. Vilmart, and K. C. Zygalakis, *High weak order methods for stochastic differential equations based on modified equations*, SIAM J. Sci. Comput., 34 (2012), pp. A1800–A1823, https://doi.org/10.1137/110846609.

[2] C. Anton, J. Deng, and Y. S. Wong, *Weak symplectic schemes for stochastic Hamiltonian equations*, Electron. Trans. Numer. Anal., 43 (2014/15), pp. 1–20.

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

[4] N. Bou-Rabee and H. Owhadi, *Boltzmann–Gibbs Preserving Langevin Integrators*, preprint, https://arxiv.org/abs/0712.4123v3, 2017.

[5] N. Bou-Rabee and H. Owhadi, *Long-run accuracy of variational integrators in the stochastic context*, SIAM J. Numer. Anal., 48 (2010), pp. 278–297, https://doi.org/10.1137/090758842.

[6] P. Chartier, E. Hairer, and G. Vilmart, *Numerical integrators based on modified differential equations*, Math. Comp., 76 (2007), pp. 1941–1953, https://doi.org/10.1090/S0025-5718-07-01967-9.

[7] C. Chen, J. Hong, and X. Wang, *Approximation of invariant measure for damped stochastic nonlinear Schrödinger equation via an ergodic numerical scheme*, Potential Anal., 46 (2017), pp 323–367, https://doi.org/10.1007/s11118-016-9583-9.

[8] W. T. Coffey and Yuri P. Kalmykov, *The Langevin Equation: With Applications to Stochastic Problems in Physics, Chemistry and Electrical Engineering*, 3rd ed., World Scientific, Hackensack, NJ, 2012.

[9] G. Da Prato, *An Introduction to Infinite-Dimensional Analysis*, Universitext, Springer-Verlag, Berlin, 2006. Revised and extended from the 2001 original by Da Prato.

[10] D. T. Gillespie, *The chemical Langevin equation*, J. Chem. Phys., 113 (2000), pp. 297–306.

[11] P. E. Kloeden and E. Platen, *Numerical Solution of Stochastic Differential Equations*, Appl. Math. (N. Y.) 23, Springer-Verlag, Berlin, 1992.

[12] B. Leimkuhler, C. Matthews, and G. Stoltz, *The computation of averages from equilibrium and nonequilibrium Langevin molecular dynamics*, IMA J. Numer. Anal., 36 (2016), pp. 13–79, https://doi.org/10.1093/imanum/dru056.

[13] J. C. Mattingly, A. M. Stuart, and D. J. Higham, *Ergodicity for SDEs and approximations: Locally Lipschitz vector fields and degenerate noise*, Stochastic Process. Appl., 101 (2002), pp. 185–232.

[14] J. C. Mattingly, A. M. Stuart, and M. V. Tretyakov, *Convergence of numerical time-averaging and stationary measures via Poisson equations*, SIAM J. Numer. Anal., 48 (2010), pp. 552–577, https://doi.org/10.1137/090770527.

[15] G. N. Milstein, *Numerical Integration of Stochastic Differential Equations*, Math. Appl. 313, Kluwer Academic Publishers, Dordrecht, 1995, https://doi.org/10.1007/978-94-015-8455-5. Translated and revised from the 1988 Russian original.

[16] G. N. Milstein, Yu. M. Repin, and M. V. Tretyakov, *Numerical methods for stochastic systems preserving symplectic structure*, SIAM J. Numer. Anal., 40 (2002), pp. 1583–1604, https://doi.org/10.1137/S0036142901395588.

[17] G. N. Milstein, Yu. M. Repin, and M. V. Tretyakov, *Symplectic integration of Hamiltonian systems with additive noise*, SIAM J. Numer. Anal., 39 (2002), pp. 2066–2088, https://doi.org/10.1137/S0036142901387440.

[18] G. N. Milstein and M. V. Tretyakov, *Quasi-symplectic methods for Langevin-type equa-*

tions, IMA J. Numer. Anal., 23 (2003), pp. 593–626, https://doi.org/10.1093/imanum/23.4.593.

[19] G. N. Milstein and M. V. Tretyakov, *Computing ergodic limits for Langevin equations*, Phys. D, 229 (2007), pp. 81–95, https://doi.org/10.1016/j.physd.2007.03.011.

[20] M. Schienbein and H. Gruler, *Langevin equation, Fokker–Planck equation and cell migration*, Bull. Math. Biol., 55 (1993), pp. 585–608.

[21] D. Talay, *Stochastic Hamiltonian systems: Exponential convergence to the invariant measure, and discretization by the implicit Euler scheme*, Markov Process. Related Fields, 8 (2002), pp. 163–198.

[22] L. Wang, *Variational Integrators and Generating Functions for Stochastic Hamiltonian Systems*, Ph.D. thesis, Karlsruhe Institute of Technology, KIT Scientific Publishing, Karlsruhe, Germany, 2007.

[23] L. Wang and J. Hong, *Generating functions for stochastic symplectic methods*, Discrete Contin. Dyn. Syst., 34 (2014), pp. 1211–1228, https://doi.org/10.3934/dcds.2014.34.1211.

[24] L. Wang, J. Hong, and L. Sun, *Modified equations for weakly convergent stochastic symplectic schemes via their generating functions*, BIT, 56 (2016), pp. 1131–1162, https://doi.org/10.1007/s10543-015-0583-8.