Applied Numerical Mathematics 129 (2018) 120–136

Contents lists available at ScienceDirect

Applied Numerical Mathematics

www.elsevier.com/locate/apnum

Symplectic Runge–Kutta methods for Hamiltonian systems driven by Gaussian rough paths

Jialin Hong [a,b], Chuying Huang [a,b,*], Xu Wang [a,b]

[a] LSEC, ICMSEC, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China
[b] School of Mathematical Sciences, University of Chinese Academy of Sciences, Beijing 100049, China

ARTICLE INFO

Article history:
Received 25 November 2017
Received in revised form 9 March 2018
Accepted 9 March 2018
Available online 14 March 2018

Keywords:
Rough path
Hamiltonian system
Symplectic Runge–Kutta method
Implicit method
Pathwise convergence rate

ABSTRACT

We consider Hamiltonian systems driven by multi-dimensional Gaussian processes in rough path sense, which include fractional Brownian motions with Hurst parameter $H \in (1/4, 1/2]$. We prove that the phase flow preserves the symplectic structure almost surely and this property could be inherited by symplectic Runge–Kutta methods, which are implicit methods in general. If the vector fields satisfy some smoothness and boundedness conditions, we obtain the pathwise convergence rates of Runge–Kutta methods. When vector fields are linear, we get the solvability of the midpoint scheme for skew symmetric cases, and obtain its pathwise convergence rate. Numerical experiments verify our theoretical analysis.

1. Introduction

Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space. We consider a stochastic differential equation (SDE)

$$dY_t = V(Y_t)dX_t, \quad t \in [0, T] \tag{1}$$

driven by multi-dimensional Gaussian process in rough path sense. Under proper assumptions, X can be lifted to a rough path almost surely and then (1) can be interpreted as a rough differential equation (RDE). For example, the lift of the fractional Brownian motion (fBm) with Hurst parameter $H > 1/4$ is constructed by piecewise linear approximations and their iterated integrals. We refer to [18,10,8,6] and references therein for more details.

The well-posedness of the RDE is given originally in [18] for the case that the vector fields belong to Lip^{γ} (see Definition 2.2), i.e., they are bounded and smooth enough with bounded derivatives. When the vector fields are linear, the equation is still well-posed (see e.g., [8]). In particular, the solution is equivalent to that in Stratonovich sense almost surely when the noise is a semi-martingale. The robustness of the solution allows numerous researches to be developed, such as the density and ergodicity of SDEs driven by non-Markovian noises (see [3,12] and references therein) and the theory of regularity structures for stochastic partial differential equations (see e.g., [6]).

The author in [4] develops an approximation approach to deal with the equation driven by non-differentiable paths. The approximations are further investigated and called the step-N Euler schemes in [8]. When the equation is driven by fBm with $H > 1/2$, modified Euler scheme, Taylor schemes and Crank–Nicolson scheme are analyzed in [15,16,14], respectively. When the noise is a Brownian motion, the pathwise convergence rate of an adaptive time-stepping Euler–Maruyama

* Corresponding author.
E-mail addresses: hjl@lsec.cc.ac.cn (J. Hong), huangchuying@lsec.cc.ac.cn (C. Huang), wangxu@lsec.cc.ac.cn (X. Wang).

https://doi.org/10.1016/j.apnum.2018.03.006
0168-9274/

method is given in [22]. For rougher noise, the simplified step-N Euler schemes [5] and modified Euler scheme [17] for fBm with $H \in (1/3, 1/2]$ are proposed to avoid simulating iterated integrals of the noise. Authors in [7] give the estimate of the pathwise convergence rate of Wong–Zakai approximations for Gaussian process under proper assumptions and obtain pathwise convergence rates of the simplified step-N Euler schemes. The results are applicable for fBm with $H \in (1/4, 1/2]$. Further, the $L^r(\Omega)$-convergence rates are obtained in [2] to reduce the complexity of multilevel Monte Carlo.

From the perspective of modeling, Gaussian noises with nontrivial correlations are more general than standard Brownian motions. For a dynamical system influenced by additional nonconservative forces in [9], which could be generally assumed as Gaussian noises, its pathwise movement rule is described by Hamiltonian equations in rough path sense. In this paper, we are concerned about a Hamiltonian system driven by multi-dimensional Gaussian process under proper assumption. The fBm with $H \in (1/4, 1/2]$ satisfies our assumption with $\rho = \frac{1}{2H}$, where ρ is a parameter related to the regularity of the noise. We prove that rough Hamiltonian system, as a generalization of the deterministic case (see [11] and references therein) and the stochastic case driven by standard Brownian motion (see e.g., [21,20]), has the characteristic property of Hamiltonian system that its phase flow preserves the symplectic structure almost surely. The symplecticity is naturally considered to be inherited by numerical methods. The simplified step-N Euler schemes, which are explicit, can not possess the discrete symplectic structure in general.

Based on the two motivations—preserving the symplectic structure of the original rough Hamiltonian system and avoiding simulating iterated integrals of the noises, we investigate the symplectic Runge–Kutta methods. For the standard Brownian setting, we refer to [1,23] and references therein for more related works. Since these methods are implicit in general, the solvability of the numerical methods should be taken into consideration, which is one of the differences between explicit methods and implicit methods. For SDEs driven by standard Brownian motions, authors in [20] use truncate technique to give the solvability of implicit methods and convergence rates in mean square sense. For Gaussian processes with nontrivial covariance, the truncation technique is not suitable and more conditions should be put on the vector fields. In [13], we prove the solvability of implicit Runge–Kutta methods when vector fields are bounded, which includes the case that they belong to Lip^γ. If the vector fields are linear, we suppose additionally that they are skew symmetric in Section 4, and prove the solvability of the 1-stage symplectic Runge–Kutta method—the midpoint scheme.

Furthermore, we analyze pathwise convergence rates for Runge–Kutta methods. For the Lip^γ case, we obtain that the convergence rate of the midpoint scheme is $(\frac{3}{2\rho} - 1 - \varepsilon)$, for arbitrary small ε and $\rho \in [1, 3/2)$. The convergence rate of another two symplectic Runge–Kutta methods, whose stages are higher, is $(\frac{1}{\rho} - \frac{1}{2} - \varepsilon)$ for $\rho \in [1, 2)$. This is limited by the convergence rate of Wong–Zakai approximations, similar to the simplified step-3 Euler scheme in [7]. If the vector fields are linear and skew symmetric, we prove the uniform boundedness of numerical solutions for the midpoint scheme and have that its convergence rate is also $(\frac{3}{2\rho} - 1 - \varepsilon)$ when $\rho \in [1, 3/2)$.

This paper is organized as follows. In Section 2, we introduce notations and definitions in rough path theory as well as the result of the well-posedness of RDEs, which contains both the Lip^γ case and the linear case. In Section 3, we derive Hamiltonian equations with additional nonconservative forces through the variational principle to obtain the rough Hamiltonian system and prove that the phase flow preserves the symplectic structure almost surely. In Section 4, we propose symplectic Runge–Kutta methods for rough Hamiltonian systems and give the solvability of them. In Section 5, we analyze pathwise convergence rates for symplectic Runge–Kutta methods. Numerical experiments in Section 6 are presented to verify our theoretical convergence rate and show the stability of implicit methods.

2. Preliminaries in rough path theory

We first recall some notations in rough path theory. For $p \in [1, \infty)$, we are interested in continuous maps $\mathbf{X} : [0, T] \to G^{[p]}(\mathbb{R}^d)$, where $[p]$ is an integer satisfying $p - 1 < [p] \le p$ and $G^{[p]}(\mathbb{R}^d)$ is the free step-$[p]$ nilpotent Lie group of $\mathbb{R}^d$, equipped with the Carnot–Carathéodory metric d.

Since the group $G^{[p]}(\mathbb{R}^d)$ is embedded in the truncated step-$[p]$ tensor algebra, i.e., $G^{[p]}(\mathbb{R}^d) \subset \bigoplus_{n=0}^{[p]} (\mathbb{R}^d)^{\otimes n}$ with $(\mathbb{R}^d)^{\otimes 0} = \mathbb{R}$, the increment of $\mathbf{X}$ is defined by $\mathbf{X}_{s,t} := \mathbf{X}_s^{-1} \otimes \mathbf{X}_t$. Denoting by $\pi_1\big(G^{[p]}(\mathbb{R}^d)\big)$ the projection of $G^{[p]}(\mathbb{R}^d)$ into $\mathbb{R}^d$, we have $\pi_1(\mathbf{X}_{s,t}) = \pi_1(\mathbf{X}_t) - \pi_1(\mathbf{X}_s)$.

Define the p-variation of $\mathbf{X}$ (see e.g., [18]) by

$$\|\mathbf{X}\|_{p\text{-}var;[s,t]} := \sup_{(t_k)\in\mathscr{D}([s,t])} \left(\sum_{t_k} \mathrm{d}(\mathbf{X}_{t_k}, \mathbf{X}_{t_{k+1}})^p \right)^{1/p},$$

where $\mathscr{D}([s,t])$ is the set of dissections of $[s,t]$. We call $\mathbf{X}$ a weak geometric p-rough path if

$$\mathbf{X} \in \mathscr{C}^{p\text{-}var}\big([0,T], G^{[p]}(\mathbb{R}^d)\big) := \{\mathbf{X} : \|\mathbf{X}\|_{p\text{-}var;[0,T]} < \infty\},$$

where the set $\mathscr{C}^{p\text{-}var}\big([0,T], G^{[p]}(\mathbb{R}^d)\big)$ contains drivers of RDEs. In addition, we say $\mathbf{X}$ is of Hölder-type if

$$\|\mathbf{X}\|_{1/p\text{-Höl};[0,T]} := \sup_{0\le s<t\le T} \frac{\mathrm{d}(\mathbf{X}_s, \mathbf{X}_t)}{|t-s|^{1/p}} < \infty.$$

This implies $|\pi_1(\mathbf{X}_{s,t})| \le \|\mathbf{X}\|_{1/p\text{-Höl};[0,T]}|t-s|^{1/p}$ for any $0 \le s < t \le T$. Similar to the $1/p$-Hölder continuity in classical case, a larger p implies lower regularity of $\mathbf{X}$.

The solution of an RDE is constructed by means of a sequence of bounded variation functions on $\mathbb{R}^d$, which are appropriate approximations of $\mathbf{X}$. Let $x : [0,T] \to \mathbb{R}^d$ be a bounded variation function, then we consider its canonical lift to a weak geometric p-rough path defined by $S_{[p]}(x)$ with

$$S_{[p]}(x)_t := \left(1, \int\limits_{0\le u_1\le t} dx_{u_1}, \cdots, \int\limits_{0\le u_1<\cdots<u_{[p]}\le t} dx_{u_1}\otimes\cdots\otimes dx_{u_{[p]}}\right) \in G^{[p]}(\mathbb{R}^d).$$

One can observe that $S_{[p]}(x)$ contains information about iterated integrals of x up to $[p]$-level, which corresponds to the regularity of $\mathbf{X}$. If $S_{[p]}(x)$ is close to $\mathbf{X}$ in p-variation sense, then x is a proper approximation of $\mathbf{X}$. In other words, the rougher $\mathbf{X}$ is, the more information is required.

We now introduce the definition of the solution of an RDE

$$dY_t = V(Y_t)d\mathbf{X}_t, \quad Y_0 = z \in \mathbb{R}^m, \tag{2}$$

where $V = (V_1, \cdots, V_d) : \mathbb{R}^m \to \mathbb{R}^{m\times d}$ is a collection of vector fields on $\mathbb{R}^m$.

Definition 2.1. (see e.g., [8]) Let $p \in [1,\infty)$ and $\mathbf{X} \in \mathscr{C}^{p\text{-}var}\big([0,T], G^{[p]}(\mathbb{R}^d)\big)$. Suppose there exists a sequence of bounded variation functions $\{x^n\}_{n=1}^\infty$ on $\mathbb{R}^d$ such that

$$\sup_{n\in\mathbb{N}} \|S_{[p]}(x^n)\|_{p\text{-}var;[0,T]} < \infty \quad \text{and} \quad \lim_{n\to\infty} \sup_{0\le s<t\le T} \mathrm{d}\big(S_{[p]}(x^n)_{s,t}, \mathbf{X}_{s,t}\big) = 0.$$

Suppose in addition $\{y^n\}_{n=1}^\infty$ are solutions of equations $dy_t^n = V(y_t^n)dx_t^n$, $y_0^n = z$, in the Riemann–Stieltjes integral sense. If y_t^n converges to Y_t in the $L^\infty([0,T])$-norm, then we call Y_t a solution of (2).

To ensure the well-posedness of an RDE, proper assumptions are given for V, which will be described by the notation Lip^γ. Throughout this paper, $|\cdot|$ is the Euclidean norm, and we will use C as generic constants, which may be different from line to line.

Definition 2.2. (see e.g., [18]) Let $\gamma > 0$ and $\lfloor\gamma\rfloor$ be the largest integer strictly smaller than γ, i.e., $\gamma - 1 \le \lfloor\gamma\rfloor < \gamma$. We say that $V \in Lip^\gamma$, if V is $\lfloor\gamma\rfloor$-Fréchet differentiable and the kth-derivative of V, D^kV, satisfies that

$$\begin{aligned} &|D^kV(y)| \le C, \ \forall\, k = 0, \cdots, \lfloor\gamma\rfloor, \quad \forall\, y \in \mathbb{R}^m, \\ &|D^{\lfloor\gamma\rfloor}V(y_1) - D^{\lfloor\gamma\rfloor}V(y_2)| \le C|y_1 - y_2|^{\gamma-\lfloor\gamma\rfloor}, \quad \forall\, y_1, y_2 \in \mathbb{R}^m, \end{aligned}$$

for some constant C. The smallest constant C satisfying these two inequalities is denoted by $|V|_{Lip^\gamma}$.

In the sequel, we give the theorem of the well-posedness of an RDE.

Theorem 2.1. *Let $\mathbf{X} \in \mathscr{C}^{p\text{-}var}\big([0,T], G^{[p]}(\mathbb{R}^d)\big)$. If $V = (V_i)_{1\le i\le d}$ is a collection of vector fields in Lip^γ with $\gamma > p$, or a collection of linear vector fields of the form $V_i(Y) = A_iY$ with $A_i \in \mathbb{R}^{m\times m}$, then (2) has a unique solution on $[0,T]$. Moreover, the Jacobian of the flow, $\frac{\partial Y_t}{\partial z}$, exists and satisfies the linear RDE*

$$d\frac{\partial Y_t}{\partial z} = \sum_{i=1}^d DV_i(Y_t)\frac{\partial Y_t}{\partial z}d\mathbf{X}_t^i, \quad \frac{\partial Y_0}{\partial z} = I_m, \tag{3}$$

where $I_m \in \mathbb{R}^{m\times m}$ is an identity matrix.

Proof. For the case of $V \in Lip^\gamma$, the existence and uniqueness of the solution follow from Theorem 10.14 and Theorem 10.26 in [8]. From Theorem 10.26 in [8] and Proposition 1 in [3], we know that $\frac{\partial Y_t}{\partial z}$ exists and satisfies (3) with a bound

$$\sup_{t\in[0,T]} \left|\frac{\partial Y_t}{\partial z}\right| \le |I_m| + C\nu_1\|\mathbf{X}\|_{p\text{-}var;[0,T]} \exp\left\{C\nu_1^p\|\mathbf{X}\|_{p\text{-}var;[0,T]}^p\right\},$$

where $\nu_1 \ge |V|_{Lip^\gamma}$ and C depends only on p.

In the linear case, for any fixed initial value $z = z_0$, the existence and uniqueness of the solution follow from Theorem 10.53 in [8]. It implies that Y will not blow up on $[0, T]$ for some $T > 0$ and

$$\sup_{t\in[0,T]} |Y_t| \le \theta + C(1+\theta)\nu_2 \|\mathbf{X}\|_{p\text{-}var;[0,T]} \exp\left\{c\nu_2^p \|\mathbf{X}\|^p_{p\text{-}var;[0,T]}\right\} =: R(\theta), \tag{4}$$

where $\nu_2 \ge \max\{|A_i| : i = 1, \cdots, d\}$, $\theta \ge |z_0|$ and C depends only on p. This guarantees that we can localize the problem as Theorem 10.21 in [8]. Considering the set $\{\tilde{z} \in \mathbb{R}^m : |\tilde{z}| \le 2\theta\}$, which is a neighborhood of z_0, and replacing V by a compactly supported $\tilde{V} \in Lip^\gamma$ which coincides with V in the ball $B_{2R(\theta)} := \{y : |y| \le 2R(\theta)\}$, we turn the problem into the first case and obtain that $\frac{\partial Y_t}{\partial z}|_{z=z_0}$ exists and satisfies (3). □

3. Symplectic structure of rough Hamiltonian systems

In this section, we derive autonomous Hamiltonian equations with additional nonconservative forces using the modified Hamilton's principle. Under Assumption 3.1, we show that these equations form rough Hamiltonian systems when the nonconservative forces are characterized by rough signals. Moreover, rough Hamiltonian systems possess the symplectic structure similar to the deterministic case.

The classical Hamilton's principle indicates that the motion $Q : [0, T] \to \mathbb{R}^m$ extremizes the action functional $S(Q) := \int_0^T L(Q, \dot{Q})dt$ under variation, where L is the Lagrangian of a deterministic Hamiltonian system. The author in [9] gives the modified Hamilton's principle through the Legendre transform $\mathscr{H}(P, Q) = P\dot{Q} - L(Q, \dot{Q})$. For a Hamiltonian system influenced by additional nonconservative forces, its Hamiltonian energy turns out to be

$$\mathscr{H}_0(P, Q) + \sum_{i=1}^{d} \mathscr{H}_i(P, Q)\dot{\chi}^i$$

instead of $\mathscr{H}(P, Q)$, where $\mathscr{H}_i : \mathbb{R}^{2m} \to \mathbb{R}$ for $i = 0, \cdots, d$. The second term is the total work done by the additional forces and $\dot{\chi}^i = \frac{dX_t^i}{dt}$ represents a formal time derivative of X^i. In this case, the action functional is

$$S(P, Q) = \int_0^T \left[P\dot{Q} - \mathscr{H}_0(P, Q) - \sum_{i=1}^{d} \mathscr{H}_i(P, Q)\dot{\chi}^i\right]dt.$$

Denote perturbation functions to P and Q by δP and δQ respectively with $\delta Q(0) = \delta Q(T) = 0$, then the modified Hamilton's principle reads

$$\begin{aligned}
&\frac{d}{d\epsilon}\bigg|_{\epsilon=0} S(P + \epsilon\delta P, Q + \epsilon\delta Q) \\
&= \int_0^T \left[P\delta\dot{Q} + \dot{Q}\delta P - \frac{\partial\mathscr{H}_0(P, Q)}{\partial P}\delta P - \frac{\partial\mathscr{H}_0(P, Q)}{\partial Q}\delta Q - \sum_{i=1}^{d}\left(\frac{\partial\mathscr{H}_i(P, Q)}{\partial P}\dot{\chi}^i\delta P + \frac{\partial\mathscr{H}_i(P, Q)}{\partial Q}\dot{\chi}^i\delta Q\right)\right]dt \\
&\equiv 0,
\end{aligned}$$

where $\delta\dot{Q}$ is the time derivative of δQ. The chain rule shows that a sufficient condition for the above equation is that P and Q satisfy the following Hamiltonian equations:

$$\begin{aligned}
\dot{P} &= -\frac{\partial\mathscr{H}_0(P, Q)}{\partial Q} - \sum_{i=1}^{d}\frac{\partial\mathscr{H}_i(P, Q)}{\partial Q}\dot{\chi}^i, \\
\dot{Q} &= \frac{\partial\mathscr{H}_0(P, Q)}{\partial P} + \sum_{i=1}^{d}\frac{\partial\mathscr{H}_i(P, Q)}{\partial P}\dot{\chi}^i,
\end{aligned}$$

which can be rewritten as

$$\begin{aligned}
dP &= -\frac{\partial\mathscr{H}_0(P, Q)}{\partial Q}dt - \sum_{i=1}^{d}\frac{\partial\mathscr{H}_i(P, Q)}{\partial Q}dX^i, \quad P_0 = p, \\
dQ &= \frac{\partial\mathscr{H}_0(P, Q)}{\partial P}dt + \sum_{i=1}^{d}\frac{\partial\mathscr{H}_i(P, Q)}{\partial P}dX^i, \quad Q_0 = q.
\end{aligned} \tag{5}$$

Let $P=(P^1,\cdots,P^m)^\top$, $Q=(Q^1,\cdots,Q^m)^\top$, $p=(p^1,\cdots,p^m)^\top$, $q=(q^1,\cdots,q^m)^\top\in\mathbb{R}^m$. In order to get a compact form for (5), we denote $Y=(P^\top,Q^\top)^\top$, $z=(p^\top,q^\top)^\top\in\mathbb{R}^{2m}$, $X_t^0=t$, $X_t=(X_t^0,X_t^1,X_t^2,\cdots,X_t^d)^\top\in\mathbb{R}^{d+1}$ and $V_i=(V_i^1,\cdots,V_i^{2m})=(-\frac{\partial\mathscr{H}_i}{\partial Q^1},\cdots,-\frac{\partial\mathscr{H}_i}{\partial Q^m},\frac{\partial\mathscr{H}_i}{\partial P^1},\cdots,\frac{\partial\mathscr{H}_i}{\partial P^m})^\top$, $i=0,\cdots,d$. Then (5) is equivalent to

$$dY_t=\sum_{i=0}^d V_i(Y_t)dX_t^i=V(Y_t)dX_t,\quad Y_0=z. \tag{6}$$

In the sequel, we consider $X_t^i=X_t^i(\omega)$, $i=1,\cdots,d$, as Gaussian noises under the following assumption.

Assumption 3.1. Let X_t^i, $i=1,\cdots,d$, be independent centered Gaussian processes with continuous sample path on $[0,T]$. There exist some $\rho\in[1,2)$ and $K\in(0,+\infty)$ such that for any $0\le s<t\le T$, the covariance of X satisfies

$$\sup_{(t_k),(t_l)\in\mathscr{D}([s,t])}\left(\sum_{t_k,t_l}\left|\mathbb{E}\left[X_{t_k,t_{k+1}}X_{t_l,t_{l+1}}\right]\right|^\rho\right)^{1/\rho}\le K|t-s|^{1/\rho},$$

where $X_{t_k,t_{k+1}}=X_{t_{k+1}}-X_{t_k}$.

For any $p>2\rho$, by piecewise linear approximations, X can be naturally lifted to a Hölder-type weak geometric p-rough path $\mathbf{X}$ almost surely, which takes values in $G^{[p]}(\mathbb{R}^{d+1})$ and $\pi_1(\mathbf{X}_{s,t})=X_{s,t}$ (see e.g., [8, Theorem 15.33]). Therefore, (5) or (6) can be transformed into an RDE almost surely and interpreted in the framework of rough path theory where the chain rule also holds for rough integrals. In this sense, we call (5) or (6) a rough Hamiltonian system.

The fBm with Hurst parameter $H\in(1/4,1/2]$ satisfies Assumption 3.1 with $\rho=\frac{1}{2H}$. It is widely studied (see e.g., [19]). When X^i, $i=1,\cdots,d$, are standard Brownian motions, i.e., $H=1/2$, Y equals to the solution of a stochastic Hamiltonian system driven by standard Brownian motions in Stratonovich sense almost surely (see e.g., [8, Theorem 17.3]).

As a function of time t and initial value z, Y is a phase flow for almost every ω. In the deterministic case [11] and the stochastic case [20,21], we know that the phase flow preserves the symplectic structure, i.e., the differential 2-form $dP\wedge dQ$ is invariant. Note that the differential here is made with respect to the initial value, which is different from the formal time derivative in (5). The geometric interpretation is that the sum of the oriented areas of two-dimensional surfaces, obtained by projecting the phase flow onto the coordinate planes $(p^1,q^1),\cdots,(p^m,q^m)$, is an integral invariant. The rough Hamiltonian system is a generalization for the deterministic case and the stochastic case, which allows for rougher noises. They share a characteristic property of Hamiltonian systems, which is proved in the next theorem.

Theorem 3.1. *The phase flow of the rough Hamiltonian system* (5) *preserves the symplectic structure:*

$$dP\wedge dQ=dp\wedge dq,\quad a.s.$$

Proof. From Theorem 2.1, we know that P and Q are differentiable with respect to p and q. Denote $P_p^{jk}=\frac{\partial P^j}{\partial p^k}$, $Q_p^{jk}=\frac{\partial Q^j}{\partial p^k}$, $P_q^{jk}=\frac{\partial P^j}{\partial q^k}$, $Q_q^{jk}=\frac{\partial Q^j}{\partial q^k}$. Since $dP^j=\sum_{k=1}^m P_p^{jk}dp^k+\sum_{l=1}^m P_q^{jl}dq^l$ and $dQ^j=\sum_{k=1}^m Q_p^{jk}dp^k+\sum_{l=1}^m Q_q^{jl}dq^l$, we have

$$\begin{aligned}dP\wedge dQ&=\sum_{j=1}^m dP^j\wedge dQ^j\\&=\sum_{j=1}^m\sum_{k=1}^m\sum_{l=1}^m\left(P_p^{jk}Q_q^{jl}-P_q^{jl}Q_p^{jk}\right)dp^k\wedge dq^l\\&\quad+\sum_{j=1}^m\sum_{k=1}^m\sum_{l=1}^{k-1}\left(P_p^{jk}Q_p^{jl}-P_p^{jl}Q_p^{jk}\right)dp^k\wedge dp^l\\&\quad+\sum_{j=1}^m\sum_{k=1}^m\sum_{l=1}^{k-1}\left(P_q^{jk}Q_q^{jl}-P_q^{jl}Q_q^{jk}\right)dq^k\wedge dq^l.\end{aligned}$$

Meanwhile, the time derivatives of P_p^{jk}, Q_p^{jk}, P_q^{jk}, Q_q^{jk} yield

$$dP_p^{jk} = \sum_{i=0}^{d}\sum_{r=1}^{m}\left(-\frac{\partial^2 \mathscr{H}_i}{\partial Q^j \partial P^r}P_p^{rk} - \frac{\partial^2 \mathscr{H}_i}{\partial Q^j \partial Q^r}Q_p^{rk}\right)dX^i, \quad P_p^{jk}(t_0) = \delta_{jk},$$

$$dQ_p^{jk} = \sum_{i=0}^{d}\sum_{r=1}^{m}\left(\frac{\partial^2 \mathscr{H}_i}{\partial P^j \partial P^r}P_p^{rk} + \frac{\partial^2 \mathscr{H}_i}{\partial P^j \partial Q^r}Q_p^{rk}\right)dX^i, \quad Q_p^{jk}(t_0) = 0,$$

$$dP_q^{jk} = \sum_{i=0}^{d}\sum_{r=1}^{m}\left(-\frac{\partial^2 \mathscr{H}_i}{\partial Q^j \partial P^r}P_q^{rk} - \frac{\partial^2 \mathscr{H}_i}{\partial Q^j \partial Q^r}Q_q^{rk}\right)dX^i, \quad P_q^{jk}(t_0) = 0,$$

$$dQ_q^{jk} = \sum_{i=0}^{d}\sum_{r=1}^{m}\left(\frac{\partial^2 \mathscr{H}_i}{\partial P^j \partial P^r}P_q^{rk} + \frac{\partial^2 \mathscr{H}_i}{\partial P^j \partial Q^r}Q_q^{rk}\right)dX^i, \quad Q_q^{jk}(t_0) = \delta_{jk},$$

where all coefficients are calculated at (P, Q). Then one can check

$$d\left(\sum_{j=1}^{m}\left(P_p^{jk}Q_q^{jl} - P_q^{jl}Q_p^{jk}\right)\right) = 0, \quad \forall\, k, l,$$

$$d\left(\sum_{j=1}^{m}\left(P_p^{jk}Q_p^{jl} - P_p^{jl}Q_p^{jk}\right)\right) = 0, \quad \forall\, k \neq l,$$

$$d\left(\sum_{j=1}^{m}\left(P_q^{jk}Q_q^{jl} - P_q^{jl}Q_q^{jk}\right)\right) = 0, \quad \forall\, k \neq l,$$

to get $\sum_{j=1}^{m} dP^j \wedge dQ^j = \sum_{j=1}^{m} dp^j \wedge dq^j$. $\square$

Because of the symplecticity of the phase flow of a rough Hamiltonian system, it is natural to construct numerical methods to inherit this property. However, the simplified step-N Euler schemes, which are similar to the N-level Taylor expansion of the solution, cannot inherit this property in general. Consequently, in the next three sections we will propose and analyze Runge–Kutta methods for rough Hamiltonian systems.

4. Runge–Kutta methods for rough Hamiltonian systems

Given a time step h, we construct s-stage Runge–Kutta methods by

$$\begin{aligned} Y_k^h(\alpha) &= Y_k^h + \sum_{\beta=1}^{s} a_{\alpha\beta} V\big(Y_k^h(\beta)\big) X_{t_k,t_{k+1}}, \\ Y_{k+1}^h &= Y_k^h + \sum_{\alpha=1}^{s} b_\alpha V\big(Y_k^h(\alpha)\big) X_{t_k,t_{k+1}} \end{aligned} \tag{7}$$

with coefficients $a_{\alpha\beta}$, b_α, $\alpha, \beta = 1, \cdots, s$, $t_k = kh$, $k = 0, \cdots, N-1$ and $Y_0^h = z$. Here, we suppose $N = T/h \in \mathbb{N}$ for simplicity. The numerical solution Y_k^h if exists is an approximation for Y_{t_k}.

Theorem 4.1. *The s-stage Runge–Kutta method (7) inherits the symplectic structure of a rough Hamiltonian system, if the coefficients satisfy*

$$a_{\alpha\beta} b_\alpha + a_{\beta\alpha} b_\beta = b_\alpha b_\beta, \ \forall\, \alpha, \beta = 1, \cdots, s.$$

Proof. Denote the jth-component of P_k^h and Q_k^h by $P_k^{h,j}$ and $Q_k^{h,j}$ respectively, $j = 1, \cdots, m$, then from (7) the exterior derivatives of $P_{k+1}^{h,j}$ and $Q_{k+1}^{h,j}$ are

$$dP_{k+1}^{h,j} = dP_k^{h,j} - \sum_{i=0}^{d}\sum_{r=1}^{m}\sum_{\alpha=1}^{s} b_\alpha \left(\frac{\partial^2 \mathscr{H}_i}{\partial Q^j \partial P^r} dP_k^{h,r}(\alpha) + \frac{\partial^2 \mathscr{H}_i}{\partial Q^j \partial Q^r} dQ_k^{h,r}(\alpha)\right) X^i_{t_k,t_{k+1}},$$

$$dQ_{k+1}^{h,j} = dQ_k^{h,j} + \sum_{i=0}^{d}\sum_{r=1}^{m}\sum_{\alpha=1}^{s} b_\alpha \left(\frac{\partial^2 \mathscr{H}_i}{\partial P^j \partial P^r} dP_k^{h,r}(\alpha) + \frac{\partial^2 \mathscr{H}_i}{\partial P^j \partial Q^r} dQ_k^{h,r}(\alpha)\right) X^i_{t_k,t_{k+1}}.$$

The exterior product performed between the above equations yields

$$\begin{aligned} dP_{k+1}^{h,j} \wedge dQ_{k+1}^{h,j} =& dP_k^{h,j} \wedge dQ_k^{h,j} + \sum_{i=0}^{d}\sum_{r=1}^{m}\sum_{\alpha=1}^{s} b_\alpha dP_k^{h,j} \wedge \left(\frac{\partial^2 \mathscr{H}_i}{\partial P^j \partial P^r} dP_k^{h,r}(\alpha) + \frac{\partial^2 \mathscr{H}_i}{\partial P^j \partial Q^r} dQ_k^{h,r}(\alpha)\right) X^i_{t_k,t_{k+1}} \\ &- \sum_{i=0}^{d}\sum_{r=1}^{m}\sum_{\alpha=1}^{s} b_\alpha \left(\frac{\partial^2 \mathscr{H}_i}{\partial Q^j \partial P^r} dP_k^{h,r}(\alpha) + \frac{\partial^2 \mathscr{H}_i}{\partial Q^j \partial Q^r} dQ_k^{h,r}(\alpha)\right) \wedge dQ_k^{h,j} X^i_{t_k,t_{k+1}} \\ &- \left(\sum_{i_1=0}^{d}\sum_{r_1=1}^{m}\sum_{\alpha_1=1}^{s} b_{\alpha_1} \left(\frac{\partial^2 \mathscr{H}_{i_1}}{\partial Q^j \partial P^{r_1}} dP_k^{h,r_1}(\alpha_1) + \frac{\partial^2 \mathscr{H}_{i_1}}{\partial Q^j \partial Q^{r_1}} dQ_k^{h,r_1}(\alpha_1)\right) X^{i_1}_{t_k,t_{k+1}}\right) \\ &\wedge \left(\sum_{i_2=0}^{d}\sum_{r_2=1}^{m}\sum_{\alpha_2=1}^{s} b_{\alpha_2} \left(\frac{\partial^2 \mathscr{H}_{i_2}}{\partial P^j \partial P^{r_2}} dP_k^{h,r_2}(\alpha_2) + \frac{\partial^2 \mathscr{H}_{i_2}}{\partial P^j \partial Q^{r_2}} dQ_k^{h,r_2}(\alpha_2)\right) X^{i_2}_{t_k,t_{k+1}}\right). \end{aligned}$$

Replacing $dP_k^{h,j}$ and $dQ_k^{h,j}$ in the second and third terms by the following expressions from (7), respectively:

$$dP_k^{h,j} = dP_k^{h,j}(\alpha) + \sum_{i=0}^{d}\sum_{r=1}^{m}\sum_{\beta=1}^{s} a_{\alpha\beta} \left(\frac{\partial^2 \mathscr{H}_i}{\partial Q^j \partial P^r} dP_k^{h,r}(\beta) + \frac{\partial^2 \mathscr{H}_i}{\partial Q^j \partial Q^r} dQ_k^{h,r}(\beta)\right) X^i_{t_k,t_{k+1}},$$

$$dQ_k^{h,j} = dQ_k^{h,j}(\alpha) - \sum_{i=0}^{d}\sum_{r=1}^{m}\sum_{\beta=1}^{s} a_{\alpha\beta} \left(\frac{\partial^2 \mathscr{H}_i}{\partial P^j \partial P^r} dP_k^{h,r}(\beta) + \frac{\partial^2 \mathscr{H}_i}{\partial P^j \partial Q^r} dQ_k^{h,r}(\beta)\right) X^i_{t_k,t_{k+1}},$$

we have

$$\begin{aligned} dP_{k+1}^{h,j} \wedge dQ_{k+1}^{h,j} =& dP_k^{h,j} \wedge dQ_k^{h,j} + \sum_{i=0}^{d}\sum_{r=1}^{m}\sum_{\alpha=1}^{s} b_\alpha dP_k^{h,j}(\alpha) \wedge \left(\frac{\partial^2 \mathscr{H}_i}{\partial P^j \partial P^r} dP_k^{h,r}(\alpha) + \frac{\partial^2 \mathscr{H}_i}{\partial P^j \partial Q^r} dQ_k^{h,r}(\alpha)\right) X^i_{t_k,t_{k+1}} \\ &- \sum_{i=0}^{d}\sum_{r=1}^{m}\sum_{\alpha=1}^{s} b_\alpha \left(\frac{\partial^2 \mathscr{H}_i}{\partial Q^j \partial P^r} dP_k^{h,r}(\alpha) + \frac{\partial^2 \mathscr{H}_i}{\partial Q^j \partial Q^r} dQ_k^{h,r}(\alpha)\right) \wedge dQ_k^{h,j}(\alpha) X^i_{t_k,t_{k+1}} \\ &- \sum_{i_1,i_2=0}^{d}\sum_{r_1,r_2=1}^{m}\sum_{\alpha,\beta=1}^{s} (b_\beta a_{\beta\alpha} + b_\alpha a_{\alpha\beta} - b_\alpha b_\beta) \left(\frac{\partial^2 \mathscr{H}_{i_1}}{\partial Q^j \partial P^{r_1}} dP_k^{h,r_1}(\alpha) + \frac{\partial^2 \mathscr{H}_{i_1}}{\partial Q^j \partial Q^{r_1}} dQ_k^{h,r_1}(\alpha)\right) \\ &\wedge \left(\frac{\partial^2 \mathscr{H}_{i_2}}{\partial P^j \partial P^{r_2}} dP_k^{h,r_2}(\beta) + \frac{\partial^2 \mathscr{H}_{i_2}}{\partial P^j \partial Q^{r_2}} dQ_k^{h,r_2}(\beta)\right) X^{i_1}_{t_k,t_{k+1}} X^{i_2}_{t_k,t_{k+1}}. \end{aligned}$$

Summing up the above equation from $j = 1, \cdots, m$, based on the symmetry of the Hessian matrix $\nabla^2 \mathscr{H}_i$, we deduce

$$\begin{aligned} \sum_{j=1}^{m} dP_{k+1}^{h,j} \wedge dQ_{k+1}^{h,j} =& \sum_{j=1}^{m} dP_k^{h,j} \wedge dQ_k^{h,j} \\ &- \sum_{i_1,i_2=0}^{d}\sum_{r_1,r_2,j=1}^{m}\sum_{\alpha,\beta=1}^{s} (b_\beta a_{\beta\alpha} + b_\alpha a_{\alpha\beta} - b_\alpha b_\beta) \left(\frac{\partial^2 \mathscr{H}_{i_1}}{\partial Q^j \partial P^{r_1}} dP_k^{h,r_1}(\alpha) + \frac{\partial^2 \mathscr{H}_{i_1}}{\partial Q^j \partial Q^{r_1}} dQ_k^{h,r_1}(\alpha)\right) \\ &\wedge \left(\frac{\partial^2 \mathscr{H}_{i_2}}{\partial P^j \partial P^{r_2}} dP_k^{h,r_2}(\beta) + \frac{\partial^2 \mathscr{H}_{i_2}}{\partial P^j \partial Q^{r_2}} dQ_k^{h,r_2}(\beta)\right) X^{i_1}_{t_k,t_{k+1}} X^{i_2}_{t_k,t_{k+1}}. \end{aligned}$$

Therefore, we obtain the symplectic condition

$$a_{\alpha\beta} b_\alpha + a_{\beta\alpha} b_\beta = b_\alpha b_\beta, \ \forall\, \alpha, \beta = 1, \cdots, s,$$

to ensure the discrete symplectic structure $dP_{k+1}^h \wedge dQ_{k+1}^h = dP_k^h \wedge dQ_k^h$. $\square$

Symplectic Runge–Kutta methods are implicit in general. As we said before, the solvability of symplectic Runge–Kutta methods should be proven. Based on Brouwer's theorem, we consider the bounded case and the linear case for the vector fields respectively.

Proposition 4.1. *(see also [13, Proposition 3.1]) If $V=(V_i)_{0\le i\le d}$ is a collection of vector fields in Lip^γ for some $\gamma>0$, then for arbitrary time step $h>0$, initial value and coefficients $\{a_{\alpha\beta}, b_\alpha : \alpha,\beta=1,\cdots,s\}$, the s-stage Runge–Kutta method (7) has at least one solution for any path.*

Proof. Fix $h>0$ and $Y_k^h\in\mathbb{R}^{2m}$. Let $Z_1,\cdots,Z_s\in\mathbb{R}^{2m}$ and $Z=(Z_1^\top,\cdots,Z_s^\top)^\top\in\mathbb{R}^{2ms}$. We define a map $\phi:\mathbb{R}^{2ms}\to\mathbb{R}^{2ms}$ with

$$
\begin{aligned}
&\phi(Z)=(\phi(Z)_1^\top,\cdots,\phi(Z)_s^\top)^\top,\\
&\phi(Z)_\alpha=Z_\alpha-Y_k^h-\sum_{\beta=1}^{s}a_{\alpha\beta}V(Z_\beta)X_{t_k,t_{k+1}},\ \alpha=1,\cdots,s.
\end{aligned}
$$

It then suffices to prove that $\phi(Z)=0$ has at least one solution. Let $c=\max\{|a_{\alpha\beta}|:\alpha,\beta=1,\cdots,s\}$, $\nu=|V|_{Lip^\gamma}$ and

$$
R=\sqrt{s}|Y_k^h|+s\sqrt{s}c\nu|X_{t_k,t_{k+1}}|+1,
$$

we have that for any $|Z|=R$,

$$
\begin{aligned}
Z^\top\phi(Z)&=\sum_{\alpha=1}^{s}Z_\alpha^\top\left(Z_\alpha-Y_k^h-\sum_{\beta=1}^{s}a_{\alpha\beta}V(Z_\beta)X_{t_k,t_{k+1}}\right)\\
&\ge|Z|\left(|Z|-\sqrt{s}|Y_k^h|-s\sqrt{s}c\nu|X_{t_k,t_{k+1}}|\right)>0.
\end{aligned}
$$

We aim to show that $\phi(Z)=0$ has a solution in the ball $B_R:=\{Z:|Z|\le R\}$. Assume by contradiction that $\phi(Z)\ne0$ for any $|Z|\le R$. We define a map ψ by $\psi(Z)=-\frac{R\phi(Z)}{|\phi(Z)|}$. Since ψ is continuous and $\psi:B_R\to B_R$, Brouwer's fixed point theorem yields that ψ has at least one fixed point Z^* such that $Z^*=\psi(Z^*)$ and $|Z^*|=R$. This leads to a contradiction since $|Z^*|^2=\psi(Z^*)^\top Z^*=-\frac{R\phi(Z^*)^\top Z^*}{|\phi(Z^*)|}<0$. Therefore, ϕ has at least one solution. □

For the case that vector fields are linear, we need to suppose they are skew symmetric in addition, and obtain the solvability of the midpoint scheme which is the 1-stage symplectic Runge–Kutta method ($a_{11}=\frac{1}{2}, b_1=1$):

$$
\begin{aligned}
Y_{k+1/2}^h&=Y_k^h+\frac{1}{2}V\big(Y_{k+1/2}^h\big)X_{t_k,t_{k+1}},\\
Y_{k+1}^h&=Y_k^h+V\big(Y_{k+1/2}^h\big)X_{t_k,t_{k+1}}.
\end{aligned}
\tag{8}
$$

Proposition 4.2. *Suppose $V=(V_i)_{0\le i\le d}$ is a collection of linear vector fields of the form $V_i(Y)=A_iY$. If A_i, $i=0,\cdots,d$ are all skew symmetric, i.e., $A_i=-A_i^\top$, then for arbitrary time step $h>0$ and initial value, the midpoint scheme is solvable for any path.*

Proof. Fix $h>0$ and $Y_k^h\in\mathbb{R}^{2m}$. Define a map $\phi:\mathbb{R}^{2m}\to\mathbb{R}^{2m}$ with

$$
\phi(Z)=Z-Y_k^h-\frac{1}{2}V(Z)X_{t_k,t_{k+1}}.
$$

Since A_i, $i=0,\cdots,d$, are all skew symmetric, we have $Z^\top A_iZ=0$, which leads to

$$
Z^\top\phi(Z)=|Z|^2-Z^\top Y_k^h-\frac{1}{2}\sum_{i=0}^{d}Z^\top A_iZX_{t_k,t_{k+1}}^i\ge|Z|(|Z|-|Y_k^h|)>0,
$$

for any $|Z|=|Y_k^h|+1=:R$. By similar technique in Proposition 4.1, we have that $\phi(Z)=0$ has at least one solution in the ball $B_R:=\{Z:|Z|\le R\}$. □

Remark 4.1. If a linear rough Hamiltonian system, i.e. (6), can be rewritten equivalently into the form (5) with $V_i(Y)=A_iY$ and $A_i\in\mathbb{R}^{2m\times2m}$ being skew symmetric, then A_i is in the form of $\begin{pmatrix}A_i^1 & -A_i^2\\ A_i^2 & A_i^1\end{pmatrix}$, where $A_i^1=-(A_i^1)^\top\in\mathbb{R}^{m\times m}$ and $A_i^2=(A_i^2)^\top\in\mathbb{R}^{m\times m}$.

5. Convergence analysis

In Definition 2.1, the solution of an RDE depends on the information about iterated integrals of the noise, which is very difficult to simulate. The Runge–Kutta methods are implementable because they only make use of the increments of the noise and omit the dissymmetric parts of iterated integrals. For instance, Lévy's area is omitted in the standard Brownian motion setting. To analyze the error caused by this, we introduce piecewise linear approximations, which is a special case of Wong–Zakai approximations in [7]. We define $x_t^h := (x_t^{h,0}, \cdots, x_t^{h,d})$ with

$$x_t^{h,i} := X_{t_k}^i + \frac{t-t_k}{h} X_{t_k,t_{k+1}}^i, \ t \in (t_k, t_{k+1}], \ \forall\, k = 1, \cdots, N, \ i = 0, \cdots, d, \tag{9}$$

and consider the following differential equation in Riemann–Stieltjes integral sense:

$$dy_t^h = V(y_t^h) dx_t^h, \quad y_0^h = z. \tag{10}$$

Theorem 5.1. *Suppose X and x^h are as in Assumption 3.1 and (9) respectively. If for some $\gamma > 2\rho$, $V = (V_i)_{0 \le i \le d}$ is a collection of vector fields in Lip^γ or a collection of linear vector fields of the form $V_i(Y) = A_i Y$, then both (6) and (10) have unique solutions Y and y^h almost surely.*

Moreover, let $\theta \ge |z|$, $\nu \ge \max\{|V|_{Lip^\gamma}, |A_i|, i = 0, \cdots, d\}$, then for any $0 \le \eta < \min\{\frac{1}{\rho} - \frac{1}{2}, \frac{1}{2\rho} - \frac{1}{\gamma}\}$, there exists a finite random variable $C(\omega)$ and a null set M such that

$$\sup_{t \in [0,T]} |y_t^h(\omega) - Y_t(\omega)| \le C(\omega) h^\eta, \quad \forall\, \omega \in \Omega \setminus M$$

holds for any $h > 0$. The finite random variable $C(\omega)$ depends on ρ, η, γ, ν, θ, K and T.

Proof. For the first case of V, see Theorem 6 and Corollary 8 in [7]. The finite random variable $C(\omega)$ depends on ρ, η, γ, ν, K and T. For the linear case, we know if $2\rho < p < \gamma$, then for any $\omega \in \Omega \setminus M$, $\|S_{[p]}(x^h)(\omega)\|_{p\text{-}var;[0,T]}$ has an upper bound for all $h > 0$ (see [8, Theorem 15.33]). According to the estimate in (4), there exists some constant $R > 0$ depending on ω, ρ, γ, ν, θ and T, such that it is a bound for Y_t and y_t^h, for all $h > 0$ on $[0, T]$. Using again localization, we replace V by a compactly supported $\tilde{V} \in Lip^\gamma$ which coincides with V in the ball $B_R := \{y : |y| \le R\}$ and then the result can be generalized from the first case to this linear case. $\square$

We are in position to analyze the convergence rate. If $V \in Lip^\gamma$, we analyze the midpoint scheme for $\rho \in [1, 3/2)$ in Theorem 5.2 and two s-stage symplectic Runge–Kutta methods ($s > 1$) for $\rho \in [1, 2)$ in Theorem 5.3. If V is linear, we focus on the convergence rate of the midpoint scheme for $\rho \in [1, 3/2)$ in Theorem 5.4, using the uniform boundedness of numerical solutions in Proposition 5.1.

Theorem 5.2. *Suppose $\rho \in [1, 3/2)$ and Y_k^h is the numerical solution of the midpoint scheme (8). If $|V|_{Lip^\gamma} \le \nu < \infty$ with $\gamma > 2\rho$, then for any $0 < \eta < \min\{\frac{3}{2\rho} - 1, \frac{1}{2\rho} - \frac{1}{\gamma}\}$, there exists a finite random variable $C(\omega)$ and a null set M such that*

$$\max_{k=1,\cdots,N} |Y_k^h(\omega) - Y_{t_k}(\omega)| \le C(\omega) h^\eta, \quad \forall\, \omega \in \Omega \setminus M$$

holds for any $0 < h \le 1$. The finite random variable $C(\omega)$ depends on ρ, η, γ, ν, K and T.

Proof. We divide the error into two parts

$$\max_{k=1,\cdots,N} |Y_k^h - Y_{t_k}| \le \max_{k=1,\cdots,N} |Y_k^h - y_{t_k}^h| + \sup_{t \in [0,T]} |y_t^h - Y_t|. \tag{11}$$

The second part has been estimated in Theorem 5.1, then it suffices to estimate the first part.

Step 1. Local error. According to the definition of x_t^h, for any $k = 1, \cdots, N$, $n \ge 1$, $i_1, \cdots, i_n \in \{0, \cdots, d\}$, we have

$$\begin{aligned}\int\limits_{t_{k-1} \le u_1 < \cdots < u_n \le t_k} dx_{u_1}^{h,i_1} \cdots dx_{u_n}^{h,i_n} &= \int\limits_{t_{k-1} \le u_1 < \cdots < u_n \le t_k} \frac{X_{t_{k-1},t_k}^{i_1}}{h} du_1 \cdots \frac{X_{t_{k-1},t_k}^{i_n}}{h} du_n \\ &= \frac{X_{t_{k-1},t_k}^{i_1} \cdots X_{t_{k-1},t_k}^{i_n}}{n!},\end{aligned}$$

which implies $\int_{t_{k-1} \le u_1 < \cdots < u_n \le t_k} dx_{u_1}^h \otimes \cdots \otimes dx_{u_n}^h = X_{t_{k-1},t_k}^{\otimes n}/n!$. Taylor expansion yields

$$
\begin{aligned}
y_{t_1}^h &= y_0^h + \int\limits_{0\le u_1\le t_1} V(y_{u_1}^h)dx_{u_1}^h \\
&= y_0^h + V(y_0^h)X_{0,t_1} + \int\limits_{0\le u_2<u_1\le t_1} DV(y_{u_2}^h)V(y_{u_2}^h)dx_{u_2}^h\otimes dx_{u_1}^h \\
&= y_0^h + V(y_0^h)X_{0,t_1} + DV(y_0^h)V(y_0^h)X_{0,t_1}^{\otimes 2}/2 + R_1
\end{aligned}
$$

with

$$R_1 = \int\limits_{0\le u_3<u_2<u_1\le t_1} \Big(DV(y_{u_3}^h)^2V(y_{u_3}^h) + D^2V(y_{u_3}^h)V(y_{u_3}^h)^2\Big)dx_{u_3}^h\otimes dx_{u_2}^h\otimes dx_{u_1}^h.$$

According to Taylor expansion and $Y_{1/2}^h = (Y_0^h + Y_1^h)/2$ implied by (8), we have that

$$
\begin{aligned}
V\big(Y_{1/2}^h\big) =& V\big(Y_0^h\big) + DV(Y_0^h)(Y_{1/2}^h - Y_0^h) + D^2V(\tilde{Y}_0^h)(Y_{1/2}^h - Y_0^h)^{\otimes 2}/2 \\
=& V\big(Y_0^h\big) + DV(Y_0^h)(Y_1^h - Y_0^h)/2 + D^2V(\tilde{Y}_0^h)(Y_1^h - Y_0^h)^{\otimes 2}/8.
\end{aligned}
$$

Here $\tilde{Y}_0^h$ is determined by Y_0^h and $Y_{1/2}^h$. Then,

$$
\begin{aligned}
Y_1^h &= Y_0^h + V\big(Y_{1/2}^h\big)X_{0,t_1} \\
&= Y_0^h + \Big[V\big(Y_0^h\big) + DV(Y_0^h)(Y_1^h - Y_0^h)/2 + D^2V(\tilde{Y}_0^h)(Y_1^h - Y_0^h)^{\otimes 2}/8\Big]X_{0,t_1}.
\end{aligned}
$$

Replacing $DV(Y_0^h)(Y_1^h - Y_0^h)$ in the above equation by

$$DV(Y_0^h)\Big[V\big(Y_0^h\big) + DV(Y_0^h)(Y_1^h - Y_0^h)/2 + D^2V(\tilde{Y}_0^h)(Y_1^h - Y_0^h)^{\otimes 2}/8\Big]X_{0,t_1},$$

we get

$$Y_1^h = Y_0^h + V\big(Y_0^h\big)X_{0,t_1} + DV(Y_0^h)V(Y_0^h)X_{0,t_1}^{\otimes 2}/2 + R_2$$

with

$$R_2 = DV(Y_0^h)^2(Y_1^h - Y_0^h)X_{0,t_1}^{\otimes 2}/4 + DV(Y_0^h)D^2V(\tilde{Y}_0^h)(Y_1^h - Y_0^h)^{\otimes 2}X_{0,t_1}^{\otimes 2}/16 + D^2V(\tilde{Y}_0^h)(Y_1^h - Y_0^h)^{\otimes 2}X_{0,t_1}/8.$$

Noticing $|Y_1^h - Y_0^h| \le |V\big(Y_{1/2}^h\big)\|X_{0,t_1}|$ and $\gamma > 2$, we obtain that the difference between Y_1^h and $y_{t_1}^h$ is

$$|Y_1^h - y_{t_1}^h| \le |R_2 - R_1| \le \max\{|V|_{Lip^\gamma}^4, 1\}(|X_{0,t_1}|^3 + |X_{0,t_1}|^4).$$

For any η satisfying that $\max\{0, \frac{3}{\gamma} - 1\} < \eta < \frac{3}{2\rho} - 1$, there exists a $2\rho < p < \gamma$ such that $\eta = \frac{3}{p} - 1$. Since for any $\omega \in \Omega\setminus M$, $|X_{0,t_1}(\omega)| = |\pi_1(\mathbf{X}_{0,t_1}(\omega))| \le \|\mathbf{X}(\omega)\|_{1/p\text{-Höl};[0,T]}h^{1/p}$, the local error is

$$|Y_1^h(\omega) - y_{t_1}^h(\omega)| \le C(\omega)h^{3/p}, \quad \forall\, h \le 1.$$

Step 2. Global error. We use the notation $\pi(t_0, y_0, x^h)_t$, $t \ge t_0$, to represent the solution of (10) with the driven noise x^h and the initial value y_0 at time t_0. Similar to Theorem 10.30 in [8],

$$
\begin{aligned}
|Y_k^h(\omega) - y_{t_k}^h(\omega)| &\le \sum_{s=1}^k |\pi(t_s, Y_s^h, x^h)_{t_k} - \pi(t_{s-1}, Y_{s-1}^h, x^h)_{t_k}| \\
&= \sum_{s=1}^k |\pi(t_s, Y_s^h, x^h)_{t_k} - \pi(t_s, \pi(t_{s-1}, Y_{s-1}^h, x^h)_{t_s}, x^h)_{t_k}|.
\end{aligned}
$$

Since the solution is continuously dependent on the initial value, which is given by the estimate in [8, Theorem 10.26], we have for any $1 \le s \le k \le N$,

$$|\pi(t_s, Y_s^h, x^h)_{t_k} - \pi(t_s, \pi(t_{s-1}, Y_{s-1}^h, x^h)_{t_s}, x^h)_{t_k}| \le C\exp\{C\nu^p\|S_{[p]}(x^h(\omega))\|_{p\text{-}var;[t_s,t_k]}^p\}|Y_s^h - \pi(t_{s-1}, Y_{s-1}^h, x^h)_{t_s}|,$$

where $C = C(\gamma, p)$. Because for any $\omega \in \Omega \setminus M$, $\|S_{[p]}(x^h)(\omega)\|^p_{p\text{-}var;[0,T]}$ has an upper bound for $h > 0$ (see [8, Theorem 15.33]), the estimate for the local error yields

$$|Y^h_k(\omega) - y^h_{t_k}(\omega)| \le C(\omega)h^{3/p-1}.$$

Therefore, for any $\omega \in \Omega \setminus M$, $0 < \eta < \frac{3}{2\rho} - 1$,

$$\max_{k=1,\cdots,N} |Y^h_k(\omega) - y^h_{t_k}(\omega)| \le C(\omega)h^{\eta}. \tag{12}$$

Combining Theorem 5.1 and the fact that $\frac{3}{2\rho} - 1 \le \frac{1}{\rho} - \frac{1}{2}$ for $\rho \in [1, 3/2)$, we complete the proof. □

When the components of the noise are independent standard Brownian motions, i.e., $\rho = 1$, the convergence order is (almost) consistent with the mean square convergence rate of SDEs in Stratonovich sense with multiplicative noise. Moreover, note that the estimate in Theorem 5.2 is valid for the case when $\rho \in [1, 3/2)$, which is similar to the simplified step-2 Euler scheme in [7]. To construct a numerical scheme for $\rho \in [1, 2)$ and fill the gap between the convergence rates of two parts in (11), in the following theorem we use higher stage symplectic Runge–Kutta methods with local order $\tau \ge 4$ when applied to classical ordinary differential equations.

Theorem 5.3. *Suppose $\rho \in [1, 2)$ and Y^h_k represents the solution of one of the two s-stage Runge–Kutta methods with coefficients expressed in the Butcher tableaus below.*

Method I ($s = 2$):

$$\begin{array}{c|cc} (3-\sqrt{3})/6 & 1/4 & (3-2\sqrt{3})/12 \\ (3+\sqrt{3})/6 & (3+2\sqrt{3})/12 & 1/4 \\ \hline & 1/2 & 1/2 \end{array}.$$

Method II ($s = 3$):

$$\begin{array}{c|ccc} a/2 & a/2 & & \\ 3a/2 & a & a/2 & \\ 1/2+a & a & a & 1/2-a \\ \hline & a & a & 1-2a \end{array},$$

where $a = 1.351207$ is the real root of $6x^3 - 12x^2 + 6x - 1 = 0$.

If $|V|_{Lip^{\gamma}} \le \nu < \infty$ with $\gamma > \max\{2\rho, 3\}$, then for any $0 < \eta < \min\{\frac{1}{\rho} - \frac{1}{2}, \frac{1}{2\rho} - \frac{1}{\gamma}\}$, there exists a finite random variable $C(\omega)$ and a null set M such that

$$\max_{k=1,\cdots,N} |Y^h_k(\omega) - Y_{t_k}(\omega)| \le C(\omega)h^{\eta}, \quad \forall\, \omega \in \Omega \setminus M.$$

In addition, both of the two methods inherit the symplectic structure of a rough Hamiltonian system.

Proof. We adopt a similar strategy as in Theorem 5.2.

Since $V \in Lip^{\gamma}$ with $\gamma > 3$, a Taylor expansion of the solution $y^h_{t_1}$ in (10) leads to

$$y^h_{t_1} = y^h_0 + V(y^h_0)X_{0,t_1} + DV(y^h_0)V(y^h_0)X^{\otimes 2}_{0,t_1}/2 + DV(y^h_0)^2V(y^h_0)X^{\otimes 3}_{0,t_1}/6 + D^2V(y^h_0)V(y^h_0)^2X^{\otimes 3}_{0,t_1}/6 + R_1 \tag{13}$$

with $|R_1| \le \max\{|V|^6_{Lip^{\gamma}}, 1\}|X_{0,t_1}|^4/24$.

For scheme (7), a Taylor expansion shows for any $\alpha = 1, \cdots, s$,

$$\begin{aligned} V\big(Y^h_0(\alpha)\big) =& V\big(Y^h_0\big) + DV(Y^h_0)(Y^h_0(\alpha) - Y^h_0) + D^2V(Y^h_0)(Y^h_0(\alpha) - Y^h_0)^{\otimes 2}/2 \\ &+ D^3V(\tilde{Y}^h_0(\alpha))(Y^h_0(\alpha) - Y^h_0)^{\otimes 3}/6, \end{aligned}$$

where $\tilde{Y}^h_0(\alpha)$ is determined by $Y^h_0(\alpha)$ and Y^h_0. Plugging it into $Y^h_1 = Y^h_0 + \sum_{\alpha=1}^s b_\alpha V\big(Y^h_0(\alpha)\big)X_{0,t_1}$, we get

$$\begin{aligned} Y^h_1 =& Y^h_0 + \sum_{\alpha=1}^s b_\alpha V\big(Y^h_0\big)X_{0,t_1} + \sum_{\alpha=1}^s b_\alpha DV(Y^h_0)(Y^h_0(\alpha) - Y^h_0)X_{0,t_1} \\ &+ \sum_{\alpha=1}^s b_\alpha D^2V(Y^h_0)(Y^h_0(\alpha) - Y^h_0)^{\otimes 2}X_{0,t_1}/2 + \sum_{\alpha=1}^s b_\alpha D^3V(\tilde{Y}^h_0(\alpha))(Y^h_0(\alpha) - Y^h_0)^{\otimes 3}X_{0,t_1}/6. \end{aligned}$$

Denoting $c_\alpha = \sum_{\beta=1}^{s} a_{\alpha\beta}$, we have from Taylor expansion that

$$Y_0^h(\alpha) - Y_0^h = \sum_{\beta=1}^{s} a_{\alpha\beta} V\big(Y_0^h(\beta)\big) X_{0,t_1} = c_\alpha V\big(Y_0^h\big) X_{0,t_1} + \sum_{\beta=1}^{s} a_{\alpha\beta} DV(Y_0^h) c_\beta V\big(Y_0^h\big) X_{0,t_1}^{\otimes 2} + R_2$$

with $|R_2| \le C(a_{\alpha\beta}, |V|_{Lip^\gamma})|X_{0,t_1}|^3$. Then

$$\begin{aligned} Y_1^h = & Y_0^h + \sum_{\alpha=1}^{s} b_\alpha V\big(Y_0^h\big) X_{0,t_1} + \sum_{\alpha=1}^{s} b_\alpha DV(Y_0^h)\Big[c_\alpha V\big(Y_0^h\big) X_{0,t_1} + \sum_{\beta=1}^{s} a_{\alpha\beta} DV(Y_0^h) c_\beta V\big(Y_0^h\big) X_{0,t_1}^{\otimes 2}\Big] X_{0,t_1} \\ & + \sum_{\alpha=1}^{s} b_\alpha D^2 V(Y_0^h)\Big[c_\alpha^2 V(y_0^h)^2 X_{0,t_1}^{\otimes 2}\Big] X_{0,t_1}/2 + R_4, \end{aligned}$$

with $|R_4| \le C(a_{\alpha\beta}, |V|_{Lip^\gamma})|X_{0,t_1}|^4$. Comparing it with (13), we get the order conditions

$$\sum_{\alpha=1}^{s} b_\alpha = 1, \quad \sum_{\alpha=1}^{s} b_\alpha c_\alpha = \frac{1}{2}, \quad \sum_{\alpha=1}^{s}\sum_{\beta=1}^{s} b_\alpha a_{\alpha\beta} c_\beta = \frac{1}{6}, \quad \sum_{\alpha=1}^{s} b_\alpha c_\alpha^2 = \frac{1}{3},$$

for

$$|Y_1^h - y_{t_1}^h| \le C(a_{\alpha\beta}, |V|_{Lip^\gamma})|X_{0,t_1}|^4,$$

which are satisfied by Method I and Method II (see also [11, Section II.1.1]). For any η satisfying that $\max\{0, \frac{4}{\gamma} - 1\} < \eta < \frac{2}{\rho} - 1$, there exists some $p \in (2\rho, \gamma)$ such that $\eta = \frac{4}{p} - 1$. Since for any $\omega \in \Omega \setminus M$, it holds that $|X_{0,t_1}(\omega)| = |\pi_1(\mathbf{X}_{0,t_1}(\omega))| \le \|\mathbf{X}(\omega)\|_{1/p\text{-Höl};[0,T]} h^{1/p}$, the local error shows

$$|Y_1^h(\omega) - y_{t_1}^h(\omega)| \le C(\omega) h^{4/p}, \quad \forall\, h \le 1.$$

A similar proof of Theorem 5.2 yields the estimate for the global error that for any $\rho \in [1, 2)$, $0 < \eta < \min\{\frac{1}{\rho} - \frac{1}{2}, \frac{1}{2\rho} - \frac{1}{\gamma}\}$ and $\omega \in \Omega \setminus M$,

$$\max_{k=1,\cdots,N} |Y_k^h(\omega) - Y_{t_k}(\omega)| \le C(\omega) h^\eta.$$

Moreover, it can be checked that the symplectic condition in Theorem 4.1 is also satisfied by Method I and Method II. $\square$

Remark 5.1. This convergence rate equals to that of the simplified step-3 Euler scheme in [7] for γ large enough. Indeed, using Runge–Kutta methods with a higher stage will not improve the convergence rate since the error of piecewise linear approximations persists.

The next proposition is essential to get the convergence rate of the midpoint scheme in the linear case.

Proposition 5.1. *If V in (6) is a collection of skew symmetric linear vector fields of the form $V_i(Y) = A_i Y$, then numerical solutions of the midpoint scheme (8) are uniformly bounded. More precisely, $|Y_k^h| = |z|$, $k = 1, \cdots, N$.*

Proof. For any $k = 0, \cdots, N-1$, since $Y_{k+1}^h = Y_k^h + \sum_{i=0}^{d} A_i\big(Y_{k+1/2}^h\big) X_{t_k,t_{k+1}}^i$, we have

$$\begin{aligned} |Y_{k+1}^h|^2 = (Y_{k+1}^h)^\top Y_{k+1}^h = & |Y_k^h|^2 + \sum_{i=0}^{d} (Y_{k+1/2}^h)^\top A_i^\top Y_k^h X_{t_k,t_{k+1}}^i + \sum_{i=0}^{d} (Y_k^h)^\top A_i Y_{k+1/2}^h X_{t_k,t_{k+1}}^i \\ & + \sum_{i_1=0}^{d}\sum_{i_2=0}^{d} (Y_{k+1/2}^h)^\top A_{i_1}^\top A_{i_2} Y_{k+1/2}^h X_{t_k,t_{k+1}}^{i_1} X_{t_k,t_{k+1}}^{i_2}. \end{aligned}$$

Substituting $Y_k^h = Y_{k+1/2}^h - \frac{1}{2}\sum_{i=0}^{d} A_i Y_{k+1/2}^h X_{t_k,t_{k+1}}^i$ into the second and third term, we get

$$|Y_{k+1}^h|^2 = |Y_k^h|^2 + 2\sum_{i=0}^{d} (Y_{k+1/2}^h)^\top A_i Y_{k+1/2}^h X_{t_k,t_{k+1}}^i.$$

Therefore, $|Y_{k+1}^h|^2 = |Y_k^h|^2$ results from $(Y_{k+1/2}^h)^\top A_i Y_{k+1/2}^h = 0$, since A_i is skew symmetric. $\square$

Theorem 5.4. *Suppose $\rho \in [1, 3/2)$. Let V be as in Proposition 5.1 and Y_k^h represents the numerical solution of the midpoint scheme. If $\theta \geq |z|$, $\nu \geq \max\{|A_i| : i = 0, \cdots, d\}$, then for any $\gamma > 2\rho$, $0 < \eta < \min\{\frac{3}{2\rho} - 1, \frac{1}{2\rho} - \frac{1}{\gamma}\}$, there exists a finite random variable $C(\omega)$ and a null set M such that*

$$\max_{k=1,\cdots,N} |Y_k^h(\omega) - Y_{t_k}(\omega)| \leq C(\omega)h^{\eta}, \quad \forall\, \omega \in \Omega \setminus M$$

holds for any $0 < h \leq 1$. The finite random variable $C(\omega)$ depends on $\rho, \eta, \gamma, \nu, \theta, K$ and T.

Proof. Proposition 5.1 implies that there exists some constant $R > 0$ depending on ω, ρ, γ, ν, θ and T, such that $|\pi(t_{k-1}, Y_{k-1}^h, x^h)_t| \leq R$ for all h, k, t. Then the localization technique is applicable. Using the same approach as in Theorem 5.1 and Theorem 5.2, we complete the proof. □

6. Numerical experiments

In this section, we illustrate our theoretical results by two examples. In the first example, we verify the theoretical convergence rate given in Theorem 5.3. In the second one, we consider a linear model and compare the performance of the midpoint scheme, the simplified step-2 Euler scheme and the simplified step-3 Euler scheme.

6.1. Example 1

We consider a rough Hamiltonian system in $\mathbb{R}^2$ with

$$\mathscr{H}_0(P, Q) = \sin(P)\cos(Q), \ \mathscr{H}_1(P, Q) = \cos(P), \ \mathscr{H}_2(P, Q) = \sin(Q).$$

More precisely, the corresponding Hamiltonian equations are

$$\begin{aligned} dP &= \sin(P)\sin(Q)dt - \cos(Q)dX^2, \quad P_0 = p, \\ dQ &= \cos(P)\cos(Q)dt - \sin(P)dX^1, \quad Q_0 = q, \end{aligned}$$

where X^1 and X^2 are independent fBms with Hurst parameter $H \in (1/4, 1/2]$.

Pathwise convergence rate. Since the vector fields are bounded with bounded derivatives, the theoretical convergence rates of both Method I and Method II in Theorem 5.3 are $(0.3 - \varepsilon)$, $(0.2 - \varepsilon)$, $(0.1 - \varepsilon)$ for arbitrary small ε, when the Hurst parameter H are $0.4, 0.35, 0.3$ respectively. Fig. 1 shows the maximum error in the discretization points of Method I, $\max_{k=1,\cdots,N} |Y_k^h(\omega) - Y_{t_k}(\omega)|$, which we call the pathwise maximum error for short. Pictures in the same row are from three sample paths with the same H. In addition, the fact that the error becomes larger, as H decreases, implies the influence of the roughness of the noise.

6.2. Example 2

For the linear case, we consider the following model

$$\begin{aligned} dP &= -Q\,dt - \sigma \sum_{i=1}^{3} Q\,dX^i, \quad P_0 = p, \\ dQ &= P dt + \sigma \sum_{i=1}^{3} P dX^i, \quad Q_0 = q, \end{aligned} \tag{14}$$

where X^1, X^2 and X^3 are independent fBms with Hurst parameter $H = 0.4$.

We have the expression of the exact solution

$$\begin{aligned} P &= p\cos(t + \sigma \sum_{i=1}^{3} X_t^i) - q\sin(t + \sigma \sum_{i=1}^{3} X_t^i), \\ Q &= q\cos(t + \sigma \sum_{i=1}^{3} X_t^i) + p\sin(t + \sigma \sum_{i=1}^{3} X_t^i). \end{aligned}$$

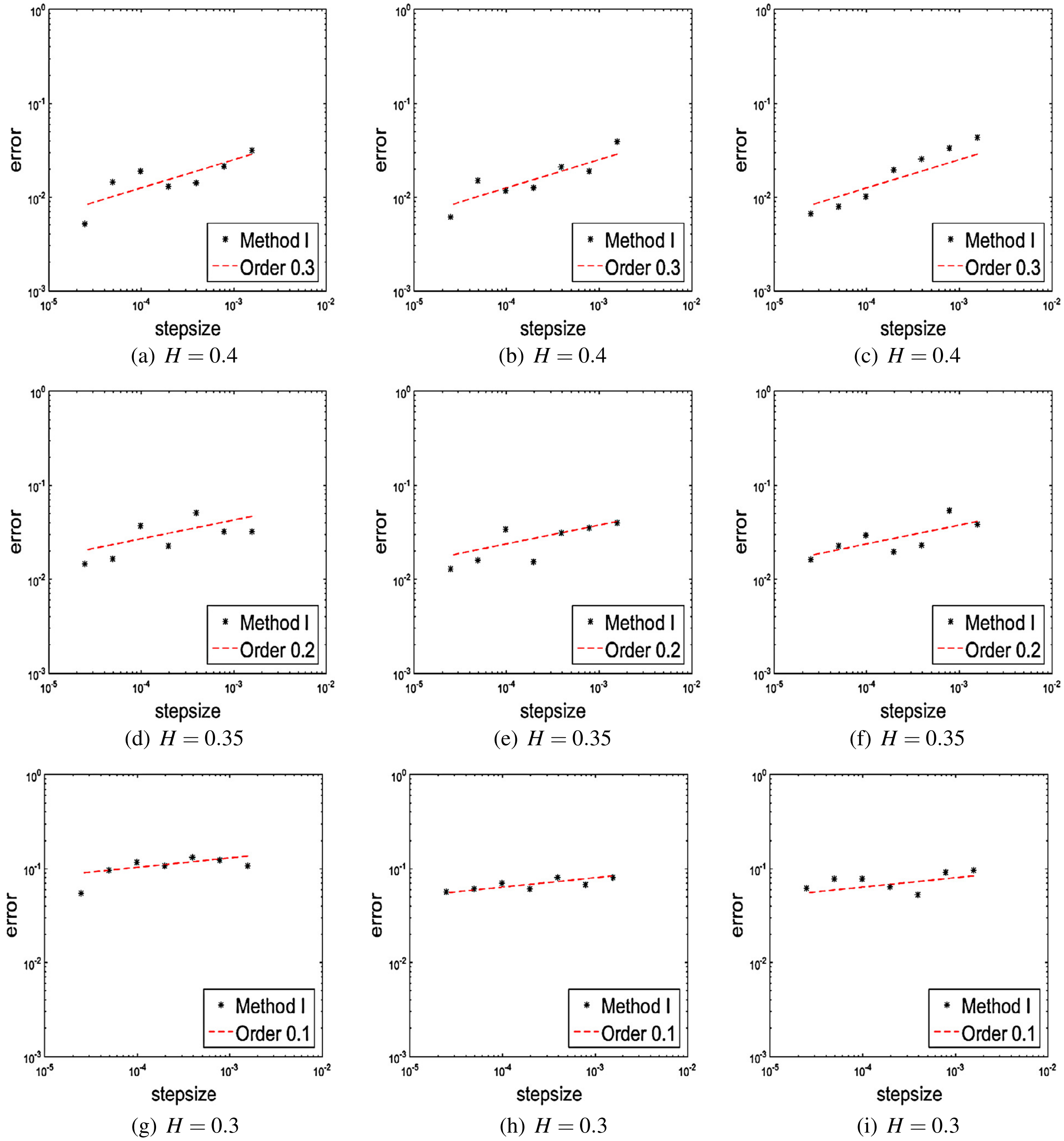


Fig. 1. Pathwise maximum error vs. step size for three sample paths with $p = 1$, $q = 2$ and $T = 0.1$.

Denoting $Y = (P, Q)^\top \in \mathbb{R}^2$ and $X^0 := t$, we have that (14) is equivalent to $dY = \sum_{i=0}^{3} A_i Y dX^i$, where $A_0 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ and $A_i = \begin{pmatrix} 0 & -\sigma \\ \sigma & 0 \end{pmatrix}$, $i = 1, 2, 3$. Since A_i are all skew symmetric, the midpoint scheme is solvable. We compare it with the simplified step-2 and step-3 Euler schemes in the following experiments.

Evolution of domains. Fig. 2 shows the evolution of domains in the phase plane for one sample path. The initial domain is a square with four corners at $(1, 1)$, $(2, 1)$, $(2, 2)$ and $(1, 2)$. Images at $t = 0.4, 1.6, 8$, are presented under the exact mapping and the three numerical methods. One can observe that the exact mapping is area preserving, which is equivalent to that it preserves the symplectic structure of the system. This property is inherited by the midpoint scheme, since the areas of

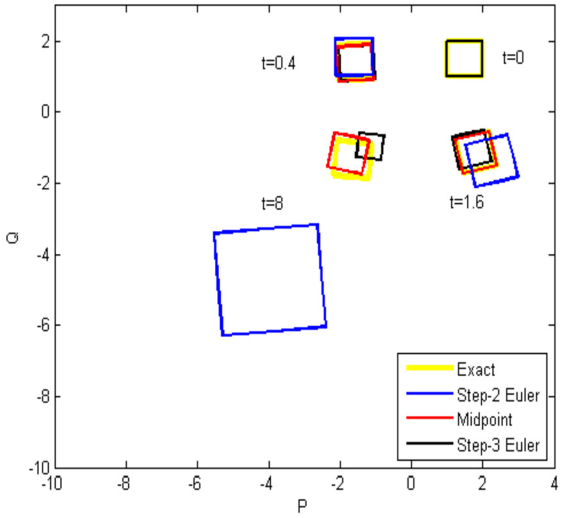
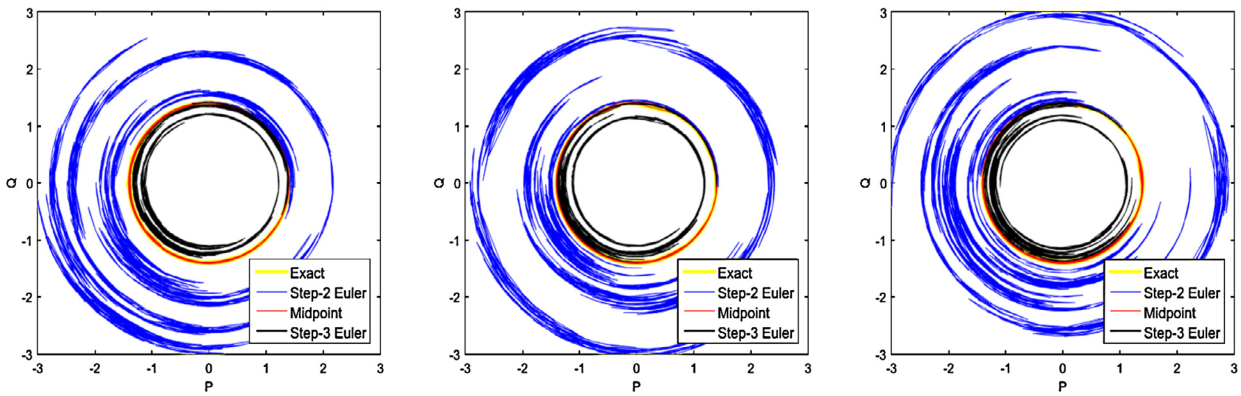


Fig. 2. Evolution of domains in the phase plane for one sample path with $\sigma = 1.5$, $T = 10$ and $h = 0.002$. (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)



(a) (b) (c)

Fig. 3. Phase trajectories for three sample paths with $p = 1$, $q = 1$, $\sigma = 1$, $T = 10$ and $h = 0.002$.

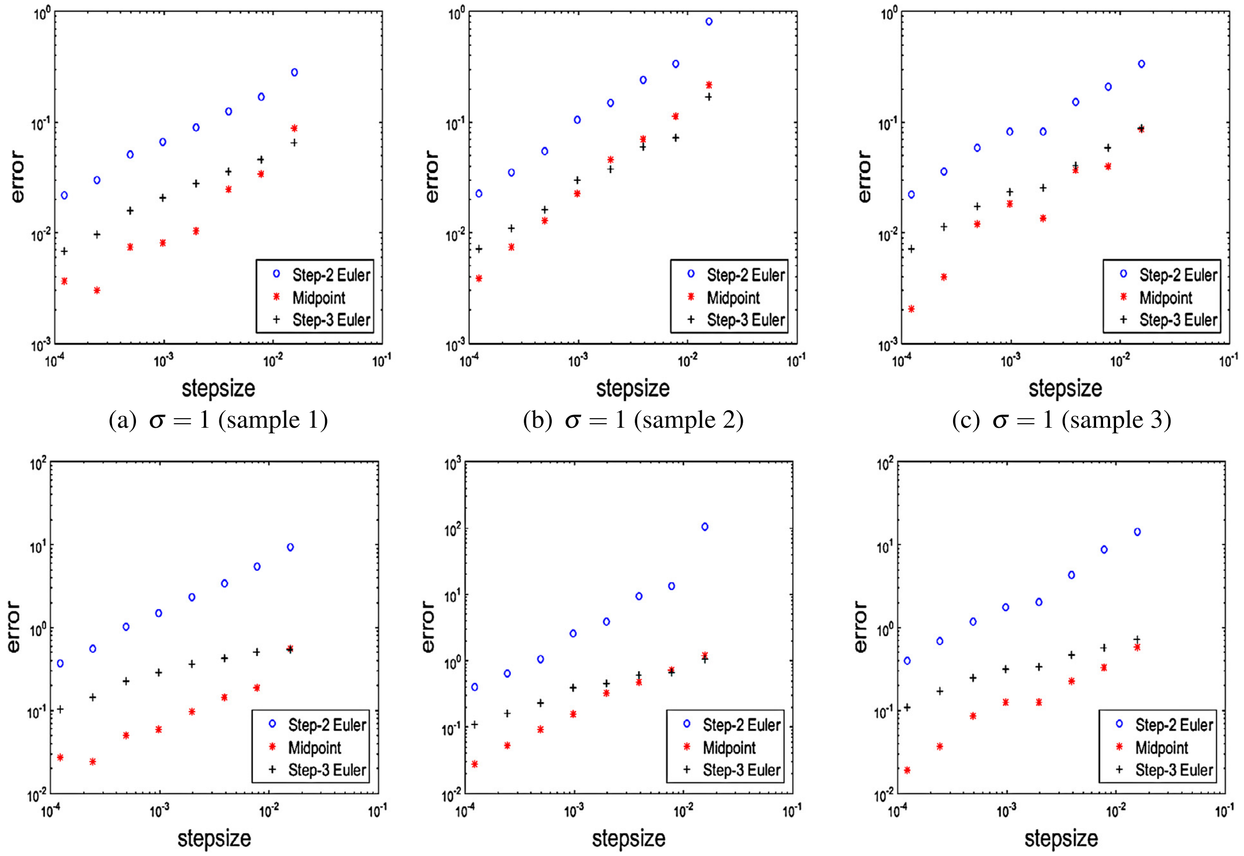
red squares are the same. However, the areas of images under the simplified step-2 and step-3 Euler schemes increase and decrease respectively, which proves that they are not symplectic methods.

Conservation of quadratic invariant. For the initial value $(1, 1)$, the exact solution is on the circle with center at the origin and radius $r = \sqrt{2}$, which implies another invariant of this system (see also [2, Example 1]). We present the phase trajectories for three sample paths in Fig. 3. The midpoint scheme preserves this property such that its solutions are always on that circle, which is proved in Proposition 5.1. The phase trajectories of the simplified step-2 Euler scheme deviate from the exact one a lot and those of the simplified step-3 Euler scheme shrink gradually. Fig. 2 and Fig. 3 verify that the midpoint scheme, as a symplectic Runge–Kutta method, preserves the invariants of the system (14).

Stability for the size of the noise. We investigate the influence of the size σ of the noise in Fig. 4. The pictures in the same column are from the same sample path. As σ becomes larger, the pathwise maximum error of the midpoint scheme is smaller in general. This implies the stability of an implicit method.

Acknowledgements

The authors gratefully thank the anonymous referees for valuable comments and suggestions in improving this paper. This work is supported by National Natural Science Foundation of China (No. 91530118, No. 91130003, No. 11021101, No. 91630312 and No. 11290142).



(a) $\sigma = 1$ (sample 1) (b) $\sigma = 1$ (sample 2) (c) $\sigma = 1$ (sample 3)

(d) $\sigma = 2$ (sample 1) (e) $\sigma = 2$ (sample 2) (f) $\sigma = 2$ (sample 3)

Fig. 4. Pathwise maximum error vs. step size for three sample paths with $p = 1$, $q = 1$ and $T = 1$.

References

[1] S. Anmarkrud, A. Kværnø, Order conditions for stochastic Runge–Kutta methods preserving quadratic invariants of Stratonovich SDEs, J. Comput. Appl. Math. 316 (2017) 40–46.
[2] C. Bayer, P. Friz, S. Riedel, J. Schoenmakers, From rough path estimates to multilevel Monte Carlo, SIAM J. Numer. Anal. 54 (3) (2016) 1449–1483.
[3] T. Cass, P. Friz, Densities for rough differential equations under Hörmander's condition, Ann. Math. (2) 171 (3) (2010) 2115–2141.
[4] A. Davie, Differential equations driven by rough paths: an approach via discrete approximation, Appl. Math. Res. Express 40 (2) (2007) abm009.
[5] A. Deya, A. Neuenkirch, S. Tindel, A Milstein-type scheme without Lévy area terms for SDEs driven by fractional Brownian motion, Ann. Inst. Henri Poincaré Probab. Stat. 48 (2) (2012) 518–550.
[6] P. Friz, M. Hairer, A Course on Rough Paths, Universitext, Springer, Cham, 2014, with an introduction to regularity structures.
[7] P. Friz, S. Riedel, Convergence rates for the full Gaussian rough paths, Ann. Inst. Henri Poincaré Probab. Stat. 50 (1) (2014) 154–194.
[8] P. Friz, N. Victoir, Multidimensional Stochastic Processes as Rough Paths. Theory and Applications, Cambridge Studies in Advanced Mathematics, vol. 120, Cambridge University Press, Cambridge, 2010.
[9] H. Goldstein, Classical Mechanics, second edition, Addison-Wesley Series in Physics, Addison-Wesley Publishing Co., Reading, Mass, 1980.
[10] M. Gubinelli, Controlling rough paths, J. Funct. Anal. 216 (1) (2004) 86–140.
[11] E. Hairer, C. Lubich, G. Wanner, Structure-preserving algorithms for ordinary differential equations, in: Geometric Numerical Integration, second edition, in: Springer Series in Computational Mathematics, vol. 31, Springer-Verlag, Berlin, 2006.
[12] M. Hairer, N. Pillai, Regularity of laws and ergodicity of hypoelliptic SDEs driven by rough paths, Ann. Probab. 41 (4) (2013) 2544–2598.
[13] J. Hong, C. Huang, X. Wang, Strong convergence rate for two classes of implementable methods for SDEs driven by fractional Brownian motions, arXiv:1711.02907.
[14] Y. Hu, Y. Liu, D. Nualart, Crank–Nicolson scheme for stochastic differential equations driven by fractional Brownian motions, arXiv:1709.01614.
[15] Y. Hu, Y. Liu, D. Nualart, Rate of convergence and asymptotic error distribution of Euler approximation schemes for fractional diffusions, Ann. Appl. Probab. 26 (2) (2016) 1147–1207.
[16] Y. Hu, Y. Liu, D. Nualart, Taylor schemes for rough differential equations and fractional diffusions, Discrete Contin. Dyn. Syst., Ser. B 21 (9) (2016) 3115–3162.
[17] Y. Liu, S. Tindel, First-order Euler scheme for SDEs driven by fractional Brownian motions: the rough case, arXiv:1703.03625.
[18] T. Lyons, Differential equations driven by rough signals, Rev. Mat. Iberoam. 14 (2) (1998) 215–310.
[19] B. Mandelbrot, J. Van Ness, Fractional Brownian motions, fractional noises and applications, SIAM Rev. 10 (1968) 422–437.
[20] G. Milstein, Yu. Repin, M. Tretyakov, Numerical methods for stochastic systems preserving symplectic structure, SIAM J. Numer. Anal. 40 (4) (2002) 1583–1604.

[21] G. Milstein, Yu. Repin, M. Tretyakov, Symplectic integration of Hamiltonian systems with additive noise, SIAM J. Numer. Anal. 39 (6) (2002) 2066–2088.
[22] Tony Shardlow, Phillip Taylor, On the pathwise approximation of stochastic differential equations, BIT 56 (3) (2016) 1101–1129.
[23] P. Wang, J. Hong, D. Xu, Construction of symplectic Runge–Kutta methods for stochastic Hamiltonian systems, Commun. Comput. Phys. 21 (1) (2017) 237–270.