Applied Mathematics and Computation 467 (2024) 128488

Contents lists available at ScienceDirect

Applied Mathematics and Computation

journal homepage: www.elsevier.com/locate/amc

Strong backward error analysis of symplectic integrators for stochastic Hamiltonian systems

Raffaele D'Ambrosio *, Stefano Di Giovacchino

Department of Information Engineering and Computer Science and Mathematics, University of L'Aquila, Italy

A R T I C L E   I N F O

MSC:
65C30
65L07
60H10

Keywords:
Stochastic Hamiltonian systems
Modified differential equations
Symplectic methods
Strong backward error analysis

A B S T R A C T

Backward error analysis is a powerful tool in order to detect the long-term conservative behavior of numerical methods. In this work, we present a long-term analysis of symplectic stochastic numerical integrators, applied to Hamiltonian systems with multiplicative noise. We first compute and analyze the associated stochastic modified differential equations. Then, suitable bounds for the coefficients of such equations are provided towards the computation of long-term estimates for the Hamiltonian deviations occurring along the aforementioned numerical dynamics. This result generalizes Benettin-Giorgilli Theorem to the scenario of stochastic symplectic methods. Finally, specific numerical methods are considered, in order to provide a numerical evidence confirming the effectiveness of the theoretical investigation.

1. Introduction

Deterministic and stochastic Hamiltonian systems are prototypical conservative models widely employed in the scientific literature [4,17,27,38,51,26,50]. Although deterministic models are effective tools to capture the intrinsic nature of physical systems, the introduction of the stochasticity, in such models, has allowed several scientists to describe physical phenomenon with a much greater detail; for example, the non-differentiability of the Wiener process, largely spread in the stochastic modeling, allows to take into account the irreversibility of the time arrow and to include stochastic perturbation that may occur in nature [4]. This paper is fully devoted to a numerical investigation of stochastic Hamiltonian systems, under a *structure-preserving perspective*, falling in the scope of Geometric Numerical Integration (GNI), whose study has been widely conducted through the most recent years [17,27,50]. In this paper, for a given integer $d \geq 1$, we denote by $p \in \mathbb{R}^d$ the vector of generalized momenta and $q \in \mathbb{R}^d$ that of generalized coordinates and we call the Hamiltonian function a sufficiently smooth function $\mathcal{H} : \mathbb{R}^{2d} \to \mathbb{R}$.

We now recall the definition of symplectic map, both in the deterministic and stochastic settings [27,33,42,43].

Definition 1.1. A deterministic map $\phi : \mathbb{R}^{2d} \to \mathbb{R}$ is called symplectic if

$$\phi'(x)^{\intercal} J \phi'(x) = J, \tag{1.1}$$

being ϕ' the Jacobian matrix of ϕ, while a stochastic map $\varphi : \mathbb{R}^{2d} \to \mathbb{R}$ is called symplectic if it satisfies (1.1) almost surely, i.e., if

$$\varphi'(x)^{\intercal} J \varphi'(x) = J, \quad \text{a.s.,}$$

* Corresponding author.
E-mail addresses: raffaele.dambrosio@univaq.it (R. D'Ambrosio), stefano.digiovacchino@graduate.univaq.it (S. Di Giovacchino).

https://doi.org/10.1016/j.amc.2023.128488
Received 5 September 2023; Received in revised form 21 November 2023; Accepted 27 November 2023

Available online 14 December 2023

being $J = \begin{bmatrix} 0 & -I_d \\ I_d & 0 \end{bmatrix}$, where I_d is the identity matrix of dimension d.

1.1. From deterministic to stochastic Hamiltonian systems

In the deterministic setting, Hamiltonian systems are notoriously described by the following system of ordinary differential equations

$$\dot{x}(t) = J\nabla\mathcal{H}(x(t)), \quad x(t_0) = x_0, \quad t \in [t_0, T]. \tag{1.2}$$

Based on Definition 1.1, we recall the following result characterizing the exact flow of Hamiltonian problems (1.2) (see [17,27] and reference therein).

Theorem 1.1. *The exact flow $\varphi_t(x_0)$ of the Hamiltonian system (1.2) is a symplectic map. Moreover, we have*

$$\mathcal{H}(\varphi_t(x_0)) = \mathcal{H}(x_0), \quad \forall\, t \geq t_0. \tag{1.3}$$

A numerical integrator for (1.2) is a map $\Phi_h : \mathbb{R}^{2d} \to \mathbb{R}^{2d}$ such that $x_{n+1} = \Phi_h(x_n),\ n = 0, 1, \dots,$ where $x_n \approx x(nh),\ n = 1, \dots, N,$ being $h = T/N$.

The conservation of the Hamiltonian along the discretized dynamics arising from suitable numerical methods for (1.2) has captured wide attention in the scientific literature, which has revealed that, in general, symplectic methods are good candidates to numerically preserve the Hamiltonian function. Indeed, the Hamiltonian error remains bounded, with size $\mathcal{O}(h^p)$, being p the order of the method, over exponentially long time windows. This result is well-known as *Benettin-Giorgilli theorem* [5,17,28,27,50].

Let us now move to the stochastic case. We consider the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with normal filtration $\left(\mathcal{F}_t\right)_{t\in[t_0,T]}$. Then, stochastic Hamiltonian system can be formulated as a perturbation of (1.2), as follows [42,43]

$$\mathrm{d}X(t) = J\nabla\mathcal{H}(X(t))\mathrm{d}t + \sum_{r=1}^{m} J\nabla\mathcal{H}_r(X(t))\circ\mathrm{d}W_r(t), \quad X(t_0) = X_0, \tag{1.4}$$

where $X_0, X(t) \in \mathbb{R}^{2d}, \mathcal{H}_r : \mathbb{R}^{2d} \to \mathbb{R},\ r = 1, \dots, m,$ are sufficiently smooth Hamiltonian functions. Moreover, each $W_r : \Omega \times [t_0, T] \to \mathbb{R},\ r = 1, \dots, m,$ is a standard $\left(\mathcal{F}_t\right)_{t\in[t_0,T]}$-Wiener process, with continuous sample paths on $(\Omega, \mathcal{F}, \mathbb{P})$. Here, the symbol $\circ$ stands for Stratonovich integration.

A numerical method for (1.4) is a map $\Phi_h : \mathbb{R}^{2d} \times \mathbb{R}^m \to \mathbb{R}^{2d}$ such that

$$X_{n+1} = \Phi_h(X_n, \xi_n),\ n = 0, 1, \dots,$$

where ξ_n is a random vector and $X_n \approx X(t_n),\ n = 1, 2, \dots, N,$ being $N = T/h$.

Let $\varphi_t(X_0, \xi_t)$ be the exact flow of (1.4), where ξ_t is a given stochastic process, depending on the Wiener processes involved in (1.4). In [42,43], the following result has been provided.

Theorem 1.2. *The phase flow $\varphi_t(X_0, \xi_t)$ of system (1.4) is a symplectic map.*

Is well-known that, in general, none of the Hamiltonian functions $\mathcal{H}, \mathcal{H}_1, \dots, \mathcal{H}_r$ is conserved along the exact flow of (1.4), while Casimir functions are constant along such dynamics [42,43]. Through this work, we consider a proper reformulation of (1.4), ensuring that the Hamiltonian function $\mathcal{H}$ is conserved along the exact flow [42,43,45]). We indeed consider $r = 1$ in (1.4) and we take $\mathcal{H}_1 = \sigma\mathcal{H}$, where $\sigma \in \mathbb{R}$. Then, Equation (1.4) reads

$$\mathrm{d}X(t) = J\nabla\mathcal{H}(X(t))\mathrm{d}t + \sigma J\nabla\mathcal{H}(X(t))\circ\mathrm{d}W(t). \tag{1.5}$$

An example of physical system described by the SDE (1.5) has been provided, for example, in [45], where the author has presented the *stochastic Harmonic oscillator*, whose flow satisfies the Hamiltonian SDE

$$\mathrm{d}\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} = \begin{pmatrix} X_2 \\ -X_1 \end{pmatrix}(\mathrm{d}t + \circ\mathrm{d}W(t)).$$

In [42,43,45], the authors have shown that the Hamiltonian function $\mathcal{H}$ is conserved along the flow of (1.5), i.e.,

$$\mathcal{H}(X(t)) = \mathcal{H}(X_0),\ t \geq t_0, \quad \text{a.s.} \tag{1.6}$$

It is worth observing that, even if the link between problem (1.5) and the underlying deterministic Hamiltonian problem (1.2) is very strict, we convey that a full treatment of the strong backward error analysis related to the stochastic case deserves a proper attention, motivating this paper.

1.2. Motivation and targets of the paper

The long-term analysis of discretizations to Hamiltonian systems, both in the deterministic and in the stochastic setting, has raised a significant interest in the scientific literature. As aforementioned, the deterministic scenario is addressed in several monographs [6,8,17,27,28,34,35,39,40,48,50]. In the stochastic setting, several researchers have oriented their studies to the construction and the analysis of numerical integrators aimed to preserve the characteristic trend of the Hamiltonian function along the flow of stochastic Hamiltonian systems [2,10,13–15,42,43,45,19–21], while first attempts to perform a backward error analysis of selected stochastic numerical integrators for stochastic conservative systems appeared in [1–3,49,53], where the authors introduced procedures to develop the so-called *weak modified equations*, i.e., stochastic differential equations, associated to specific numerical schemes, whose exact flow coincides with the numerical one in weak sense. In [3,18,22], weak modified equations have been used to perform a deep study of numerical methods for stochastic Hamiltonian systems. In particular, in [16], the case of linear Itô Hamiltonian equations has been addressed, while [18] is fully oriented to the long-term analysis of Hamiltonian errors in a more general scenario, where the authors have proven that the expected Hamiltonian error may generally grow exponentially in time. A backward error analysis with respect to the strong measure, at least for Euler-Maruyama scheme, has been introduced in [23] and exploited in [33] also in other contexts.
More in general, backward error analysis has been revealed itself to be a powerful instrument that has found its application also in other problems, mainly arising from optimization theory, ergodic theory and differential geometry; see, for example, [22,24,31,32,36,37]. This paper aims to rigorously provide the long-term analysis of numerical methods for stochastic Hamiltonian problems (1.5); such schemes are intended as stochastic perturbation of symplectic deterministic integrators. At least at the best of our knowledge, this aspect has remained a challenging open question in the scientific literature for several years. Through this work, we aim to provide an exhaustive answer to the following question, that in deterministic case admits a positive feedback (see, for instance, [5,17,27,50]):

Are stochastic symplectic numerical schemes suitable candidates in order to achieve a good behavior in terms of numerical Hamiltonian conservation?

Then, the goal of this paper is to provide a stochastic counterpart of the *Benettin-Giorgilli theorem* [5,27] that explains the long-term behavior of numerical Hamiltonians in the deterministic scenario. We specialize the framework of strong stochastic modified equations [23,33] to the setting of Stratonovich Hamiltonian systems that, along the stochastic flow, exhibit the kind of conservation described in Equation (1.6). The employed strategy generalized the results contained in [27] to the stochastic case and the well-known *Wong-Zakay approximation* of the Wiener process [33,52] will be also employed.

The paper proceeds as follows. In Section 2, we introduce fundamental concepts on the symplectic integration in the stochastic setting. In Section 3, the construction and the analysis of the stochastic modified equations is presented. Next, Section 4 contains the main results of this paper, providing long-term estimates for the Hamiltonian deviation associated to stochastic symplectic integrators, while Section 5 analyzes the non-symplectic case. Finally, selected examples and numerical experiments, confirming the effectiveness of the theoretical analysis, are provided in Section 6 and some conclusions and possible future developments on this line of research are reported in Section 7. Appendix A contains the proof of the results needed overall the treatise.

2. Background and preliminary results

In this section, we recall and introduce some preliminary results on the symplectic integration of stochastic Hamiltonian systems (1.5). To this purpose, we focus our attention to Stratonovich stochastic differential equation (SDE) of the form

$$\mathrm{d}X(t) = f(X(t))\big(\mathrm{d}t + \sigma \circ \mathrm{d}W(t)\big), \tag{2.7}$$

with vector field f sufficiently smooth. Correspondingly, a one-step numerical method for (2.7) is defined by the nonlinear map

$$X_{n+1} = \Phi_h(X_n, \Delta W_n), \tag{2.8}$$

where $\Delta W_n \sim N(0,h)$. Let us now recall the well-known definition of mean-square order [25,30].

Definition 2.1. A one-step numerical method (2.8) for (2.7) has mean-square order p if

$$\mathbb{E}\|X(h) - \Phi_h(X_0, \Delta W_0)\|^2 = \mathcal{O}\left(h^{p+1}\right), \tag{2.9}$$

for sufficiently small values of h.

A relevant example of numerical scheme for (2.7) is given by stochastic Runge-Kutta methods [11,9,46,47], having the following form

$$
\begin{aligned}
X^i &= X_n + \xi_n \sum_{j=1}^{s} a_{ij} f(X^j), \quad i = 1, \ldots, s, \\
X_{n+1} &= X_n + \xi_n \sum_{i=1}^{s} b_i f(X^i), \quad \xi_n = h + \sigma \Delta W_n.
\end{aligned} \tag{2.10}
$$

Clearly, if the diffusive part in (2.7) has $\sigma = 0$, stochastic Runge-Kutta methods (2.10) reduce to their underlying deterministic Runge-Kutta scheme

$$
\begin{aligned}
X^i &= X_n + h \sum_{j=1}^{s} a_{ij} f(X^j), \quad i = 1, \ldots, s, \\
X_{n+1} &= X_n + h \sum_{i=1}^{s} b_i f(X^i).
\end{aligned} \tag{2.11}
$$

For methods (2.11), the following assumption will be adopted.

Assumption 2.1. The underlying deterministic method (2.11) of the stochastic Runge-Kutta (2.10) is symplectic.

We are ready to state the following result.

Theorem 2.1. *Let us consider stochastic Runge-Kutta methods (2.10) under Assumption 2.1. Then, methods (2.10) are symplectic according to Definition 1.1.*

We aim to provide a proof of Theorem 2.1 by inheriting and properly remodulating the underlying deterministic scenario [27]. Its proof is a direct consequence of the following two lemmas, also requiring the following definition.

Definition 2.2. A scalar function $C : \mathbb{R}^d \to \mathbb{R}$ is a conserved quantity along the trajectory $X(t)$, solution to (2.7), if

$$
C(X(t)) = C(X_0), \quad t \geq t_0, \quad \text{a.s},
$$

with $X_0 = X(t_0)$, i.e., if

$$
\nabla C(x)^{\intercal} f(x) = 0, \quad \forall\, x \in \mathbb{R}^{2d}.
$$

We prove the following lemma.

Lemma 2.1. *Let us consider stochastic Runge-Kutta methods (2.10). Then, if Assumption 2.1 holds true, then methods (2.10) preserve all quadratic invariants for the flow of SDE (2.7).*

Proof. See Appendix A.1. □

Given the exact flow $\varphi_t(X_0, \xi_t)$ of the SDE (1.5), we associate the following quantity

$$
M(\varphi_t(X_0, \xi_t)) := \frac{\partial \varphi_t(X_0, \xi_t)}{\partial X_0} \tag{2.12}
$$

and we derive a differential equation, known as *variational equation* [27], satisfied by $M(\varphi_t(X_0, \xi_t))$. We start by considering that the flow $\varphi_t(X_0, \xi_t)$ of the Hamiltonian SDE (1.5) can be written as

$$
\varphi_t(X_0, \xi_t) = \varphi_0(X_0, \xi_0) + \int_{t_0}^{t} J \nabla \mathcal{H}(\varphi_s(X_0, \xi_s)) \mathrm{d}s + \sigma \int_{t_0}^{t} J \nabla \mathcal{H}(\varphi_s(X_0, \xi_s)) \circ \mathrm{d}W(s). \tag{2.13}
$$

Differentiating (2.13) leads to

$$
\begin{aligned}
\frac{\partial \varphi_t(X_0, \xi_t)}{\partial X_0} &= I_{2d} + \frac{\partial}{\partial X_0} \int_{t_0}^{t} J \nabla \mathcal{H}(\varphi_s(X_0, \xi_s)) \mathrm{d}s + \sigma \frac{\partial}{\partial X_0} \int_{t_0}^{t} J \nabla \mathcal{H}(\varphi_s(X_0, \xi_s)) \circ \mathrm{d}W(s) \\
&= I + \int_{t_0}^{t} J \nabla^2 \mathcal{H}(\varphi_s(X_0, \xi_s)) \frac{\partial}{\partial X_0} \varphi_s(X_0, \xi_s) \mathrm{d}s + \sigma \int_{t_0}^{t} J \nabla^2 \mathcal{H}(\varphi_s(X_0, \xi_s)) \frac{\partial}{\partial X_0} \varphi_s(X_0, \xi_s) \circ \mathrm{d}W(s).
\end{aligned}
$$

Since $M(\varphi_0(X_0, \xi_0)) = I_{2d}$, we conclude that $M(\varphi_t(X_0, \xi_t))$ solves the SDE

$$\mathrm{d}M(\varphi_t(X_0,\xi_t)) = J\nabla^2\mathcal{H}(\varphi_t(X_0,\xi_t))M(\varphi_t(X_0,\xi_t))\,(\mathrm{d}t + \sigma\circ\mathrm{d}W(t))\,. \tag{2.14}$$

Lemma 2.2. *The quantity $(M^\intercal J M)(\varphi_t(X_0,\xi_t))$ is a quadratic conservative quantity for the exact flow of the variational SDE (2.14), i.e.,*

$$(M^\intercal J M)(\varphi_t(X_0,\xi_t)) = (M^\intercal J M)(\varphi_t(X_0,\xi_0)), \quad a.s.,$$

where $\varphi_t(X_0,\xi_t)$ solves (1.5).

Proof. Applying the stochastic chain rule [29] to (2.14) and using the properties of the matrix J yields

$$\begin{aligned}\mathrm{d}\left(M^\intercal J M\right) &= \mathrm{d}M^\intercal\,(JM) + M^\intercal J\,\mathrm{d}M\\ &= M^\intercal\nabla^2\mathcal{H}(\varphi_t(X_0,\xi_t))J^\intercal J M\,(\mathrm{d}t + \sigma\circ\mathrm{d}W(t))\\ &+ M^\intercal J J\nabla^2\mathcal{H}(\varphi_t(X_0,\xi_t))M\,(\mathrm{d}t + \sigma\circ\mathrm{d}W(t)) = 0,\end{aligned}$$

where, by the sake of simplicity of treatment, we have omitted the argument of the function M. □

According to Theorem 2.1, the stochastic perturbation (2.10) of a symplectic Runge-Kutta method (2.11) is able of preserving its symplecticity.

The symplectic integration of stochastic Hamiltonian systems (1.4), i.e., the study of numerical integrators for (1.4) able to retain the symplecticity along their numerical dynamics, has been the object of several research papers [10,41–43] and of the monographs [17,33], all mostly devoted to the construction and the analysis of the order of convergence of new stochastic symplectic numerical methods, suitable for systems (1.4), is also worth remarking that several papers in the scientific literature have highlighted the long-term behavior of numerical discretizations to SDEs, and in particular, to stochastic Hamiltonian systems, via weak backward error analysis tools; see, e.g., [16,18,22,53] and reference therein. As a consequence, at least at the best of our knowledge, the strong backward error analysis of symplectic schemes solving (1.5) has remained an open question.

To this purpose, the starting point of our analysis regards the computation of the stochastic modified differential equations, that is the topic of next section.

3. Stochastic modified equations for Hamiltonian SDEs

In this section, based on the general idea proposed in [33] and inspired by the theory proposed in the deterministic setting [27], we aim to derive a stochastic modified equation whose exact flow equals the numerical solution $\Phi_h(X_0,\Delta W_0)$ to (2.7), at grid points $t_n, n = 0,1,\dots,N$. The starting ingredient of such analysis is the Wong-Zakay approximation [33,52], that allows us to achieve an approximation of the Wiener process $W(t)$ that is differentiable. We indeed proceed as follows.

For $t \in [t_n,t_{n+1}]$, let us define the following approximation

$$W(t) \approx \widetilde{W}(t) := W(t_n) + \frac{t-t_n}{h}\Delta W_n, \quad t\in[t_n,t_{n+1}], \tag{3.15}$$

with $\Delta W_n = W(t_{n+1}) - W(t_n)$, so that

$$\dot{\widetilde{W}}(t) = \frac{1}{h}\Delta W_n, \quad t\in[t_n,t_{n+1}]. \tag{3.16}$$

Note that, for any $n = 0,1,\dots,N$, we have $W(t_n) \equiv \widetilde{W}(t_n)$, a.s.. Then, the procedure (also developed in [33], Section III.1.3) consists in considering the SDE

$$dX(t) = f(X(t))\left(dt + \sigma\circ d\widetilde{W}(t)\right), \quad t\in[t_n,t_{n+1}], \tag{3.17}$$

instead of the SDE (2.7). Thanks to the differentiability of the approximating process $\widetilde{W}(t)$ and by (3.16), last equation provides the random ordinary differential equation

$$\dot{X}(t) = f(X(t))\left(1 + \frac{\sigma\Delta W_n}{h}\right) = f(X(t))\frac{1}{h}\left(h + \sigma\Delta W_n\right), \quad t\in[t_n,t_{n+1}],$$

that is,

$$\dot{X}(t) = f(X(t))\frac{\xi_n}{h}, \quad t\in[t_n,t_{n+1}], \tag{3.18}$$

where $\xi_n = h + \sigma\Delta W_n$.

Remark 3.1. The solution to the random ODE (3.18) is not equal to the stochastic process given by the solution to the SDE (2.7), but it can be seen as a convenient approximation. In particular, when the vector field is that of the Hamiltonian SDE, i.e., $f(x) = J\nabla H(x)$, the exact flows of both equations (2.7) and (3.18) share the property of being energy-preserving and symplectic. Moreover, it is immediate to see that the numerical methods (2.10), applied to the random ODE (3.18), provide the same numerical solutions if they were applied to the SDE (2.7). Then, it makes sense to use the random ODE (3.18) to derive results on the Hamiltonian errors of methods (2.10), applied to the Hamiltonian SDE of the form (1.5).

We now introduce the following stochastic modified equation

$$\dot{\widetilde{X}}(t) = \left(f(\widetilde{X}(t)) + \xi_0 f_2(\widetilde{X}(t)) + \xi_0^2 f_3(\widetilde{X}(t)) + \dots \right) \frac{\xi_0}{h}, \quad t \in [t_0, t_0 + h], \tag{3.19}$$

where the coefficients $f_j, j = 2, \dots$, are such that $\widetilde{X}(h) = X_1$. It is worth observing that the power series defined in (3.19) is only formal since, in general, it may not be convergent [27]. We will discuss again this issue and define a suitable truncation of the modified vector field at the beginning of Section 4. By Taylor series arguments, we get

$$\begin{aligned}\widetilde{X}(t_0 + h) = X_0 &+ h\left(f + \xi_0 f_2 + \xi_0^2 f_3 + \dots \right) \frac{\xi_0}{h} \\ &+ \frac{h^2}{2!}\left(f' + \xi_0 f_2' + \xi_0^2 f_3' \right)\left(f + \xi_0 f_2 + \xi_0^2 f_3 \right) \frac{\xi_0^2}{h^2} + \frac{h^3}{3!}\left(f''(f, f) + f' f' f + \dots \right) \frac{\xi_0^3}{h^3} + \dots,\end{aligned}$$

where each function in the right-hand side is evaluated in X_0. Hence,

$$\widetilde{X}(t_0 + h) = X_0 + \xi_0 f + \xi_0^2 \left[f_2 + \frac{1}{2!} f' f \right] + \xi_0^3 \left[f_3 + \frac{1}{2!}(f' f_2 + f_2' f) + \frac{1}{3!}(f''(f, f) + f' f' f) \right] + \dots \tag{3.20}$$

If we assume that the numerical method $\Phi_h(X_0, \Delta W_0)$ admits the following expansion

$$\Phi_h(X_0, \Delta W_0) = X_0 + \xi_0 f(X_0) + \xi_0^2 d_2(X_0) + \xi_0^3 d_3(X_0) + \dots, \tag{3.21}$$

then, the modified coefficients f_j are computed as follows

$$f_2 = d_2 - \frac{1}{2!} f' f, \quad f_3 = d_3 - \frac{1}{3!}\left(f''(f, f) + f' f' f \right) - \frac{1}{2!}\left(f' f_2 + f_2' f \right), \quad \dots \tag{3.22}$$

It is worth noting that stochastic Runge-Kutta methods (2.10) satisfy an expansion of the form (3.21), as we will verify for selected numerical methods in Appendix A.5 and Appendix A.6.

Remark 3.2. Equations in (3.22) are formally equal to those given in [12,27] for the deterministic case. This is due to the structure of Equation (2.7) that allows us to perform the expansions in power series of the stochastic parameter ξ_0.

The following result holds true.

Theorem 3.1. *For numerical methods (2.8) of mean-square order $p \geq 1$, we have that all coefficients f_j, $j = 2, \dots, p$, in (3.19) are identically equal to 0.*

Proof. The result straightforwardly descends by comparing the same powers of ξ_0 in (3.19) and (3.21), taking into account (3.22) and considering that

$$\mathbb{E}\left[|\xi_0^p|^2 \right] = \mathcal{O}(h^p), \quad p \geq 1. \quad \square$$

The following integrability lemma [27] is now reported. The reader can find a complete proof of this result in Lemma VI.2.7 in [27].

Lemma 3.1. *Let $D \in \mathbb{R}^m$ be open, $f : D \to \mathbb{R}^m$ be continuously differentiable, and assume that the Jacobian $f'(y)$ is symmetric for all $y \in D$. Then, for every $y_0 \in D$ there exists an open neighborhood of y_0, say $N \subset D$, and a function $\mathcal{H} : \mathbb{R}^m \to \mathbb{R}$, such that*

$$f(y) = \nabla \mathcal{H}(y),$$

for $y \in N$.

Now, we are ready to prove the following result that generalizes the deterministic theory about the existence of modified Hamiltonians for symplectic schemes [27] to the stochastic case.

Theorem 3.2. *Let us consider a stochastic Hamiltonian system (1.5) and apply a symplectic numerical method (2.8). Then, the stochastic modified equation (3.19) is Hamiltonian, i.e., there exist smooth functions $\mathcal{H}_j(y)$ such that $f_j(y) = J\nabla\mathcal{H}_j(y)$, $j = 2, 3, \dots$, with*

$$\widetilde{\mathcal{H}}(y) = \mathcal{H}(y) + \xi_0 \mathcal{H}_2(y) + \xi_0^2 \mathcal{H}_3(y) + \dots. \tag{3.23}$$

Proof. We proceed by induction on j. Since $f_1(y) = f(y) = J\nabla\mathcal{H}(y)$, then we assume there exist functions $f_2, \dots, f_r$ such that the truncated system

$$\widetilde{X}^{[r]}(t) = \left(f(\widetilde{X}^{[r]}(t)) + \xi_0 f_2(\widetilde{X}^{[r]}(t)) + \dots + \xi_0^{r-1} f_r(\widetilde{X}^{[r]}(t)) \right) \frac{\xi_0}{h}$$

is Hamiltonian with Hamiltonian function $\widetilde{\mathcal{H}}^{[r]}(y) = \mathcal{H}(y) + \xi_0\mathcal{H}_2(y) + \dots + \xi_0^{r-1}\mathcal{H}_r(y)$. Let us denote by $\widetilde{\varphi}_{r,h}(X_0, \Delta W_0)$ its flow. Then, by construction, we get that the numerical flow, i.e., the flow of (3.19), satisfies

$$\Phi_h(X_0, \Delta W_0) = \widetilde{\varphi}_{r,h}(X_0, \Delta W_0) + \xi_0^{r+1} f_{r+1}(X_0) + \mathcal{O}\left(\xi_0^{r+2}\right), \tag{3.24}$$

and (see [27], Section IX.3.1)

$$\Phi'_h(X_0, \Delta W_0) = \widetilde{\varphi}'_{r,h}(X_0, \Delta W_0) + \xi_0^{r+1} f'_{r+1}(X_0) + \mathcal{O}\left(\xi_0^{r+2}\right).$$

Since both $\Phi_h(X_0, \Delta W_0)$ and $\widetilde{\varphi}_{r,h}(X_0, \Delta W_0)$ are symplectic maps, we obtain

$$\begin{aligned} J &= \Phi'_h(X_0, \Delta W_0)^{\intercal} J \Phi_h(X_0, \Delta W_0)' \\ &= J + \xi_0^{r+1}\left[f'_{r+1}(X_0)^{\intercal} J \widetilde{\varphi}'_{r,h}(X_0, \Delta W_0) + \widetilde{\varphi}'_{r,h}(X_0, \Delta W_0)^{\intercal} J f'_{r+1}(X_0)\right] + \mathcal{O}\left(\xi_0^{r+2}\right), \text{ a.s.} \end{aligned}$$

Since $\widetilde{\varphi}'_{r,h}(X_0, \Delta W_0) = I + \mathcal{O}(\xi_0)$, we get

$$J = J + \xi_0^{r+1}\left[f'_{r+1}(X_0)^{\intercal} J + J f'_{r+1}(X_0)\right] + \mathcal{O}\left(\xi_0^{r+2}\right), \quad \text{a.s.}$$

Hence, the matrix Jf'_{r+1} is symmetric and the result on f_{r+1} holds true by Lemma 3.1. □

Remark 3.3. It is worth remarking that the expansion provided in (3.23) is only formal. We operate a truncation in Section 4.

Remark 3.4. The $\mathcal{O}$ term in (3.24) is referred to the mean-square topology, i.e., it encloses terms

$$K_{r+2}(X_0)\xi_0^{r+2} + K_{r+3}(X_0)\xi_0^{r+3} + \dots + K_{r+N}(X_0)\xi_0^{r+N} + \dots,$$

such that the term $K_{r+2}(X_0)\xi_0^{r+2}$ is dominant in the mean-square sense.

The existence of modified Hamiltonian in the stochastic setting has also been treated in [41,33,42], using alternative techniques and in the setting of *Rough stochastic systems*.

Remark 3.5. The stochastic Runge-Kutta methods (2.10) are implicit both in the drift and diffusion terms of the SDE. Then, the unboundedness of the random variable ΔW_n may lead to a lack of uniqueness of solutions to the nonlinear system of equations defined by the relation for the stages (2.10), making such methods not well-defined. This was also observed, for example, in [1,33,42,44,52].

In [33,44,52], the author showed that, for methods of strong order less or equal than 1, a way to overcome this issue may be to replace the unbounded random variable ΔW_n in the method with the bounded random

$$\bar{\xi}_n = h + \sigma\sqrt{h}Z_n, \quad Z_n = \begin{cases} Y_n, & |Y_n| \le A_h, \\ A_h, & Y_n > A_h, \\ -A_h, & Y_n < -A_h, \end{cases} \tag{3.25}$$

being $Y_n \sim N(0,1)$ and $A_h = \sqrt{k|\ln h|}$, where k is typically chosen equal to 4 (see, for example [33,42,44,52]). We will follow this approach in the numerical experiments of the present paper.

For higher order methods, i.e., for $p > 2$ in Definition (2.1), the approximation (3.25) may not work well [42,44] and one may obtain the well-posedness of such methods by looking at truncated approximations of iterated stochastic integrals; see, for example, [29,42,44,52] and reference therein. However, in the sequel of the present research, we adopt the approximation (3.25) since, as it is well-known, obtaining stochastic methods with $p > 2$ (in Definition 2.1) may be a challenging issue [11,33,44] and, hence, we find reasonable focusing on the most popular scenario.

Under (3.25), the corresponding amended stochastic Runge-Kutta methods (2.10) read

$$X_{n+1} = X_n + \bar{\xi}_n \sum_{i=1}^{s} b_i f(X^i), \qquad X^i = X_n + \bar{\xi}_n \sum_{j=1}^{s} a_{ij} f(X^j), \qquad i = 1, \dots, s. \tag{3.26}$$

The following technical lemma [33] provides a bound for the truncated increment Z_n, for any $n \ge 0$.

Lemma 3.2. *For any $0 < \epsilon < 1/2$ and $k \geq 1$, there exists a constant $\bar{h} = \bar{h}(\epsilon, k)$ such that*

$$|Z_n| \leq h^{-\epsilon}, \quad \forall\, h \leq \bar{h}, \quad \forall\, n \geq 0, \quad \text{a.s.}$$

Due to Lemma 3.2, for any $n \geq 0$, the random variable $\bar{\xi}_n$ satisfies

$$|\bar{\xi}_n| \leq h + \sigma h^{1/2-\epsilon}, \quad h \leq \bar{h}, \quad \text{a.s.} \tag{3.27}$$

The following assumption will be used in the sequel [27].

Assumption 3.1. There exist constants $M, R > 0$ such that $\|f(y)\| \leq M$, for any $\|y - y_0\| \leq 2R$.

Moreover, with reference to methods (2.10), we introduce the notation

$$\mu = \sum_{i=1}^{s} |b_i|, \quad \lambda = \max_{i=1,\dots,d} \sum_{j=1}^{s} |a_{ij}|. \tag{3.28}$$

In addition, given $R \in \mathbb{R}$, $y_0 \in \mathbb{R}^m$, we call

$$B_R(y_0) = \{y \in \mathbb{R}^m : \|y - y_0\| \leq R\}.$$

The following technical lemma provides a bound for coefficients $d_j(y)$ given in (3.21), that generalize the underlying deterministic scenario [27] to the stochastic case.

Lemma 3.3. *If $f(y)$ is analytic in $B_{2R}(y_0)$ and satisfies Assumption 3.1, then the coefficients $d_j(y)$ in (3.21) are analytic functions in $B_R(y_0)$ and satisfy*

$$\left\|d_j(y)\right\| \leq \mu M \left(\frac{2\lambda M}{R}\right)^{j-1}, \tag{3.29}$$

for $\|y - y_0\| \leq R$, as long as h satisfies $h + \sigma h^{1/2-\epsilon} \leq R/(2\lambda M)$.

Proof. See Appendix A.2. □

The following technical lemma relies on the notion of Lie derivative [27] that, for a fixed index i, is given by

$$(D_i g)(y) := g'(y) f_i(y),$$

where $f_1(y) = f(y)$.

Lemma 3.4. *If the map Φ in (2.8) can be expanded as in (3.21), then the functions $f_j(y)$ of the stochastic modified equation (3.19) satisfy*

$$f_j(y) = d_j(y) - \sum_{i=2}^{j} \frac{1}{i!} \sum_{k_1+\cdots+k_i=j} \left(D_{k_1} \dots D_{k_{i-1}} f_{k_i}\right)(y), \tag{3.30}$$

where each k_m is greater or equal than 1.

Proof. See Appendix A.3. □

We conclude this subsection presenting the following bound on the coefficients of the stochastic modified equation in (3.19).

Lemma 3.5. *Let $f(y)$ be analytic in $B_{2R}(y_0)$, let the Taylor series coefficients of the numerical method (3.21) be analytic in $B_R(y_0)$. If Assumption 3.1 and (3.29) hold, then*

$$\|f_j(y)\| \leq \ln 2\, \eta M \left(\frac{\eta M j}{R}\right)^{j-1} \quad \text{for} \quad \|y - y_0\| \leq R/2, \tag{3.31}$$

where $\eta = 2\max\{\lambda, \mu/(2\ln 2 - 1)\}$, as long as h satisfies $h + \sigma h^{1/2-\epsilon} \leq R/(2\lambda M)$.

Proof. The proof arises as direct consequence of Lemma 3.3, Lemma 3.4, together with Lemma IX 7.4 and Lemma IX 7.5 in [27]. □

4. Long-term energy conservation of stochastic symplectic methods

In this section, using the bounds provided in the previous one, we aim to provide all core results of this work, i.e., the long-term error estimates associated to numerical Hamiltonians computed by symplectic methods.

As announced in Section 3, the formal series provided in (3.19) is generally non convergent [27]. Hence, we need to consider a truncated stochastic modified equation, i.e., an equation of the form

$$\dot{\widetilde{X}}(t)=F_N(\widetilde{X}(t))\frac{\bar{\xi}_n}{h},\quad t\in[t_n,t_{n+1}],\tag{4.32}$$

where $F_N(y)=f(y)+\bar{\xi}_n f_2(y)+\cdots+\bar{\xi}_n^{N-1}f_N(y)$. We present the following proposition on conservative quantities along the solution of the truncated modified SDE (4.32).

Proposition 4.1. *Let us assume that the scalar quantity $C:\mathbb{R}^m\to\mathbb{R}$ is conserved along the flow of* (2.7) *and also along the dynamics provided by a numerical method $\Phi_h(X_0,\Delta W_0)$. Then, it remains constant also along the flow of its truncated modified SDE* (4.32)*, say $\widetilde{\varphi}_{N,h}(X_0)$.*

Proof. The proof proceeds by induction. First of all, let us take $N=2$. Since C is conserved along the flow of (2.7), we have

$$\nabla C(x)^{\intercal}f(x)=0,\quad\forall\, x\in\mathbb{R}^m.\tag{4.33}$$

Let us consider the following truncated modified equation

$$\dot{\widetilde{X}}(t)=\left(f(\widetilde{X}(t))+\xi_0 f_2(\widetilde{X}(t))\right)\frac{\xi_0}{h},\quad t\in[t_0,t_1].\tag{4.34}$$

Hence, for the flow of (4.34), taking into account (4.33), we can write

$$\frac{\mathrm{d}}{\mathrm{d}t}C(\widetilde{X})=\nabla C(\widetilde{X})^{\intercal}\left(f(\widetilde{X}(t))+\xi_0 f_2(\widetilde{X}(t))\right)\frac{\xi_0}{h}=\frac{\xi_0^2}{h}\nabla C(\widetilde{X})^{\intercal}f_2(\widetilde{X}(t)).$$

Hence, the goal is to show that

$$\nabla C(x)^{\intercal}f_2(x)=0,\quad\forall\, x\in\mathbb{R}^m.\tag{4.35}$$

By assumption (3.21), we have

$$\begin{aligned}C(\Phi_h(X_0,\Delta W_0))&=C\left(X_0+\xi_0 f(X_0)+\xi_0^2 d_2(X_0)+\mathcal{O}(\xi_0^3)\right)\\&=C(X_0)+\nabla C(X_0)^{\intercal}\left(\xi_0 f(X_0)+\xi_0^2 d_2(X_0)+\mathcal{O}(\xi_0^3)\right)\\&\qquad+\frac{1}{2}\xi_0^2 f(X_0)^{\intercal}\nabla^2 C(X_0)f(X_0)+\mathcal{O}(\xi_0^3).\end{aligned}$$

Hence, by (4.33), we get

$$C(\Phi_h(X_0,\Delta W_0))=C(X_0)+\xi_0^2\left[\nabla C(X_0)^{\intercal}d_2(X_0)+\frac{1}{2}f(X_0)^{\intercal}\nabla^2 C(X_0)f(X_0)\right]+\mathcal{O}(\xi_0^3).$$

Then, the conservation of C along $\Phi_h(X_0,\Delta W_0)$ implies

$$\nabla C(x)^{\intercal}d_2(x)+\frac{1}{2}f(x)^{\intercal}\nabla^2 C(x)f(x)=0,\quad\forall\, x\in\mathbb{R}^m,$$

that, by (3.22), reduces to

$$\nabla C(x)^{\intercal}\left(f_2(x)+\frac{1}{2}f'(x)f(x)\right)+\frac{1}{2}f(x)^{\intercal}\nabla^2 C(x)f(x)=0,\quad\forall\, x\in\mathbb{R}^m.\tag{4.36}$$

Then, in view of (4.35), it remains to show that

$$\nabla C(x)^{\intercal}f'(x)f(x)+f(x)^{\intercal}\nabla^2 C(x)f(x)=0,\quad\forall\, x\in\mathbb{R}^m.\tag{4.37}$$

This can be achieved by differentiating in time the left-hand side of (4.33) along the flow of (2.7). We indeed have

$$\begin{aligned}0=\frac{\mathrm{d}}{\mathrm{d}t}\left(\nabla C(x)^{\intercal}f(x)\right)&=\nabla\left(\nabla C(x)^{\intercal}f(x)\right)\cdot f(x)\\&=\left(\nabla^2 C(x)f(x)\right)\cdot f(x)+\nabla C(x)^{\intercal}f'(x)f(x),\end{aligned}$$

that gives (4.37).

Now, for a given integer $2\le r\le N-1$, let $\widetilde{\varphi}_{r,h}(X_0,\Delta W_0)$ be the flow of the truncated modified equation of order r, that is, the flow of

$$\dot{\widetilde{X}}_r(t)=F_r(\widetilde{X}_r(t))\frac{\xi_0}{h},\qquad t\in[t_0,t_0+h],$$

where $F_r(x) = f(x) + \xi_0 f_2(x) + \cdots + \xi_0^{r-1} f_r(x)$, and, as inductive hypothesis, we assume that the quantity C is conserved along such flow. The goal is then to show that

$$C\left(\widetilde{\varphi}_{r+1,h}(X_0, \Delta W_0)\right) = C(X_0) \qquad \text{a.s.,} \tag{4.38}$$

where $\widetilde{\varphi}_{r+1,h}(X_0, \Delta W_0)$ is the flow of

$$\dot{\widetilde{X}}_{r+1}(t) = \left(F_r(\widetilde{X}_{r+1}(t)) + \xi_0^r f_{r+1}(\widetilde{X}_{r+1}(t)\right)\frac{\xi_0}{h}, \qquad t \in [t_0, t_0 + h].$$

By the inductive assumption, (4.38) is achieved if

$$\nabla C(x)^\intercal f_{r+1}(x) = 0, \qquad \forall\, x \in \mathbb{R}^m. \tag{4.39}$$

By the construction of the modified equation and by assumptions, we have

$$\begin{aligned} C(X_0) &= C\left(\Phi_h(X_0, \Delta W_0)\right) = C\left(\widetilde{\varphi}_{r,h}(X_0, \Delta W_0) + \xi_0^{r+1} f_{r+1}(X_0) + \mathcal{O}\left(\xi_0^{r+2}\right)\right) \\ &= C\left(\widetilde{\varphi}_{r,h}(X_0, \Delta W_0)\right) + \xi_0^{r+1} \nabla C(\widetilde{\varphi}_{r,h}(X_0, \Delta W_0))^\intercal f_{r+1}(X_0) + \mathcal{O}\left(\xi_0^{r+2}\right) \\ &= C(X_0) + \xi_0^{r+1} \nabla C(\widetilde{\varphi}_{r,h}(X_0, \Delta W_0))^\intercal f_{r+1}(X_0) + \mathcal{O}\left(\xi_0^{r+2}\right), \end{aligned}$$

that implies

$$\nabla C(\widetilde{\varphi}_{r,h}(X_0, \Delta W_0))^\intercal f_{r+1}(X_0) = 0, \qquad \text{a.s., } \forall\, X_0 \in \mathbb{R}^m.$$

Since

$$\nabla C(\widetilde{\varphi}_{r,h}(X_0, \Delta W_0)) = \nabla C(X_0) + \mathcal{O}\left(\xi_0\right),$$

we get (4.38), that concludes the proof. □

In the scientific literature, a possible truncation strategy relies on the choice of N minimizing the term $\|\bar{\xi}_n^{N-1} f_N(y)\|$ (see, for instance, [17,27]). Using (3.31), we obtain

$$\begin{aligned} \|\bar{\xi}_n^{N-1} f_N(y)\| &\leq |\bar{\xi}_n|^{N-1} \|f_N(y)\| \leq |\bar{\xi}_n|^{N-1} \ln 2\, \eta M \left(\frac{\eta M N}{R}\right)^{N-1} \\ &= \ln 2\, \eta M \left(\frac{\eta M |\bar{\xi}_n| N}{R}\right)^{N-1}, \quad \text{a.s.} \end{aligned} \tag{4.40}$$

This term is minimized for $\bar{\xi}_n N = h_0$, where $h_0 = R/(\eta M e)$. Hence, N is chosen as the largest integer such that $(h + \sigma h^{1/2-\epsilon})N \leq h_0$. Correspondingly, we introduce the following useful lemma.

Lemma 4.1. *Under Assumption 3.1 and (3.31), for a numerical method of mean-square order p, we have*

$$\|F_N(y)\| \leq M(1 + 1.65\eta), \quad \|F_N(y) - f(y)\| \leq c M |\bar{\xi}|^p, \quad \text{a.s.,} \tag{4.41}$$

for $\|y - y_0\| \leq R/2$, as long as $(h + \sigma h^{2-\epsilon})N \leq \mathrm{e} h_0$. The constant c only depends on the coefficients of the method.

Proof. See Appendix A.4. □

Now, let $\widetilde{\varphi}_{N,h}(X_0, Z_0)$ be the flow of the truncated stochastic modified equation (4.32). We state the following result.

Theorem 4.1. *Let us suppose that $f(y)$ is analytic in $B_{2R}(y_0)$ and the coefficients $d_j(y)$ of (3.21) are analytic in $B_R(y_0)$. Assume that Assumption 3.1 and (3.29) hold true. Then, as long as $h + \sigma h^{1/2-\epsilon} \leq h_0/4$, we have*

$$\|\Phi_h(X_0, Z_0) - \widetilde{\varphi}_{N,h}(X_0, Z_0)\| \leq (h + \sigma h^{1/2-\epsilon}) c_1 M \mathrm{e}^{-h_0/(h + \sigma h^{1/2-\epsilon})}, \quad \text{a.s,} \tag{4.42}$$

with c_1 depending only on the method.

Proof. We define the auxiliary random variable

$$g(\bar{\xi}_0) := \Phi_h(X_0, \bar{\xi}_0) - \widetilde{\varphi}_{N,h}(X_0, \bar{\xi}_0).$$

We note that, for any fixed realization of Z_0, the random variable $g(\bar{\xi}_0)$ is an analytic function of $\bar{\xi}_0$, for sufficiently small values of $\bar{\xi}_0$, since $\Phi_h(X_0, \bar{\xi}_0)$ and $\widetilde{\varphi}_{N,h}(X_0, \bar{\xi}_0)$ are analytic as well. By Equation (3.22), the leading term of $g(\bar{\xi}_0)$ has size $\mathcal{O}\left(\bar{\xi}_0^{N+1}\right)$. Then, the maximum principle for analytic functions, applied to $g(\bar{\xi}_0)/\bar{\xi}_0^{N+1}$, gives

$$\left\| \frac{g(\bar{\xi}_0)}{\bar{\xi}_0^{N+1}} \right\| = \frac{1}{|\bar{\xi}_0|^{N+1}} \|g(\bar{\xi}_0)\| \leq \frac{1}{\varepsilon^{N+1}} \max_{|z|\leq\varepsilon} \|g(z)\|, \ |\bar{\xi}_0| \leq \varepsilon, \quad \text{a.s.},$$

i.e.,

$$\|g(\bar{\xi}_0)\| \leq \left(\frac{|\bar{\xi}_0|}{\varepsilon} \right)^{N+1} \max_{|z|\leq\varepsilon} \|g(z)\|, \ |\bar{\xi}_0| \leq \varepsilon, \quad \text{a.s.}, \tag{4.43}$$

if $g(z)$ is analytic for $z \leq \varepsilon$. Taking $\varepsilon = \mathrm{e}h_0/N$ and since

$$\|g(z)\| \leq \|\Phi_h(X_0, z) - X_0\| + \|\widetilde{\varphi}_{N,h}(X_0, z) - X_0\| \quad \text{a.s.}, \tag{4.44}$$

we separately estimate the two terms in the right-hand side of (4.44). We start with the first, by looking at the i-th term of the summation in (3.21). Applying (3.29), we have

$$z^i d_i(X_0) \leq |z|^i \|d_i(X_0)\| \leq |z|^i \mu M \left(\frac{2\lambda M}{R} \right)^{i-1} = |z|^i \frac{\mu M R}{2\lambda M} \left(\frac{2\lambda M}{R} \right)^i, \quad \text{a.s.} \tag{4.45}$$

Then, the power-series in (3.21) is almost surely bounded by

$$\sum_{i=2}^{\infty} \frac{\mu R}{2\lambda} \left(\frac{|z| 2\lambda M}{R} \right)^i,$$

that is convergent for $|z| \leq R/(4\lambda M)$, i.e., $\Phi_h(X_0, z)$ is analytic for $|z| \leq R/(4\lambda M)$. Since $\eta > 2\lambda$ and, as a consequence of $h \leq h_0/4$, we have $N \geq 4$, then we conclude that $\Phi_h(X_0, z)$ is analytic for $|z| \leq \varepsilon$. Applying again (3.29) and Assumption 3.1, for $|z| \leq \varepsilon$, gives

$$\|\Phi_h(X_0, z) - X_0\| \leq |z| \|f(X_0)\| + \sum_{i=2}^{\infty} |z|^j \|d_j(X_0)\| \leq |z| M \left(1 + \mu \sum_{i=2}^{\infty} \left(\frac{|z| 2\lambda M}{R} \right)^{j-1} \right), \text{ a.s.} \tag{4.46}$$

Since

$$\frac{2\lambda M}{R} |z| \leq \frac{2\lambda M}{R} \varepsilon = \frac{2\lambda}{\mu N} \leq 1, \quad \text{a.s.},$$

then the series in (4.46) is almost surely convergent. Hence, we get

$$\|\Phi_h(X_0, z) - X_0\| \leq |z| M \left(1 + \mu \frac{q}{1-q} \right), \quad \text{a.s.},$$

where $|q| = (|z| 2\lambda M/R) \leq 1$, a.s. Then, we conclude

$$\|\Phi_h(X_0, z) - X_0\| \leq |z| M (1 + \mu), \ |z| \leq \varepsilon, \quad \text{a.s.} \tag{4.47}$$

Similar computations and the use of the first bound in (4.41) show

$$\|\widetilde{\varphi}_{N,h}(X_0, z) - X_0\| \leq |z| M (1 + 1.65\eta), \ |z| \leq \varepsilon, \quad \text{a.s.} \tag{4.48}$$

We note that $\widetilde{\varphi}_{N,h}(X_0, z)$ is analytic in $|z| \leq \varepsilon$. Then, inserting (4.46) and (4.47) into (4.44), and taking into account that $\varepsilon = \mathrm{e}h_0/N$, yields

$$\|g(\bar{\xi}_0)\| \leq \left(\frac{|\bar{\xi}_0|}{\varepsilon} \right)^{N+1} \varepsilon M C = |\bar{\xi}_0| M C \left(\frac{|\bar{\xi}_0|}{\varepsilon} \right)^N = |\bar{\xi}_0| M C \left(\frac{|\bar{\xi}_0| N}{e h_0} \right)^N \leq |\bar{\xi}_0| M C \mathrm{e}^{-N}, \quad \text{a.s.}, \tag{4.49}$$

because of $|\bar{\xi}_0| N \leq h_0$ a.s., where $C = 2 + 1.65\eta + \mu$. Since

$$N \leq \frac{h_0}{h + \sigma h^{1/2-\epsilon}} \leq N + 1 \quad \text{a.s.},$$

i.e.,

$$-N \leq 1 - \frac{h_0}{h + \sigma h^{1/2-\epsilon}} \quad \text{a.s.},$$

leading to the result. □

The result obtained in Theorem 4.1, together with the bounds provided by the previous section, allows us to state the following theorem, regarding the long-term energy preservation for symplectic schemes along numerical paths.

Let us consider the truncated stochastic modified equation (3.19), associated to a symplectic method of mean-square order p, that is Hamiltonian by Theorem 3.2, i.e.,

$$\dot{\widetilde{X}}(t) = J \nabla \widetilde{H}^{[N]}(\widetilde{X}(t)) \frac{\xi_n}{h}, \quad t \in [t_n, t_{n+1}], \tag{4.50}$$

with modified Hamiltonian given by

$$\widetilde{\mathcal{H}}^{[N]}(y) = \mathcal{H}(y) + \xi_n^p \mathcal{H}_{p+1}(y) + \cdots + \xi_n^{N-1}\mathcal{H}_N(y). \tag{4.51}$$

Proposition 4.2. *For a symplectic method of mean-square order p,*

$$\mathbb{E}\left[\left\|\widetilde{\mathcal{H}}^{[N]}(\widetilde{X}(t)) - \mathcal{H}(X(t))\right\|^2\right] = \mathcal{O}(h^p),$$

where $\widetilde{\mathcal{H}}^{[N]}(y)$ is given in (4.51) and $\widetilde{X}(t)$ solves (4.50).

Proof. By (4.50) and (4.51), and since $\widetilde{X}_0 = X_0$, we have

$$\begin{aligned}\widetilde{\mathcal{H}}^{[N]}(\widetilde{X}(t)) &= \widetilde{\mathcal{H}}^{[N]}(\widetilde{X}_0) = \widetilde{\mathcal{H}}^{[N]}(X_0) = \mathcal{H}(X_0) + \xi_n^p \mathcal{H}_{p+1}(X_0) + \cdots + \xi_n^{N-1}\mathcal{H}_N(X_0)\\ &= \mathcal{H}(X_0) + \xi_n^p\left(\mathcal{H}_{p+1}(X_0) + \xi_n^2 \mathcal{H}_{p+2}(X_0) + \cdots + \xi_n^{N-1-p}\mathcal{H}_N(X_0)\right).\end{aligned}$$

In the mean-square topology, we see that the summation

$$\xi_n^2 \mathcal{H}_{p+2}(X_0) + \cdots + \xi_n^{N-1-p}\mathcal{H}_N(X_0)$$

is negligible with respect to the term $\mathcal{H}_{p+1}(X_0)$, i.e., we can write

$$\xi_n^p\left(\mathcal{H}_{p+1}(X_0) + \xi_n^2 \mathcal{H}_{p+2}(X_0) + \cdots + \xi_n^{N-1-p}\mathcal{H}_N(X_0)\right) = \mathcal{O}\left(\xi_n^p\right).$$

This, together with the fact that $\mathcal{H}(X_0) = \mathcal{H}(X(t))$, yields

$$\widetilde{\mathcal{H}}^{[N]}\left(\widetilde{X}(t)\right) - \mathcal{H}(X(t)) = \mathcal{O}\left(\xi_n^p\right). \tag{4.52}$$

Side-by-side squaring and taking expectation in (4.52), leads to the thesis. □

Theorem 4.2. *Consider a Stratonovich Hamiltonian system (1.5) with analytic Hamiltonian function given by $\mathcal{H} : D \to D \subset \mathbb{R}^{2d}$, D an open set, and a symplectic method $\Phi_h(X_0, \Delta W_0)$, with mean-square order p and stepsize h. If the numerical solution evolves in a compact set $K \subset D$, then there exist h_0 and N such that*

$$\begin{aligned}&\widetilde{\mathcal{H}}^{[N]}(X_n) = \widetilde{\mathcal{H}}^{[N]}(X_0) + \mathcal{O}\left(\exp\left(-\frac{h_0}{2(h+\sigma h^{1/2-\epsilon})}\right)\right), \quad \text{a.s.},\\ &\mathcal{H}(X_n) = \mathcal{H}(X_0) + \mathcal{O}\left((h+\sigma h^{1/2-\epsilon})^p\right), \quad \text{a.s.,}\end{aligned} \tag{4.53}$$

as long as

$$nh \le \frac{h\cdot\exp\left(\frac{h_0}{2(h+\sigma h^{1/2-\epsilon})}\right)}{h+\sigma h^{1/2-\epsilon}}.$$

Proof. The proof extends to the stochastic case the procedure drawn in [27] for the deterministic one. We have

$$\begin{aligned}\widetilde{\mathcal{H}}^{[N]}(X_n) - \widetilde{\mathcal{H}}^{[N]}(X_0) &= \sum_{j=1}^{n} \widetilde{\mathcal{H}}^{[N]}(X_j) - \widetilde{\mathcal{H}}^{[N]}(X_{j-1}) = \sum_{j=1}^{n} \widetilde{\mathcal{H}}^{[N]}(X_j) - \widetilde{\mathcal{H}}^{[N]}(\varphi_{N,h}(X_{j-1}, \Delta W_{j-1}))\\ &= \sum_{j=1}^{n} \mathcal{O}\left((h+\sigma h^{1/2-\epsilon})c_1 M \mathrm{e}^{-h_0/(h+\sigma h^{1/2-\epsilon})}\right)\\ &= \mathcal{O}\left(n(h+\sigma h^{1/2-\epsilon})c_1 M \mathrm{e}^{-h_0/(h+\sigma h^{1/2-\epsilon})}\right),\end{aligned}$$

that gives the first equation in (4.53). The second relation comes from the fact that the modified Hamiltonian in X_n is given by

$$\widetilde{\mathcal{H}}^{[N]}(X_n) = \mathcal{H}(X_n) + \bar{\xi}_{n-1}^p\left[\mathcal{H}_{p+1}(X_n) + \bar{\xi}_{n-1}\mathcal{H}_{p+2} + \cdots + \bar{\xi}_{n-1}^{N-p-1}\mathcal{H}_N(X_n)\right]. \tag{4.54}$$

We see that $\mathcal{H}_{p+1}(X_n) + \bar{\xi}_{n-1}\mathcal{H}_{p+2} + \cdots + \bar{\xi}_{n-1}^{N-p-1}\mathcal{H}_N(X_n)$ is almost surely uniformly bounded in K, independently on h and N. By the proof of Lemma 3.1 (see [27], Lemma VI.2.7), we have that close to $X_0 \in D$,

$$\mathcal{H}_{p+1}(X_n) + \bar{\xi}_{n-1}\mathcal{H}_{p+2} + \cdots + \bar{\xi}_{n-1}^{N-p-1}\mathcal{H}_N(X_n) = \sum_{j=0}^{N-p-1}\int_0^1 \bar{\xi}_{n-1}^j\left\langle X, f_{p+j+1}(tX)\right\rangle dt + K_0,$$

where K_0 is a constant. By (3.31), the left-hand side of the above equation is a.s. bounded by

$$\sum_{j=0}^{N-p-1} |\bar{\xi}_{n-1}|^j \|X_n\| \|f_{p+j+1}(tX_n)\| + K_0 \le \ln 2\eta M K_1 \sum_{j=0}^{N-p-1} |\bar{\xi}_{n-1}|^j \left(\frac{\eta M(p+j+1)}{R}\right)^{p+j} + K_0$$
$$\le \ln 2\eta M K_1 \sum_{j=0}^{N-p-1} \left(\frac{|\bar{\xi}_{n-1}|\eta M(p+j+1)}{R}\right)^j \widetilde{C}_j + K_0,$$

where K_1 is a bound for $\|X\|$, with $X \in K$ and $\widetilde{C}_j = \left((\eta M(p+j+1))/R\right)^p$. Moreover, since $|\bar{\xi}|_{n+1}/(\mathrm{e}h_0) \le 1/N$, we end up with

$$\|\mathcal{H}_{p+1}(X_n) + \bar{\xi}_{n-1}\mathcal{H}_{p+2} + \cdots + \bar{\xi}_{n-1}^{N-p-1}\mathcal{H}_N(X_n)\| \le K_0 + \ln 2\eta M K_1 \sum_{j=0}^{N-p-1} \left(\frac{p+j+1}{N}\right)^j \widetilde{C}_j.$$

Defining

$$K_2 := \max_N \left(\frac{p+j+1}{N}\right)^j \widetilde{C}_j,$$

we see that

$$\|\mathcal{H}_{p+1}(X_n) + \bar{\xi}_{n-1}\mathcal{H}_{p+2} + \cdots + \bar{\xi}_{n-1}^{N-p-1}\mathcal{H}_N(X_n)\| \le K_3, \quad X \in K, \quad \text{a.s.}, \tag{4.55}$$

where $K_3 = K_0 + \ln 2\eta M K_1 K_2$. Hence, we get

$$\widetilde{\mathcal{H}}^{[N]}(X_n) = \mathcal{H}(X_n) + \mathcal{O}\left(\bar{\xi}_{n-1}^p\right), \quad \text{a.s.} \tag{4.56}$$

Then, from (4.56) and the first relation in (4.53), we achieve

$$\mathcal{H}(X_n) + \mathcal{O}\left(\bar{\xi}_{n-1}^p\right) = \widetilde{\mathcal{H}}^{[N]}(X_0) + \mathcal{O}\left(\mathrm{e}^{-h_0/2(h+\sigma h^{1/2-\epsilon})}\right)$$
$$= \mathcal{H}(X_0) + \mathcal{O}\left(\bar{\xi}_0^p\right) + \mathcal{O}\left(\mathrm{e}^{-h_0/2(h+\sigma h^{1/2-\epsilon})}\right).$$

As long as the term $(h+\sigma h^{1/2-\epsilon})^p$ dominates the exponential term $\mathrm{e}^{-h_0/(2(h+\sigma h^{1/2-\epsilon})}$, we get

$$\mathcal{H}(X_n) = \mathcal{H}(X_0) + \mathcal{O}\left((h+\sigma h^{1/2-\epsilon})^p\right), \quad \text{a.s.},$$

that concludes the proof. □

Remark 4.1. Looking at the term

$$\left(h+\sigma h^{1/2-\epsilon}\right)^p = \sum_{k=0}^{p} \binom{p}{k} h^{p-k}\sigma^k h^{k(1/2-\epsilon)} = h^p + \sum_{k=1}^{p}\binom{p}{k}\sigma^k h^{p-k(1/2+\epsilon)},$$

we observe that the Hamiltonian deviation in (4.53) can be rewritten as

$$\mathcal{H}(X_n) = \mathcal{H}(X_0) + \mathcal{O}(h^p) + \mathcal{O}\left(\sum_{k=1}^{p}\binom{p}{k}\sigma^k h^{p-k(1/2+\epsilon)}\right), \quad \text{a.s.} \tag{4.57}$$

Equation (4.57) allows us to make a comparison with the deterministic case, that obeys to Benettin-Giorgilli Theorem [17,27]. One can indeed observe the presence of an extra term, i.e., the summation term, that may affect the $\mathcal{O}(h^p)$ Hamiltonian error, typically visible in symplectic schemes, due to the presence of the stochastic parameter σ and the Wiener process $W(t)$. As long as σ becomes larger, the Hamiltonian error may grow, even if we use symplectic schemes. We will test this issue in Section 6.

Remark 4.2. It is worth observing that the estimates provided in Theorem 4.2 are valid in time intervals of length

$$t_n \le \frac{h\cdot\exp\left(\frac{h_0}{2(h+\sigma h^{1/2-\epsilon})}\right)}{h+\sigma h^{1/2-\epsilon}}. \tag{4.58}$$

Hence, we observe a significant reduction of such time interval if compared to the deterministic scenario, in which the boundedness of the Hamiltonian errors hold in intervals of size $\mathcal{O}\left(\mathrm{e}^{h_0/2h}\right)$ [27].

4.1. The case of purely stochastic Hamiltonian system

A particular case of study is represented by the purely stochastic Hamiltonian system, i.e., the SDE of the form

$$dX = J\nabla H(X)\circ dW, \tag{4.59}$$

that is, the only diffusive term appears. Under the Wong-Zakay approximation (3.15), the SDE (4.59) reads

$$\dot{X}(t) = J\nabla H(X(t))\frac{\Delta W_n}{h}, \quad t \in [t_n, t_{n+1}], \quad \Delta W_n = W(t_{n+1}) - W(t_n).$$

Here, we have $\xi_n = \Delta W_n$. Of course, we can apply all the theoretical results presented in this paper to SDE (4.59), such as the existence of a modified truncated Hamiltonian for symplectic schemes of mean-square order p (see Definition (2.1)). It turns to be of the form

$$\widetilde{\mathcal{H}}^{[N]}(y) = \mathcal{H}(y) + \Delta W_n^p \mathcal{H}_{p+1}(y) + \cdots + \Delta W_n^{N-1}\mathcal{H}_N(y).$$

In situations in which one can adopt the truncation (3.25), as those discussed in Remark (3.5), ΔW_n is replaced by the bounded random variable $\bar{\xi}_n = \sqrt{h}Z_n$, where Z_n is defined in (3.25) and satisfies, by Lemma 3.2, the inequality

$$|\bar{\xi}_n| \leq \sqrt{h}h^{-\epsilon} = h^{1/2-\epsilon}, \quad \forall\, h \leq \bar{h}, \quad \text{a.s..}$$

Then, by similar computations as those in Theorem 4.2, it follows that, for symplectic methods applied to (4.59), the truncated modified Hamiltonian $\widetilde{\mathcal{H}}^{[N]}$ satisfies

$$\widetilde{\mathcal{H}}^{[N]}(X_n) - \widetilde{\mathcal{H}}^{[N]}(X_0) = \mathcal{O}\left(nh^{1/2-\epsilon}\mathrm{e}^{-h_0/(h^{1/2-\epsilon})}\right) = \mathcal{O}\left(\mathrm{e}^{-h_0/(2h^{1/2-\epsilon})}\right),$$

as long as

$$nh^{1/2-\epsilon} \leq \mathrm{e}^{h_0/2h^{1/2-\epsilon}},$$

i.e.,

$$t_n \leq h^{1/2+\epsilon}\mathrm{e}^{h_0/2h^{1/2-\epsilon}}. \tag{4.60}$$

It follows that, for time windows of size given by (4.60), the numerical Hamiltonian of symplectic methods for the SDE (4.59) satisfies

$$H(X_n) = H(X_0) + \mathcal{O}\left(h^{p/2-p\epsilon}\right), \quad \text{a.s..} \tag{4.61}$$

While the $h^{-p\epsilon}$ factor in the $\mathcal{O}$ term in (4.60) takes into account the effect of the truncation (3.25), the factor $h^{p/2}$ highlights a reduction of a factor of a half respect to the deterministic accuracy $\mathcal{O}(h^p)$. This may be explained by looking, for example, at the implicit midpoint method applied to (4.60)

$$X_{n+1} = X_n + \sqrt{h}Z_n J\nabla H\left(\frac{X_{n+1}+X_n}{2}\right),$$

with Z_n as in (3.25). We can read the above method as

$$X_{n+1} = X_n + \sqrt{h}J\nabla \widehat{H}\left(\frac{X_{n+1}+X_n}{2}\right), \quad \widehat{H}(X) = Z_n H(X),$$

so that is can be seen as the application of the implicit midpoint to the Hamiltonian system with Hamiltonian $\widehat{H}$, with stepsize $\hat{h} = \sqrt{h}$. This justifies the presence of the factor $h^{p/2}$ in the estimate (4.61).

5. Mean-square Hamiltonian growth for non-symplectic schemes

In this section, we present a result providing the time growth of the numerical Hamiltonian deviation associated to non-symplectic methods applied to (1.5). It is worth recalling that, in the deterministic scenario [27], the authors have proven that a linear growth of the Hamiltonian deviation occurs. To some extent, the following theorem aims to derive a stochastic counterpart of such a deterministic evidence.

Theorem 5.1. *For a non-symplectic scheme $\Phi_h(X_0, \Delta W_0)$ of mean-square order p, applied to (1.5), the following error estimate holds true*

$$\mathbb{E}\left[\left\|\mathcal{H}(X_n) - \mathcal{H}(X_0)\right\|^2\right] = \mathcal{O}\left(t_n h^p\right), \tag{5.62}$$

being $X_n \approx X(t_n)$ the numerical solution obtained after the application of n steps of the method.

Proof. Denoting by $\varphi_h(X_0, \Delta W_0)$ the exact flow of (1.5), we can write

$$\mathcal{H}(X_n) - \mathcal{H}(X_0) = \sum_{j=1}^{n}\mathcal{H}(X_j) - \mathcal{H}(X_{j-1}) = \sum_{j=1}^{n}\mathcal{H}(X_j) - \mathcal{H}(\varphi_h(X_{j-1}, \Delta W_{j-1})),$$

so that

$$\left\|\mathcal{H}(X_n) - \mathcal{H}(X_0)\right\|^2 = \left\|\sum_{j=1}^{n}\mathcal{H}(X_j) - \mathcal{H}(\varphi_h(X_{j-1}, \Delta W_{j-1}))\right\|^2.$$

Since $\|\cdot\|_2 := \mathbb{E}[\|\cdot\|^2]^{1/2}$ is a norm, using the triangular inequality, we get

$$\begin{aligned}\left\|\mathcal{H}(X_n)-\mathcal{H}(X_0)\right\|_2^2 &= \left\|\sum_{j=1}^{n}\mathcal{H}(X_j)-\mathcal{H}(\varphi_h(X_{j-1},\Delta W_{j-1}))\right\|_2^2\\ &\le \sum_{j=1}^{n}\left\|\mathcal{H}(X_j)-\mathcal{H}(\varphi_h(X_{j-1},\Delta W_{j-1}))\right\|_2^2,\end{aligned}$$

that is,

$$\mathbb{E}\left[\left\|\mathcal{H}(X_n)-\mathcal{H}(X_0)\right\|^2\right] \le \sum_{j=1}^{n}\mathbb{E}\left[\left\|\mathcal{H}(X_j)-\mathcal{H}(\varphi_h(X_{j-1},\Delta W_{j-1}))\right\|^2\right].$$

Then, by standard local error estimation arguments [42,43], we obtain

$$\mathbb{E}\left[\left\|\mathcal{H}(X_n)-\mathcal{H}(X_0)\right\|^2\right] = \sum_{j=1}^{n}\mathcal{O}\left(h^{p+1}\right) = \mathcal{O}\left(nh^{p+1}\right) = \mathcal{O}(nh\cdot h^p),$$

that concludes the proof. □

6. Numerical experiments

In this section, we aim to provide a selection of numerical tests confirming the sharpness and the effectiveness of the theoretical analysis developed in the previous sections.

6.1. Analysis of the stochastic midpoint rule

We test the stochastic midpoint method (1s-Gaussian Runge-Kutta method)

$$X_{n+1} = X_n + \bar{\xi}_n J\nabla\mathcal{H}\left(\frac{X_{n+1}+X_n}{2}\right), \quad \bar{\xi}_n = h+\sigma\sqrt{h}Z_n, \tag{6.63}$$

where the truncated random variable Z_n is defined as in (3.25). From [33], we conclude that the mean-square order of method (6.63), according to Definition 2.1, is $p=2$ and, from [33,42], we learn that it is maintained by taking $k=4$ in (3.25), i.e., $A_h=\sqrt{4|\ln h|}$. If we define

$$M_{n+1} := \frac{\partial X_{n+1}}{\partial X_n},$$

then, differentiation of (6.63) yields

$$\frac{\partial X_{n+1}}{\partial X_n} = I + \frac{\bar{\xi}_n}{2}J\nabla^2\mathcal{H}\left(\frac{X_{n+1}+X_n}{2}\right)\left(\frac{\partial X_{n+1}}{\partial X_n}+I\right),$$

where the identity matrix I has dimension $2d$. Hence, we obtain

$$\left(I-\frac{\bar{\xi}_n}{2}J\nabla^2\mathcal{H}(\bar{X})\right)M_{n+1} = I+\frac{\bar{\xi}_n}{2}J\nabla^2\mathcal{H}(\bar{X}), \tag{6.64}$$

that is,

$$M_{n+1} = \left(I-\frac{\bar{\xi}_n}{2}J\nabla^2\mathcal{H}(\bar{X})\right)^{-1}\left(I+\frac{\bar{\xi}_n}{2}J\nabla^2\mathcal{H}(\bar{X})\right),$$

where we have defined $\bar{X} := \left(X_{n+1}+X_n\right)/2$.
A straightforward computation, similar to the one presented in the deterministic setting and contained in [17,27], shows that

$$M_n^{\intercal} J M_n = J, \quad \text{a.s.}, \quad n\ge 0. \tag{6.65}$$

Hence, by Theorem 3.2, the stochastic midpoint rule (6.63) is symplectic. We test Equation (6.65) for the Hamiltonian function of the mathematical pendulum

$$\mathcal{H}(p,q) = \frac{1}{2}p^2-\cos(q), \tag{6.66}$$

and we take $\sigma=0.1, T=100, h=0.1$. In Fig. 1, we plot the quantity $\|M_n^{\intercal}JM_n-J\|_\infty$, $n=0,1,\dots,N=T/h$, computed along the numerical trajectory of (6.63) applied to SDE (1.5) with (6.66), confirming the validity of (6.65).

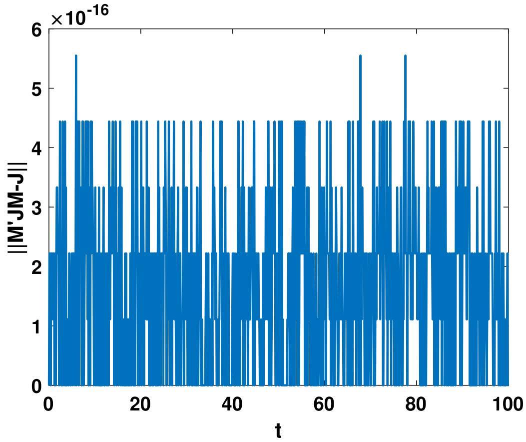


Fig. 1. Pattern of $\|M_n^{\intercal} J M_n - J\|_\infty$, computed over a single trajectory computed by applying the stochastic midpoint method (6.63) to (1.5) with Hamiltonian function (6.66). Here $h = 0.1$, $\sigma = 0.1$ and the initial values are $p(0) = 1, q(0) = 2$.

By the calculations provided in Appendix A.5, for separable Hamiltonians quadratic in p, the first order stochastic modified equation associated to method (6.63) reads

$$\begin{cases} d\widetilde{p} = \left[-V'(\widetilde{q}) + \dfrac{\xi_n^2}{24} \left(\widetilde{p}^2 V'''(\widetilde{q}) + 2V''(\widetilde{q}) V'(\widetilde{q}) \right) \right] (dt + \circ dW), \\ d\widetilde{q} = \left[\widetilde{p} - \dfrac{\xi_n^2}{12} V''(\widetilde{q}) \widetilde{p} \right] (dt + \circ dW), \quad t \in [t_n, t_{n+1}], \end{cases} \tag{6.67}$$

$$\widetilde{\mathcal{H}}^{[3]}(p, q) = \mathcal{H}(p, q) - \frac{\xi_n^2}{24} \left[V''(q) p^2 + V'(q)^2 \right]. \tag{6.68}$$

Hence, for the Hamiltonian defined in (6.66), we obtain

$$\widetilde{\mathcal{H}}^{[3]}(p, q) = \mathcal{H}(p, q) - \frac{\xi_n^2}{24} \left[\cos(q) p^2 + \sin(q)^2 \right].$$

It is worth noting that, for quadratic Hamiltonians $\mathcal{H}(p, q) = (p^2 + q^2)/2$, the corresponding modified Hamiltonian (6.68) becomes

$$\widetilde{\mathcal{H}}^{[3]}(p, q) = \mathcal{H}(p, q) - \frac{\xi_n^2}{24} \left[p^2 + q^2 \right] = \left(1 - \frac{\xi_n^2}{12} \right) \mathcal{H}(p, q)$$

and one can directly verify that the modified flow of (6.67) with quadratic potentials satisfies

$$\mathrm{d}\mathcal{H}(p, q) = 0,$$

i.e., we have the energy conservation, as we expect by the symplecticity of the stochastic midpoint method (6.63), confirming the result in Proposition 4.1. Moreover, we also note that the modified flow of (6.67), when the Hamiltonian function is quadratic, satisfies

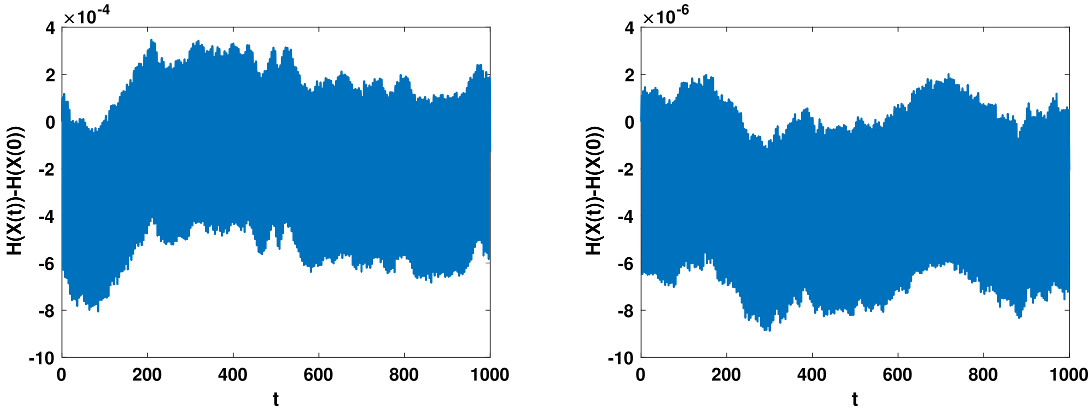
$$\widetilde{\mathcal{H}}^{[3]}(\widetilde{X}(t)) = \left(1 - \frac{\xi_n^2}{12} \right) \mathcal{H}(\widetilde{X}(t)) = \left(1 - \frac{\xi_n^2}{12} \right) \mathcal{H}(X_0) = \mathcal{H}(X(t)) + \mathcal{O}\left(\xi_n^2 \right),$$

that gives

$$\mathbb{E} \left[\widetilde{\mathcal{H}}^{[3](\widetilde{X}(t))} - \mathcal{H}(X(t)) \|^2 \right] = \mathcal{O}\left(h^2 \right),$$

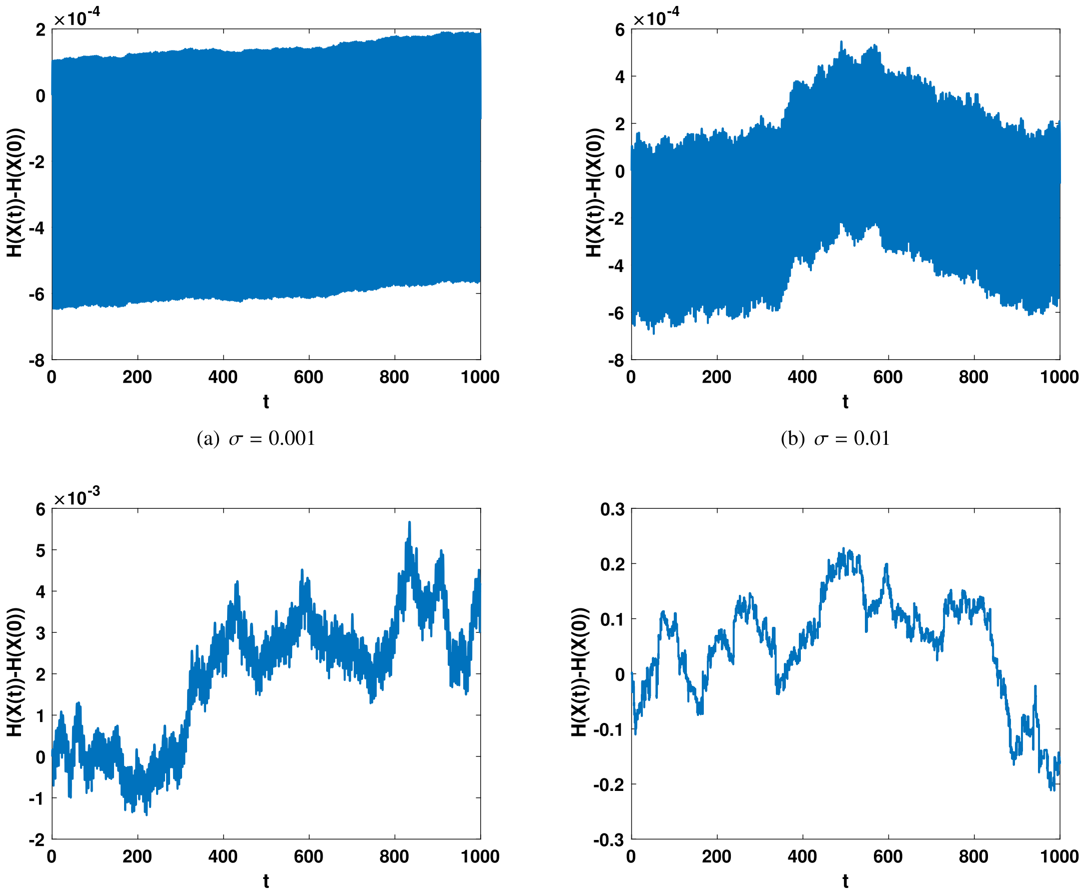
as stated in Proposition 4.2.

In Fig. 2, we plot the Hamiltonian deviations of single trajectories for method (6.63) applied to the Stratonovich Hamiltonian system (1.5) with Hamiltonian function given in (6.66), with two different stepsizes $h = 0.1$ and $h = 0.01$. We take $\sigma = 0.01$ in both cases. Theorem 4.2 states that, for reasonably small values of h, the Hamiltonian deviation remains bounded over long time. This is visible in Fig. 2, where its reduction of size h^2 ($p = 2$) is confirmed. In Fig. 3, we plot four single trajectories of (6.63), with $h = 0.1$, for selected increasing values of σ. In accordance with (4.57), we observe a growth of the Hamiltonian deviation for increasing values of σ.



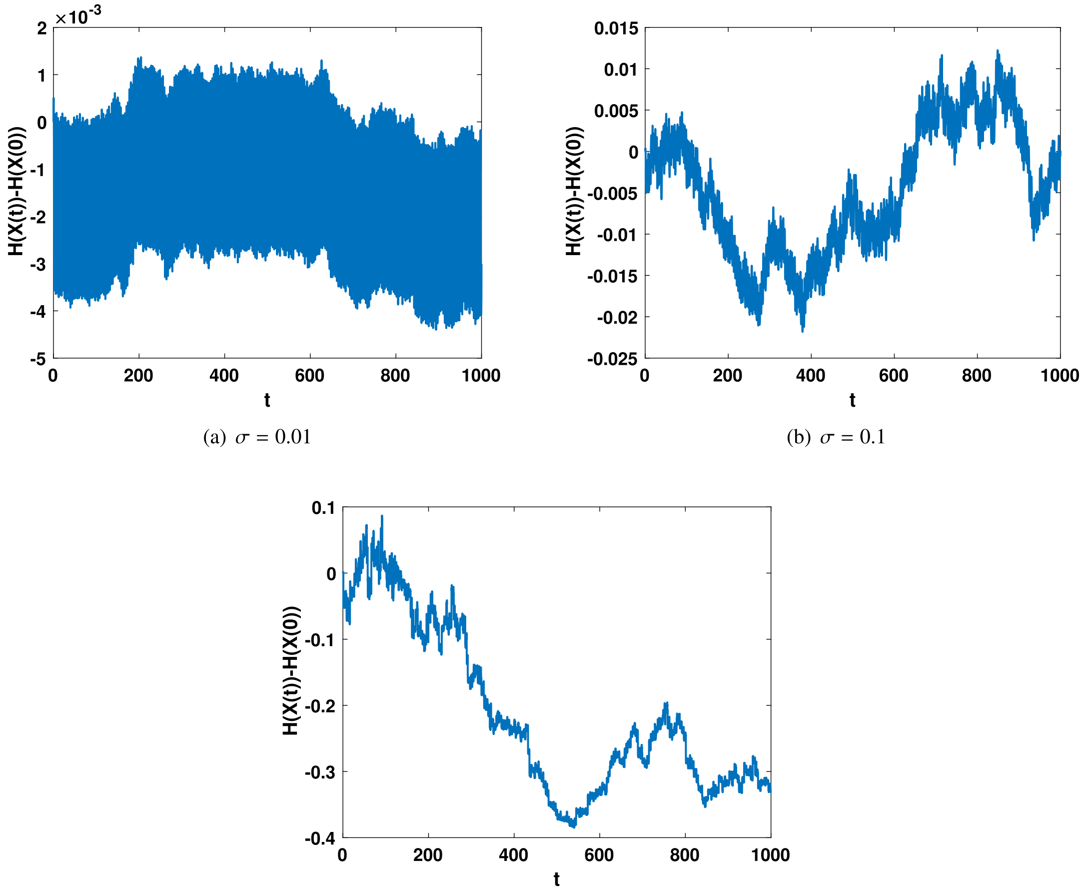
(a) $h = 0.1, \sigma = 0.01$ (b) $h = 0.01, \sigma = 0.01$

Fig. 2. Hamiltonian deviations over a single trajectory computed by the stochastic midpoint method (6.63) applied to the stochastic mathematical pendulum with Hamiltonian function $\mathcal{H}$ given in (6.66).



(a) $\sigma = 0.001$ (b) $\sigma = 0.01$ (c) $\sigma = 0.1$ (d) $\sigma = 1$

Fig. 3. Hamiltonian deviations over single trajectories computed by applying the stochastic midpoint method (6.63) applied to the stochastic mathematical pendulum with Hamiltonian function $\mathcal{H}$ given in (6.66), with $h = 0.1$, for selected values of σ.



(a) $\sigma = 0.01$

(b) $\sigma = 0.1$

(c) $\sigma = 0.5$

Fig. 4. Hamiltonian deviations over single trajectories computed by applying the stochastic midpoint method (6.63) applied to the double well potential (6.69), with $h = 0.1$, for selected values of σ.

Finally, in Fig. 4, we plot the Hamiltonian deviation for the stochastic midpoint method (6.63) applied to the Hamiltonian SDE (1.5) with *double-well potential* [13], i.e., with the quartic potential $\mathcal{V}$ given by

$$\mathcal{V}(q) = \frac{1}{4}q^4 - \frac{1}{2}q^2, \tag{6.69}$$

for different values of σ, confirming the results in Theorem 4.2.

6.2. Analysis of the stochastic trapezoidal method

Let us now consider the stochastic trapezoidal method

$$X_{n+1} = X_n + \frac{\bar{\xi}_n}{2}\Big(f(X_n) + f(X_{n+1})\Big), \qquad \bar{\xi}_n = h + \sigma\sqrt{h}Z_n, \tag{6.70}$$

that can be recovered as belonging to the family of stochastic Runge-Kutta methods for Stratonovich SDEs [11]. The computations presented in Appendix A.6 show that method (6.70) has mean-square order p equal to 2.

When applied to the Hamiltonian SDE (1.5), method (6.70) assumes the following formulation

$$X_{n+1} = X_n + \frac{\bar{\xi}_n}{2}\Big(J\nabla\mathcal{H}(X_n) + J\nabla\mathcal{H}(X_{n+1})\Big), \qquad \bar{\xi}_n = h + \sigma\sqrt{h}Z_n. \tag{6.71}$$

It is well-known (see, for instance, [27] and references therein) that, in the deterministic setting (i.e., for $\sigma = 0$), method (6.71) is not symplectic. Hence, it cannot be symplectic in the stochastic setting. Indeed, by differentiating (6.71) we have

$$\frac{\partial X_{n+1}}{\partial X_n} = I_{2d} + \frac{\bar{\xi}_n}{2}\left(J\nabla^2\mathcal{H}(X_n) + J\nabla^2\mathcal{H}(X_{n+1})\frac{\partial X_{n+1}}{\partial X_n}\right),$$

that is

$$\left(I_{2d}-\frac{\bar{\xi}_n}{2}J\nabla^2\mathcal{H}(X_{n+1})\right)\frac{\partial X_{n+1}}{\partial X_n}=I_{2d}+\frac{\bar{\xi}_n}{2}J\nabla^2\mathcal{H}(X_n),$$

leading to

$$M_{n+1}=\frac{\partial X_{n+1}}{\partial X_n}=\left(I_{2d}-\frac{\bar{\xi}_n}{2}J\nabla^2\mathcal{H}(X_{n+1})\right)^{-1}\left(I_{2d}+\frac{\bar{\xi}_n}{2}J\nabla^2\mathcal{H}(X_n)\right). \tag{6.72}$$

Similarly to [27], one can verify that the matrix M_n, $n=0,1,\dots,$ defined by Equation (6.72), does not satisfy the symplecticity condition $M_n^\intercal J M_n = J$, a.s.

However, method (6.71) is able to a.s.-preserve quadratic Hamiltonian functions of the form

$$\mathcal{H}(p,q)=\frac{1}{2}(p^2+q^2), \tag{6.73}$$

along the numerical dynamics. Indeed, for such Hamiltonian, method (6.71) reads

$$X_{n+1}=X_n+\frac{\bar{\xi}_n}{2}\left(JX_n+JX_{n+1}\right),$$

that is,

$$X_{n+1}=\left(I_{2d}-\frac{\bar{\xi}_n}{2}J\right)^{-1}\left(I_{2d}+\frac{\bar{\xi}_n}{2}J\right)X_n,$$

from which we get

$$\|X_{n+1}\|^2=\|X_n\|^2,\quad n=0,1,\dots,\qquad \text{a.s.},$$

i.e., the conservation of energy occurs. Last equality can be achieved since, for any fixed value of $\bar{\xi}_n$, the matrix $\left(I_{2d}-\frac{\bar{\xi}_n}{2}J\right)^{-1}\left(I_{2d}+\frac{\bar{\xi}_n}{2}J\right)$ is an orthogonal matrix, i.e., it preserves the Euclidean norm.

Finally, through the computations given in Appendix A.6 and taking into account (3.22), we obtain for the stochastic trapezoidal method (6.70) that

$$f_2(y)\equiv 0,\quad f_3(y)=\frac{1}{12}\left(f'(y)f'(y)f(y)+f''(y)(f(y),f(y))\right). \tag{6.74}$$

Specializing (6.74) to the case $f(y)=J\nabla H(y)$, with the same computations as in Appendix A.5, we see that

$$f_3(p,q)=\frac{1}{12}\begin{bmatrix}-p^2V'''(q)+V'(q)V''(q)\\ -V''(q)p.\end{bmatrix}$$

It is of direct verification to check that, in general, the corresponding modified system

$$d\widetilde{X}=\left(J\nabla\mathcal{H}(\widetilde{X}(t))+\xi^2f_3(\widetilde{X})\right)d\xi,\quad \xi=t+\sigma W(t), \tag{6.75}$$

is not Hamiltonian. However, if the Hamiltonian is quadratic, as in (6.73), the modified system (6.75) becomes

$$d\widetilde{X}(t)=\left(1-\frac{\xi_n^2}{12}\right)J\mathcal{H}(\widetilde{X}(t)),\quad t\in[t_n,t_{n+1}]$$

and, consequently,

$$\|\widetilde{X}(t_{n+1})\|^2=\|\widetilde{X}(t_n)\|^2,\quad n=0,1,\dots,\quad \text{a.s.},$$

as predicted by Proposition 4.1. In Fig. 5, we can observe a linear growth of the mean-square deviation for the numerical Hamiltonian over the solutions computed by the stochastic trapezoidal method (6.70), applied to the Hamiltonian given in (6.66), confirming the result stated in Theorem 5.1.

In Fig. 6, we report the mean-square Hamiltonian provided by the stochastic trapezoidal method (6.70), applied to the Hamiltonian function with potential defined in (6.69).

7. Conclusions and future developments

In this paper, we have addressed our attention to providing long-term estimates characterizing the behavior of numerical Hamiltonians along symplectic stochastic numerical discretizations, applied to stochastic Hamiltonian systems with multiplicative noise (1.5). It is well-known that the flow of such systems preserves the Hamiltonian function, so that the investigation has been oriented

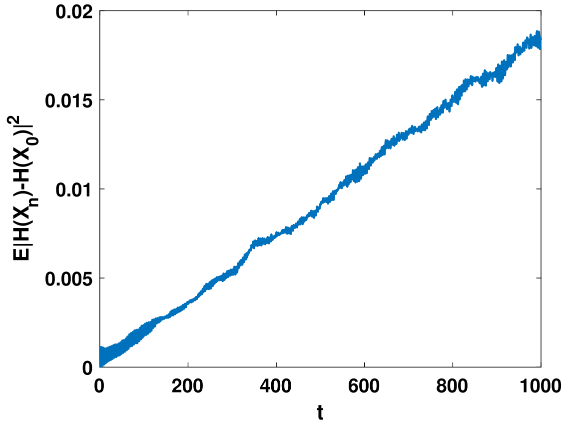


Fig. 5. Mean-square deviation of the numerical Hamiltonian provided by the stochastic trapezoidal method (6.70), applied to the Hamiltonian function defined in (6.66). Here, we have used $h = 0.5, \sigma = 0.1$ and $X(0) = \begin{bmatrix} 1 & \sqrt{2} \end{bmatrix}^\intercal$. The expectation has been performed using $M = 500$ i.i.d. paths.

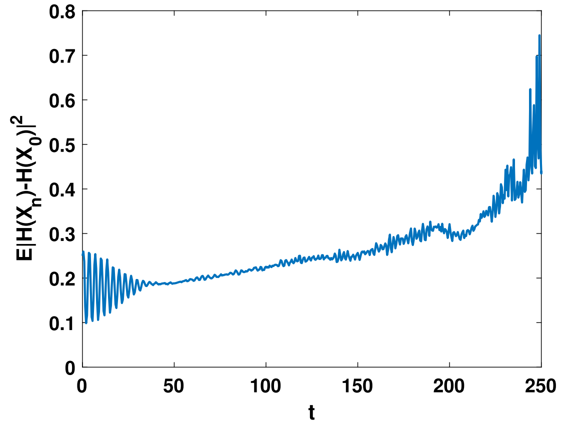


Fig. 6. Mean-square deviation of the numerical Hamiltonian provided by the stochastic trapezoidal method (6.70), applied to the Hamiltonian function with potential defined in (6.69). Here, we have used $h = 0.5, \sigma = 0.1$ and $X(0) = \begin{bmatrix} 1 & \sqrt{2} \end{bmatrix}^\intercal$. The expectation has been performed using $M = 500$ i.i.d. paths.

to the detection of this property also along numerical dynamics provided by the aforementioned discretizations. We have first constructed stochastic modified equations associated to such systems and, then, developed a strong backward error analysis, relying on such modified equations, aimed to capture the long-term behavior of the numerical Hamiltonians. The investigation has proved that stochastic perturbations (2.10) of symplectic Runge-Kutta schemes (2.11) are symplectic stochastic numerical methods. In addition, such schemes admit an associated Hamiltonian stochastic modified equation, in analogy with the deterministic setting. Moreover, for symplectic schemes, it turned out that the Hamiltonian deviation remains reasonably bounded with size $\mathcal{O}(h^p)$, where p is the mean-square order of the method, provided that the stochastic strength σ of the system is suitably small. This result provides the stochastic counterpart of the well-known Benettin-Giorgilli Theorem [5,27]. Moreover, it is worth remarking that the underlying Hamiltonian conservation is no longer visible over exponentially big time intervals as in the deterministic case. Finally, we have also proven that, for non-symplectic schemes, the mean-square Hamiltonian deviation grows linearly in time, as revealed in Theorem 5.1.

The research conducted in the present paper leaves several future glimmers open, since, as it is well-known, stochastic backward error analysis is generally a challenging ad hot topic in the recent scientific literature [49]. First of all, a backward error analysis for numerical discretizations to stochastic Poisson and Lie-Poisson systems [7,15,16] seems to be a very emergent open question. It is well-known, indeed, that the exact solution of such systems preserves the *Casimirs* along its dynamics. Hence, it would be interesting to analyze the long-term numerical conservation of these quantities along the numerical dynamics computed by proper numerical schemes.

Finally, another open issue at the best of our knowledge regards the use of strong stochastic modified equations to construct new numerical methods able of preserving Hamiltonian functions and Casimirs with a higher order of accuracy, extending what has been already done in the weak sense [1,2] and for the invariant measure [2].

Data availability

No data was used for the research described in the article.

Acknowledgements

The authors are members of the INdAM Research group GNCS. This work is supported by GNCS-INDAM project and by PRIN2017-MIUR project 2017JYCLSF "Structure preserving approximation of evolutionary problems". The authors are very grateful to the anonymous referee for the valuable report.

Appendix A

A.1. Proof of Lemma 2.1

Let us take $C(x) = x^{\intercal} D x$, with $D \in \mathbb{R}^{2d\times 2d}$ being a constant symmetric positive definite matrix. Then, according to Definition 2.2, the following condition holds true

$$x^{\intercal} D f(x) = 0, \quad \forall\, x \in \mathbb{R}^{2d}. \tag{A.1}$$

For method (2.10) applied to SDE (2.7), we have that, for $n = 0, 1, \dots,$

$$\begin{aligned} C(X_{n+1}) &= X_{n+1}^{\intercal} D X_{n+1} \\ &= \left(X_n^{\intercal} + \xi_n \sum_{i=1}^{s} b_i f(X_i)^{\intercal} \right) D \left(X_n + \xi_n \sum_{i=1}^{s} b_i f(X_i) \right) \\ &= C(X_n) + 2\xi_n \sum_{i=1}^{s} b_i f(X_i)^{\intercal} D X_n + \xi_n^2 \sum_{i,j=1}^{s} b_i b_j f(X_i)^{\intercal} D f(X_j). \end{aligned} \tag{A.2}$$

Inserting the second equation in (2.10) and taking into account (A.1), we get

$$C(X_{n+1}) = C(X_n) + \xi_n^2 \sum_{i,j=1}^{s} (b_i b_j - 2 b_i a_{ij}) f(X_i)^{\intercal} D f(X_j), \tag{A.3}$$

that gives

$$C(X_{n+1}) = C(X_n) \quad n \ge 0, \quad \text{a.s.},$$

since the summation in (A.3) is null by Assumption 2.1.

A.2. Proof of Lemma 3.3

Let us consider methods in the form (3.26). For $y \in B_{3R/2}(y_0)$ and $\|\Delta y\| \le 1$, the function $\alpha(z) = f(y + z\Delta y)$ is analytic for $|z| \le R/2$ and bounded by M. Indeed, since

$$\|y + z\Delta y - z_0\| \le 2R$$

implies

$$\|y - y_0\| + |z|\|\Delta y\| \le \frac{3R}{2} + |z|,$$

that provides the analyticity of $\alpha(z)$ for $|z| < R/2$. Then, by Cauchy's estimate, we get

$$\|\alpha'(0)\| = \|f'(y)\Delta y\| \le \frac{2M}{R}.$$

Then, we have

$$\begin{aligned} \|f'(y)\| &= \inf\{c \ge 0 : \|f'(y)\Delta y\| \le C\|\Delta y\|,\ \forall\, \Delta y\} \\ &\le \inf\{c \ge 0 : \|f'(y)\Delta y\| \le C,\ \forall\, \Delta y\} \le \frac{2M}{R}, \end{aligned}$$

i.e., we get $\|f'(y)\| \le 2M/R$ in the operator norm.
If we define the map

$$\phi : \left(X_1, \dots, X_s\right) \to \left(\phi(X)_1, \dots, \phi(X)_s\right) \tag{A.4}$$

with

$$\phi(X)_i = X_n + \bar{\xi}_n \sum_{i=1}^{s} a_{ij} f(X_j), \tag{A.5}$$

then, for $X_0 \in B_R(y_0)$ and $\gamma < 1$, the solution of method (3.26) is uniquely determined if the map (A.4)-(A.5) is a contraction map. We see that this is the case. We indeed have

$$\begin{aligned}\|\phi(X)_i - \phi(\widetilde{X})_i\| &\leq |\bar{\xi}_n| \sum_{j=1}^{s} |a_{ij}| \|f(X_j) - f(\widetilde{X}_j)\| \\ &\leq |\bar{\xi}_n| \frac{2M}{R} \sum_{j=1}^{s} |a_{ij}| \|X_j - \widetilde{X}_j\| \\ &\leq |\bar{\xi}_n| \frac{2M}{R} \max_{j=1,\ldots,s} \|X_j - \widetilde{X}_j\| \sum_{j=1}^{s} |a_{ij}|.\end{aligned}$$

Hence,

$$\|\phi(X) - \phi(\widetilde{X})\| = \max_{i=1,\ldots,s} \|\phi(X)_i - \phi(\widetilde{X}_i)\| \leq |\bar{\xi}_n| \left(\frac{2M}{R}\right) \mu \| \|X - \widetilde{X}\|.$$

In other terms, if

$$|\bar{\xi}_n| \leq \frac{\gamma R}{2\lambda M}, \tag{A.6}$$

then (A.4)-(A.5) is a contraction map. By Lemma 3.2, last condition is almost surely satisfied if

$$h + \sigma h^{1/2-\epsilon} \leq \frac{\gamma R}{2\lambda M}. \tag{A.7}$$

Moreover, since

$$\|\Phi_h(y, Z_0) - y\| = \|\bar{\xi}_0 \sum_{i=1}^{s} b_i f(g_i)\| \leq |\bar{\xi}_0| \mu M, \text{ a.s.}$$

again Cauchy's estimate applied to the function $z(\bar{\xi}_0) = \Phi_h(y, Z_0) - y$ yields

$$\|d_j(y)\| = \frac{1}{j!} \left\| \frac{d^j}{d\bar{\xi}_0^j} \left(\Phi_h(y, Z_0) - y\right)\Big|_0 \right\| \leq \frac{1}{j!} \left(\frac{2\lambda M}{\gamma R}\right)^j j! |\bar{\xi}_0| \mu M \leq \mu M \left(\frac{2\lambda M}{\gamma R}\right)^{j-1}, \text{ a.s.}$$

For $\gamma \to 1$, we get the result.

A.3. Proof of Lemma 3.4

The stochastic modified equation (3.19) can be written as

$$\dot{\widetilde{X}}(t) = F(\widetilde{X}(t)) \frac{\xi}{h}, \quad t \in [t_0, t_0 + h],$$

where $F(y) = f(y) + \xi_0 f_2(y) + \xi_0^2 f_3(y) + \ldots$. Hence, the stochastic process $\widetilde{X}(t)$ admits the following expansion

$$\begin{aligned}\widetilde{X}(t_0 + h) &= X_0 + hF(X_0)\frac{\xi_0}{h} + \frac{h^2}{2!} F(X_0)' \frac{\xi_0}{h} + \frac{h^3}{3!} F(X_0)'' \frac{\xi_0}{h} + \ldots \\ &= X_0 + \xi_0 F(X_0) + \frac{\xi_0^2}{2!} F'(X_0) F(X_0) + \frac{\xi_0^3}{3!} \Big(F'(X_0) F(X_0) \Big)' F(X_0) + \ldots \\ &= X_0 + \xi_0 D^0 F(X_0) + \frac{\xi_0^2}{2!} (DF)(X_0) + \frac{\xi_0^3}{3!} \left(D^2 F\right)(X_0) + \ldots .\end{aligned}$$

Hence, the process $\widetilde{X}(t)$ satisfies

$$\widetilde{X}(t_0 + h) = X_0 + \sum_{i \geq 1} \frac{\xi_0^i}{i!} D^{i-1} F(X_0),$$

being $D = D_1 + \xi D_2 + \xi_2 D_3 + \ldots$. Moreover, we get

$$
\begin{aligned}
\widetilde{X}(t_0+h) &= X_0 + \sum_{i\geq 1} \frac{\xi_0^i}{i!}\left(\sum_{j\geq 1}\xi_0^{j-1}D_j\right)^{i-1} F(X_0)\\
&= X_0 + \sum_{i\geq 1} \frac{\xi_0^i}{i!} \sum_{k_1,\dots,k_{i-1}} \xi_0^{k_1+\cdots+k_{i-1}-i+1} D_{k_1}\cdots D_{k_{i-1}} F(X_0)\\
&= X_0 + \sum_{i\geq 1} \frac{1}{i!} \sum_{k_1,\dots,k_{i-1}} \xi_0^{k_1+\cdots+k_{i-1}+1} D_{k_1}\cdots D_{k_{i-1}} \sum_{k_i}\xi_0^{k_i-1} f_{k_i}(X_0)\\
&= X_0 + \sum_{i\geq 1} \frac{1}{i!} \sum_{k_1,\dots,k_i} \xi_0^{k_1+\dots k_i}\left(D_{k_1}\cdots D_{k_{i-1}} f_{k_i}\right)(X_0),
\end{aligned}
$$

where all $k_m \geq 1$. Since $\widetilde{X}(t_0+h) = \Phi_h(X_0, \Delta W_0)$, comparing the same powers of ξ_0 in (3.21) and (3.19), leads to the thesis.

A.4. Proof of Lemma 4.1

Assumption 3.1 and the estimate (3.31) give

$$
\begin{aligned}
\|F_N(y)\| &= \|f(y) + \sum_{j=2}^{N} \bar{\xi}_0^{j-1} f_j(y)\| \leq \|f(y)\| + \sum_{j=2}^{N} |\bar{\xi}_0|^{j-1}\|f_j(y)\|\\
&\leq M\left(1+\ln 2\,\eta\sum_{j=2}^{N}\left(\frac{\eta M|\bar{\xi}_0|j}{R}\right)^{j-1}\right), \quad \text{a.s.}
\end{aligned}
$$

Since $|\bar{\xi}_0|/(eh_0) \leq 1/N$, a.s., we have

$$
\|F_N(y)\| \leq M\left(1+\ln 2\,\eta\sum_{j=2}^{N}\left(\frac{j}{N}\right)^{j-1}\right), \quad \text{a.s.,}
$$

that gives the first result in (4.41) [27].
Moreover, by Theorem 3.1, for a method of mean-square order p, we have

$$
F_N(y) = f(y) + \bar{\xi}_0^p f_{p+1}(y) + \bar{\xi}_0^{p+1} f_{p+2}(y) + \cdots + \bar{\xi}_0^{N-1} f_N(y).
$$

Then,

$$
\|F_N(y) - f(y)\| \leq |\bar{\xi}_0|^p \sum_{j=1}^{N-p} |\bar{\xi}_0|^{j-1} f_{p+j}(y)\| \leq \ln 2\,\eta M|\bar{\xi}_0|^p \sum_{j=1}^{N-p} |\bar{\xi}_0|^{j-1}\left(\frac{\eta M(p+j)}{R}\right)^{p+j-1}, \text{ a.s.}
$$

The second result in (4.41) also follows, since

$$
\begin{aligned}
\sum_{j=1}^{N-p} |\bar{\xi}_0|^{j-1}\left(\frac{\eta M(p+j)}{R}\right)^{p+j-1} &= \sum_{j=1}^{N-p}\left(\frac{\eta M|\bar{\xi}_0|(p+j)}{R}\right)^{j-1}\left(\frac{\eta M(p+j)}{R}\right)^{p}\\
&\leq \sum_{j=1}^{N-p}\left(\frac{p+j}{N}\right)^{j-1}\left(\frac{\eta M(p+j)}{R}\right)^{p} = \text{const.}
\end{aligned}
$$

A.5. Stochastic modified equation for the implicit midpoint rule

We here report the calculations needed to determine the first-order modified equation for the stochastic implicit midpoint method (6.63), for separable Hamiltonians of the form

$$
\mathcal{H}(p,q) = \frac{1}{2}p^{\intercal}p + \mathcal{V}(q), \tag{A.8}
$$

being $\mathcal{V}: \mathbb{R}^d \to \mathbb{R}$ a smooth potential. For the sake of simplicity, in the sequel, we consider $d = 1$.

We first perform a Taylor expansion of (6.63), to determine the coefficients d_j defined in (3.21). Then, over a single step, the stochastic midpoint rule (6.63) applied to (1.5) with $\mathcal{H}$ as in (A.8) reads

$$
p_1 = p - \bar{\xi}_0 V'\left(\frac{q+q_1}{2}\right), \quad q_1 = q + \bar{\xi}_0\left(\frac{p+p_1}{2}\right), \tag{A.9}
$$

where we have denoted by $p = p(t_0)$ and $q = q(t_0)$. Then, we have

$$\begin{aligned}V'\left(\frac{q+q_1}{2}\right) &= V'(q)+\frac{1}{2}V''(q)(q_1-q)+\frac{1}{8}V'''(q)(q_1-q)^2+\mathcal{O}\left((q_1-q)^3\right)\\ &= V'(q)+\frac{\bar{\xi}_0}{4}V''(q)\left(p+p_1\right)+\frac{\bar{\xi}_0^2}{32}V'''(q)(p+p_1)^2+\mathcal{O}\left(\bar{\xi}_0^3\right)\\ &= V'(q)+\frac{\bar{\xi}_0}{2}V''(q)p+\frac{\bar{\xi}_0^2}{4}\left[\frac{1}{2}V'''(q)p^2-V''(q)V'(q)\right]+\mathcal{O}\left(\bar{\xi}_0^3\right).\end{aligned}$$

Hence, we get

$$\begin{aligned}p_1 &= p-\bar{\xi}_0V'(q)-\frac{\bar{\xi}_0^2}{2}V''(q)p-\frac{\bar{\xi}_0^3}{4}\left[\frac{1}{2}V'''(q)p^2-V''(q)V'(q)\right]+\mathcal{O}\left(\bar{\xi}_0^4\right),\\ q_1 &= q+\bar{\xi}_0p-\frac{\bar{\xi}_0^2}{2}V'(q)-\frac{\bar{\xi}_0^3}{4}V''(q)p+\mathcal{O}\left(\bar{\xi}_0^4\right).\end{aligned}$$

Correspondingly, for method (A.9), we obtain

$$d_2(p,q)=-\frac{1}{2}\begin{bmatrix}V''(q)p\\ V'(q)\end{bmatrix},\quad d_3(q,p)=-\frac{1}{4}\begin{bmatrix}\frac{1}{2}V'''(q)p^2-V''(q)V'(q)\\ V''(q)p\end{bmatrix}.$$

Since

$$f'(p,q)f(p,q)=\begin{bmatrix}0 & -V''(q)\\ 1 & 0\end{bmatrix}\begin{bmatrix}-V'(q)\\ p\end{bmatrix}=-\begin{bmatrix}V''(q)p\\ V'(q)\end{bmatrix},$$

taking into account (3.22), we immediately obtain

$$f_2\equiv 0,\quad f_3=d_3-\frac{1}{3!}\left(f''(f,f)+f'f'f\right).$$

We note that the method (6.63) has a mean-square order $p=2$. This already implies $f_2\equiv 0$. Next, we compute the term $f''(f,f)$ componentwise, i.e.,

$$\left(f''(f,f)\right)_1=\sum_{j,k=1}^{2}\frac{\partial^2 f_1}{\partial y_j\partial y_k}f_jf_k,\quad \left(f''(f,f)\right)_2=\sum_{j,k=1}^{2}\frac{\partial^2 f_2}{\partial y_j\partial y_k}f_jf_k,$$

where

$$f_1=-V'(q),\quad f_2=p,\quad y_1=p,\quad y_2=q.$$

Hence, we have

$$\left(f''(f,f)\right)_1=\frac{\partial^2 f_1}{\partial y_1^2}f_1^2+2\frac{\partial^2 f_1}{\partial y_1\partial y_2}f_1f_2+\frac{\partial^2 f_1}{\partial y_2^2}f_2^2=-p^2V'''(q)$$

and

$$\left(f''(f,f)\right)_2=\frac{\partial^2 f_2}{\partial y_1^2}f_1^2+2\frac{\partial^2 f_2}{\partial y_1\partial y_2}f_1f_2+\frac{\partial^2 f_2}{\partial y_2^2}f_2^2=\frac{\partial^2 f_2}{\partial y_1^2}f_1^2=0.$$

Moreover, we compute

$$f'f'f=\begin{bmatrix}0 & -V''(q)\\ 1 & 0\end{bmatrix}\begin{bmatrix}-V''(q)p\\ -V'(q)\end{bmatrix}=\begin{bmatrix}V'(q)V''(q)\\ -V''(q)p\end{bmatrix}.$$

Then, we finally get

$$f_3=\begin{bmatrix}\frac{1}{24}p^2V'''(q)+\frac{1}{12}V''(q)V'(q)\\ -\frac{1}{12}pV''(q)\end{bmatrix}.$$

A.6. Taylor expansion for the stochastic trapezoidal method

We finally present the computations needed to determine the coefficients $d_j(y),\ j=2,3,\dots,$ for the trapezoidal method (6.70). Over single step, it reads

$$X_1=X_0+\frac{\bar{\xi}_0}{2}\Big(f(X_0)+f(X_1)\Big). \tag{A.10}$$

We now expand the term $f(X_1)$, as follows:

$$
\begin{aligned}
f(X_1) = f\left(X_0 + \frac{\bar{\xi}_0}{2}\big(f(X_0)+f(X_1)\big)\right) &= f(X_0) + \frac{\bar{\xi}_0}{2} f'(X_0)\Big(f(X_0)+f(X_1)\Big) \\
&\quad + \frac{\bar{\xi}_0^2}{8} f''(X_0)\Big(f(X_0)+f(X_1), f(X_0)+f(X_1)\Big) + \mathcal{O}\left(\bar{\xi}_0^3\right) \\
&= f(X_0) + \frac{\bar{\xi}_0}{2} f'(X_0)f(X_0) + \frac{\bar{\xi}_0}{2} f'(X_0) f\left(X_0 + \frac{\bar{\xi}_0}{2}\big(f(X_0)+f(X_1)\big)\right) \\
&\quad + \frac{\bar{\xi}_0^2}{2} f''(X_0)\Big(f(X_0), f(X_0)\Big) + \mathcal{O}\left(\bar{\xi}_0^3\right) \\
&= f(X_0) + \bar{\xi}_0 f'(X_0)f(X_0) + \frac{\bar{\xi}_0^2}{4} f'(X_0)f'(X_0)\Big(f(X_0)+f(X_1)\Big) \\
&\quad + \frac{\bar{\xi}_0^2}{2} f''(X_0)\Big(f(X_0), f(X_0)\Big) + \mathcal{O}\left(\bar{\xi}_0^3\right).
\end{aligned}
\tag{A.11}
$$

We finally achieve

$$
f(X_1) = f(X_0) + \bar{\xi}_0 f'(X_0)f(X_0) + \frac{\bar{\xi}_0^2}{2}\Big(f'(X_0)f'(X_0)f(X_0) + f''(X_0)\big(f(X_0), f(X_0)\big)\Big) + \mathcal{O}\left(\bar{\xi}_0^3\right)
$$

leading to the following expansion

$$
X_1 = X_0 + \bar{\xi}_0 f(X_0) + \frac{\bar{\xi}_0^2}{2} f'(X_0)f(X_0) + \frac{\bar{\xi}_0^3}{4}\Big(f'(X_0)f'(X_0)f(X_0) + f''(X_0)\big(f(X_0), f(X_0)\big)\Big) + \mathcal{O}\left(\bar{\xi}_0^4\right),
$$

from which we obtain

$$
d_2(y) = \frac{1}{2!} f'(y)f(y), \; d_3 = \frac{1}{4}\Big(f'(y)f'(y)f(y) + f''(y)\big(f(y), f(y)\big)\Big).
$$

References

[1] A. Abdulle, D. Cohen, G. Vilmart, K.C. Zygalakis, High weak order methods for stochastic differential equations based on modified equations, SIAM J. Sci. Comput. 34 (2012) A1800–A1823.
[2] A. Abdulle, G. Vilmart, K.C. Zygalakis, High order numerical approximation of the invariant measure of ergodic SDEs, SIAM J. Sci. Comput. 52 (2014) 1600–1622.
[3] C. Anton, Weak backward error analysis for stochastic Hamiltonian systems, BIT Numer. Math. 59 (2019) 613–646.
[4] A. Bazzani, S. Siboni, G. Turchetti, Diffusion in Hamiltonian systems with a small stochastic perturbation, Physica D 76 (1994) 8–21.
[5] G. Benettin, A. Giorgilli, On the Hamiltonian interpolation of near to the identity symplectic mappings with application to symplectic integration algorithms, J. Stat. Phys. 74 (1994) 1117–1143.
[6] S. Blanes, F. Casas, A Concise Introduction to Geometric Numerical Integration, CRC Press, New York, 2016.
[7] C.E. Brehier, D. Cohen, T. Jahnke, Splitting integrators for stochastic Lie-Poisson systems, Math. Comput. 92 (2023) 2167–2216.
[8] L. Brugnano, F. Iavernaro, Line Integral Methods for Conservative Problems, CRC Press, New York, 2016.
[9] E. Buckwar, A. Rössler, R. Winkler, Stochastic Runge-Kutta methods for Itô SODEs with small noise, SIAM J. Sci. Comput. 32 (2010) 1789–1808.
[10] K. Burrage, P.M. Burrage, Low rank Runge-Kutta methods, symplecticity and stochastic Hamiltonian problems with additive noise, J. Comput. Appl. Math. 236 (2012) 3920–3930.
[11] K. Burrage, P.M. Burrage, Order conditions of stochastic Runge–Kutta methods by B-series, SIAM J. Numer. Anal. 38 (2000) 1626–1646.
[12] P. Chartier, E. Hairer, G. Vilmart, Algebraic structures of B-series, Found. Comput. Math. 10 (4) (2010) 407–427.
[13] C. Chen, D. Cohen, R. D'Ambrosio, A. Lang, Drift-preserving numerical integrators for stochastic Hamiltonian systems, Adv. Comput. Math. 46 (27) (2020).
[14] C. Chen, D. Cohen, J. Hong, Conservative methods for stochastic differential equations with conserved quantity, Int. J. Numer. Anal. Model. 13 (2016) 435–456.
[15] D. Cohen, G. Dujardin, Energy-preserving integrators for stochastic Poisson systems, Commun. Math. Sci. 12 (2014) 1523–1539.
[16] D. Cohen, G. Vilmart, Drift-preserving numerical integrators for stochastic Poisson systems, Int. J. Comput. Math. 99 (2022) 4–20.
[17] R. D'Ambrosio, Numerical Approximation of Ordinary Differential Problems. From Deterministic to Stochastic Numerical Methods, Springer, 2023.
[18] R. D'Ambrosio, S. Di Giovacchino, Long-term analysis of stochastic Hamiltonian systems under time discretizations, SIAM J. Sci. Comput. 45 (2023) 257–A 288.
[19] R. D'Ambrosio, S. Di Giovacchino, Numerical conservation issues for the stochastic Korteweg-de Vries equation, J. Comput. Appl. Math. 424 (2023) 114967.
[20] R. D'Ambrosio, S. Di Giovacchino, G. Giordano, B. Paternoster, On the conservative character of discretizations to Itô-Hamiltonian systems with small noise, Appl. Math. Lett. 138 (2023) 108529.
[21] R. D'Ambrosio, S. Di Giovacchino, How do Monte Carlo estimates affect stochastic geometric numerical integration?, Int. J. Comput. Math. 100 (1) (2023) 192–208.
[22] A. Debussche, E. Faou, Weak backward error analysis, SIAM J. Numer. Anal. 50 (2012) 1735–1752.
[23] J. Deng, Strong backward error analysis for Euler-Maruyama method, Int. J. Numer. Anal. Model. 13 (1) (2016) 1–21.
[24] S.D. Giovacchino, D.J. Higham, K. Zygalakis, Backward error analysis and the qualitative behaviour of stochastic optimization algorithms: application to stochastic coordinate descent, submitted for publication, arXiv:2309.02081, 2023.
[25] T.C. Gard, Introduction to Stochastic Differential Equations, Marcel Dekker Inc., New York-Basel, 1988.
[26] C.W. Gardiner, Handbook of Stochastic Methods, for Physics, Chemistry and the Natural Sciences, Springer-Verlag, 2004.
[27] E. Hairer, C. Lubich, G. Wanner, Geometric Numerical Integration. Structure-Preserving Algorithms for Ordinary Differential Equations, second edition, Springer, 2006.
[28] D.F. Griffiths, D.J. Higham, Numerical Methods for Ordinary Differential Equations: Initial Value Problems, Springer, 2010.
[29] D.J. Higham, P.E. Kloeden, An Introduction to the Numerical Simulation of Stochastic Differential Equations, SIAM, 2021.

[30] P.E. Kloeden, E. Platen, Numerical Solution of Stochastic Differential Equations, part of Applications of Mathematics, vol. 23, Springer-Verlag, Berlin, 1992.
[31] M. Kopec, Weak backward error analysis for Langevin process, BIT Numer. Math. 55 (4) (2015) 1057–1103.
[32] M. Kopec, Weak backward error analysis for overdamped Langevin processes, IMA J. Numer. Anal. 35 (2) (2015) 583–614.
[33] J. Hong, L. Sun, Symplectic Integration of Stochastic Hamiltonian Systems, Springer, 2022.
[34] A. Iserles, A First Course in the Numerical Analysis of Differential Equations, second edition, Cambridge University Press, Cambridge, 2008.
[35] F. Kang, Q. Mengzhao, Symplectic Geometric Algorithms for Hamiltonian Systems, Springer, Berlin, Heidelberg, 2010.
[36] A. Laurent, H.Z. Munthe-Kaas, The universal equivariance properties of exotic aromatic B-series, submitted for publication, arXiv:2305.10993, 2023.
[37] A. Laurent, G. Vilmart, Exotic aromatic B-series for the study of long time integrators for a class of ergodic SDEs, Math. Comput. 89 (321) (2020) 169–202.
[38] J.A. Lazaro-Cami', J.P. Ortega, Stochastic Hamiltonian dynamical systems, Rep. Math. Phys. 61 (2008) 65–122.
[39] B. Leimkuhler, S. Reich, Geometric Integrators in Hamiltonian Mechanics, Cambridge University Press, Cambridge, 2003.
[40] B. Leimkuhler, S. Reich, Simulating Hamiltonian Dynamics, Cambridge University Press, Cambridge, 2005.
[41] Q. Ma, D. Ding, X. Ding, Symplectic conditions and stochastic generating functions of stochastic Runge-Kutta methods for stochastic Hamiltonian systems with multiplicative noise, Appl. Math. Comput. 219 (2012) 635–643, Theory Probab. Appl. 30 (1986) 750–766.
[42] G.N. Milstein, Y.M. Repin, M.V. Tretyakov, Numerical methods for stochastic systems preserving symplectic structure, SIAM J. Numer. Anal. 40 (2002) 1583–1604.
[43] G.N. Milstein, Yu.M. Repin, M.V. Tretyakov, Symplectic integration of Hamiltonian systems with additive noise, SIAM J. Numer. Anal. 39 (2002) 2066–2088.
[44] G.N. Milstein, M.V. Tretyakov, Stochastic Numerics for Mathematical Physics, Scientic Computation, Springer-Verlag, Berlin, 2004.
[45] T. Misawa, Energy conservative stochastic difference scheme for stochastic H dynamical systems, Jpn. J. Ind. Appl. Math. 17 (2000) 119–128.
[46] A. Rössler, Second order Runge-Kutta methods for Itô stochastic differential equations, SIAM J. Numer. Anal. 47 (2009) 1713–1738.
[47] A. Rössler, Runge-Kutta methods for Itô stochastic differential equations with scalar noise, BIT Numer. Math. 46 (2006) 97–110.
[48] J.M. Sanz-Serna, M.P. Calvo, Numerical Hamiltonian Problems, Chapman & Hall, 1994.
[49] T. Shardlow, Modified equations for stochastic differential equations, BIT Numer. Math. 46 (2006) 111–125.
[50] A.M. Stuart, A.R. Humphries, Dynamical Systems and Numerical Analysis, part of Cambridge Monographs on Applied and Computational Mathematics, Cambridge University Press, 1999.
[51] D. Talay, Stochastic Hamiltonian systems: exponential convergence to the invariant measure, and discretization by the implicit Euler scheme, Markov Process. Relat. Fields 8 (2002) 1–36.
[52] Z. Zhang, G.E. Karniadakis, Numerical Methods for Stochastic Partial Differential Equations with White Noise, Springer, 2017.
[53] K.C. Zygalakis, On the existence and the applications of modified equations for stochastic differential equations, SIAM J. Sci. Comput. 33 (2011) 102–130.