Variational integrators for stochastic dissipative Hamiltonian systems

Michael Kraus*[1,2] and Tomasz M. Tyranowski†[1]

[1]Max-Planck-Institut für Plasmaphysik
Boltzmannstraße 2, 85748 Garching, Germany
[2]Technische Universität München, Zentrum Mathematik
Boltzmannstraße 3, 85748 Garching, Germany

Abstract

Variational integrators are derived for structure-preserving simulation of stochastic forced Hamiltonian systems. The derivation is based on a stochastic discrete Hamiltonian which approximates a type-II stochastic generating function for the stochastic flow of the Hamiltonian system. The generating function is obtained by introducing an appropriate stochastic action functional and considering a stochastic generalization of the deterministic Lagrange-d'Alembert principle. Our approach presents a general methodology to derive new structure-preserving numerical schemes. The resulting integrators satisfy a discrete version of the stochastic Lagrange-d'Alembert principle, and in the presence of symmetries, they also satisfy a discrete counterpart of stochastic forced Noether's theorem. Furthermore, mean-square and weak Lagrange-d'Alembert Runge-Kutta methods are proposed and tested numerically to demonstrate their superior long-time numerical stability and energy behavior compared to non-geometric methods. The Vlasov Fokker-Planck equation is considered as one of the numerical test cases, and a new geometric approach to collisional kinetic plasmas is presented.

1 Introduction

Stochastic differential equations (SDEs) play an important role in modeling dynamical systems subject to internal or external random fluctuations. Standard references include [9], [38], [44], [49], [65], [75]. Within this class of problems, we are interested in stochastic forced Hamiltonian systems, which take the form

$$\begin{aligned} d_t q &= \frac{\partial H}{\partial p} dt + \sum_{i=1}^{m} \frac{\partial h_i}{\partial p} \circ dW^i(t), \\ d_t p &= \left[-\frac{\partial H}{\partial q} + F(q,p) \right] dt + \sum_{i=1}^{m} \left[-\frac{\partial h_i}{\partial q} + f_i(q,p) \right] \circ dW^i(t), \end{aligned} \tag{1.1}$$

*michael.kraus@ipp.mpg.de
†tomasz.tyranowski@ipp.mpg.de

where $H = H(q,p)$ and $h_i = h_i(q,p)$ for $i = 1,\dots,m$ are the Hamiltonian functions, $F = F(q,p)$ and $f_i = f_i(q,p)$ are the forcing terms, $W(t) = (W^1(t),\dots,W^m(t))$ is the standard m-dimensional Wiener process, and $\circ$ denotes Stratonovich integration. We use d_t to denote the stochastic differential of stochastic processes (other than the Wiener process $W(t)$) to avoid confusion with the exterior derivative d of differential forms. The system (1.1) can be formally regarded as a classical forced Hamiltonian system with the randomized Hamiltonian given by $\widehat{H}(q,p,t) = H(q,p) + \sum_{i=1}^m h_i(q,p)\circ\dot{W}^i(t)$, and the randomized forcing given by $\widehat{F}(q,p,t) = F(q,p) + \sum_{i=1}^m f_i(q,p)\circ\dot{W}^i(t)$, where $H(q,p)$ and $F(q,p)$ are the deterministic Hamiltonian and forcing, respectively, and $h_i(q,p)$, $f_i(q,p)$ represent the intensity of the noise. Equation (1.1) is a generalization of stochastic Hamiltonian systems considered in [11], [35], [52], and [69]. Such systems can be used to model, e.g., mechanical systems with uncertainty, or error, assumed to arise from random forcing, limited precision of experimental measurements, or unresolved physical processes on which the Hamiltonian of the deterministic system might otherwise depend. Applications arise in many models in physics, chemistry, and biology. Particular examples include molecular dynamics (see, e.g., [10], [39], [50], [84]), dissipative particle dynamics (see, e.g., [77]), investigations of the dispersion of passive tracers in turbulent flows (see, e.g., [82], [89]), energy localization in thermal equilibrium (see, e.g., [76]), lattice dynamics in strongly anharmonic crystals (see, e.g., [27]), description of noise induced transport in stochastic ratchets (see, e.g., [51]), and collisional kinetic plasmas ([43], [85]).

As occurs for other SDEs, most Hamiltonian SDEs cannot be solved analytically and one must resort to numerical simulations to obtain approximate solutions. In principle, general purpose stochastic numerical schemes for SDEs can be applied to stochastic Hamiltonian systems. However, as for their deterministic counterparts, stochastic Hamiltonian systems possess several important geometric features: in the case of systems without forcing, their phase space flows (almost surely) preserve the *symplectic* structure ([11], [68], [69]); when the forcing terms are present, then the solutions also satisfy the stochastic *Lagrange-d'Alembert principle*, as will be shown in Section 2, and in some special cases the phase space flow may be *conformally symplectic* (see [12], [36], [70]). When simulating these systems numerically, it is therefore advisable that the numerical scheme also preserves such geometric features. Geometric integration of deterministic Hamiltonian systems has been thoroughly studied (see [28], [64], [81] and the references therein) and symplectic integrators have been shown to demonstrate superior performance in long-time simulations of Hamiltonian systems without forcing, compared to non-symplectic methods; so it is natural to pursue a similar approach for stochastic Hamiltonian systems. This is a relatively recent pursuit. Stochastic symplectic integrators are discussed in [4], [6], [7], [8], [19], [22], [25], [37], [56], [57], [68], [69], [71], [87], [94], [95], [97], [98].

Long-time accuracy and near preservation of the Hamiltonian by symplectic integrators applied to deterministic Hamiltonian systems have been rigorously studied using the so-called backward error analysis (see, e.g., [28] and the references therein). To the best of our knowledge, such general rigorous results have not yet been proved for stochastic Hamiltonian systems, but backward error analysis for SDEs is currently an active area of research. Modified SDEs associated with some particular numerical schemes are considered in [1], [24], [83], [96], and [99]. Backward error analysis for the Langevin equation with additive noise is studied for several integrators in [2], [45], and [46]. Recently, backward error analysis for a weak symplectic scheme applied to a stochastic Hamiltonian system has been presented in [5]. The numerical evidence and partial theoretical results to date are promising and suggest that stochastic geometric integrators indeed possess the property of very accurately capturing the evolution of the Hamiltonian H over long time intervals.

An important class of geometric integrators are *variational integrators*. This type of numerical

schemes is based on discrete variational principles and provides a natural framework for the discretization of Lagrangian systems, including forced, dissipative, or constrained ones. These methods have the advantage that they are symplectic when applied to systems without forcing, and in the presence of a symmetry, they satisfy a discrete version of Noether's theorem. For an overview of variational integration for deterministic systems see [60]; see also [31], [40], [42], [53], [54], [72], [73], [80], [91], [93]. Variational integrators were introduced in the context of finite-dimensional mechanical systems, but were later generalized to Lagrangian field theories (see [59]) and applied in many computations, for example in elasticity, electrodynamics, or fluid dynamics; see [55], [74], [86], [90].

Stochastic variational integrators were first introduced in [14] and further studied in [13]. However, those integrators were restricted to the special case when the Hamiltonian functions $h_i = h_i(q)$ were independent of p, and only low-order Runge-Kutta types of discretization were considered. Stochastic discrete Hamiltonian variational integrators applicable to a general class of Hamiltonian systems were proposed in [35] by generalizing the variational principle for deterministic systems introduced in [54] and applying a Galerkin type of discretization; see also [33]. In the present work we extend the ideas put forth in [35] to forced systems of the form (1.1) and propose the corresponding Lagrange-d'Alembert variational integrators.

When the forcing terms in Eq. (1.1) are linear functions of the momentum variable p, then the stochastic flow of the system is conformally symplectic (see [70] and Section 2.4). Stochastic conformally symplectic integrators for such systems were proposed in [12], [15], and [36]. Quasi-symplectic integrators were introduced in [70] and further studied in [66]. These ideas are very interesting, but at present seem to be limited only to systems that exhibit a very special form, that is, systems with separable Hamiltonians, linear forcing terms, and additive noise. The stochastic Lagrange-d'Alembert variational integrators introduced in Section 3 are applicable to the general class of systems of the form (1.1) and preserve their underlying variational structure.

Main content The main content of the remainder of this paper is, as follows.

In Section 2 we introduce a stochastic Lagrange-d'Alembert principle and a stochastic generating function suitable for considering stochastic forced Hamiltonian systems, and we discuss their properties.

In Section 3 we present a general framework for constructing stochastic Lagrange-d'Alembert variational integrators, prove the discrete stochastic Lagrange-d'Alembert principle, propose mean-square and weak stochastic Lagrange-d'Alembert partitioned Runge-Kutta methods, and present several particularly interesting examples of low-stage schemes. We also discuss connections with the idea of quasi-symplectic integrators.

In Section 4 we present the results of our numerical tests, which verify the excellent long-time performance of our integrators compared to some popular non-geometric methods. In particular, as one of the test cases we consider the Vlasov Fokker-Planck equation, which is used as a model for collisional kinetic plasmas.

Section 5 contains the summary of our work.

2 Lagrange-d'Alembert principle for stochastic forced Hamiltonian systems

The stochastic variational integrators proposed in [14] and [13] were formulated for dynamical systems which are described by a Lagrangian and which are subject to noise whose magnitude depends only on the position q. Therefore, these integrators can be extended to (1.1) only if the Hamiltonian functions $h_i = h_i(q)$ are independent of p and the Hamiltonian H is non-degenerate (i.e., the associated Legendre transform is invertible). However, in the case of general $h_i = h_i(q,p)$ the paths $q(t)$ of the system become almost surely nowhere differentiable, which poses a difficulty in interpreting the meaning of the corresponding Lagrangian. To avoid these kind of issues, in [35] an action functional based on a phase space Lagrangian was introduced, and variational integrators for unforced Hamiltonian systems were constructed. In the present work we extend the approach taken in [35] to include forced Hamiltonian systems. To begin, in the next section, we will introduce an appropriate stochastic action functional and show that it can be used to define a type-II generating function for the stochastic flow of the system (1.1).

2.1 Stochastic Lagrange-d'Alembert principle

Let the Hamiltonian functions $H : T^*Q \longrightarrow \mathbb{R}$ and $h_i : T^*Q \longrightarrow \mathbb{R}$ for $i = 1,\dots,m$ be defined on the cotangent bundle T^*Q of the configuration manifold Q, and let (q,p) denote the canonical coordinates on T^*Q. The Hamiltonian forces $F : T^*Q \longrightarrow T^*Q$ and $f_i : T^*Q \longrightarrow T^*Q$ for $i = 1,\dots,m$ are fiber-preserving mappings with the coordinate representations $F(q,p) = (q, F(q,p))$ and $f_i(q,p) = (q, f_i(q,p))$, respectively, where by a slight abuse of notation we use the same symbol to denote the force and its local representation. For simplicity, in this work we assume that the configuration manifold has a vector space structure, $Q \cong \mathbb{R}^N$, so that $T^*Q = Q \times Q^* \cong \mathbb{R}^N \times \mathbb{R}^N$ and $TQ = Q \times Q \cong \mathbb{R}^N \times \mathbb{R}^N$. In this case, the natural pairing between one-forms and vectors can be identified with the scalar product on $\mathbb{R}^N$, that is, $\langle (q,p), (q,\dot{q}) \rangle = p \cdot \dot{q}$, where $(q,\dot{q})$ denotes the coordinates on TQ. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be the probability space with the filtration $\{\mathcal{F}_t\}_{t\geq 0}$, and let $W(t) = (W^1(t),\dots,W^m(t))$ denote a standard m-dimensional Wiener process on that probability space (such that $W(t)$ is $\mathcal{F}_t$-measurable). We will assume that the Hamiltonian functions and the forcing terms are sufficiently smooth and satisfy all the necessary conditions for the existence and uniqueness of solutions to (1.1), and their extendability to a given time interval $[t_a, t_b]$ with $t_b > t_a \geq 0$. One possible set of such assumptions can be formulated by considering the Itô form of (1.1),

$$d_t z = A(z)dt + B(z)dW(t), \tag{2.1}$$

with $z = (q,p)$ and

$$A(z) = \begin{pmatrix} \frac{\partial H}{\partial p} + \frac{1}{2}\sum_{i=1}^m \left[\frac{\partial^2 h_i}{\partial p \partial q}\frac{\partial h_i}{\partial p} + \frac{\partial^2 h_i}{\partial p^2}\left(f_i - \frac{\partial h_i}{\partial q}\right)\right] \\ -\frac{\partial H}{\partial q} + F + \frac{1}{2}\sum_{i=1}^m \left[\left(\frac{\partial^2 h_i}{\partial q \partial p} - \frac{\partial f_i}{\partial p}\right)\left(\frac{\partial h_i}{\partial p} - f_i\right) - \left(\frac{\partial^2 h_i}{\partial q^2} - \frac{\partial f_i}{\partial q}\right)\frac{\partial h_i}{\partial p}\right] \end{pmatrix}, \qquad B(z) = \begin{pmatrix} \left(\frac{\partial h}{\partial p}\right)^T \\ -\left(\frac{\partial h}{\partial q}\right)^T + f \end{pmatrix}, \tag{2.2}$$

where $\partial^2 h_i/\partial q^2$, $\partial^2 h_i/\partial p^2$, and $\partial^2 h_i/\partial q \partial p$ denote the Hessian matrices of h_i, whereas $\partial h/\partial q$, $\partial h/\partial p$, $\partial f_i/\partial q$, and $\partial f_i/\partial p$ denote the Jacobian matrices of $h = (h_1,\dots,h_m)$ and f_i, respectively, and the

$n \times m$ forcing matrix f is defined as $f = (f_1, \ldots, f_m)$. For simplicity and clarity of the exposition, throughout this paper we assume that (see [9], [38], [44], [49])

(H1) H and h_i for $i = 1, \ldots, m$ are C^2 functions of their arguments

(H2) F and f_i for $i = 1, \ldots, m$ are C^1 functions of their arguments

(H3) A and B are globally Lipschitz

These assumptions are sufficient for our purposes, but could be relaxed if necessary. Define the space

$$C([t_a, t_b]) = \big\{(q,p) : \Omega \times [t_a, t_b] \longrightarrow T^*Q \,\big|\, q,\, p \text{ are almost surely continuous } \mathcal{F}_t\text{-adapted semimartingales}\big\}. \tag{2.3}$$

Since we assume $T^*Q \cong \mathbb{R}^N \times \mathbb{R}^N$, the space $C([t_a, t_b])$ is a vector space (see [75]). Therefore, we can identify the tangent space $TC([t_a, t_b]) \cong C([t_a, t_b]) \times C([t_a, t_b])$. We can now define the following stochastic action functional, $\mathcal{B} : \Omega \times C([t_a, t_b]) \longrightarrow \mathbb{R}$,

$$\mathcal{B}\big[q(\cdot), p(\cdot)\big] = p(t_b)q(t_b) - \int_{t_a}^{t_b} \Big[p \circ d_t q - H\big(q(t), p(t)\big)\, dt - \sum_{i=1}^{m} h_i\big(q(t), p(t)\big) \circ dW^i(t)\Big], \tag{2.4}$$

where $\circ$ denotes Stratonovich integration, and we have omitted writing the elementary events $\omega \in \Omega$ as arguments of functions, following the standard convention in stochastic analysis. For a given curve $\big(q(t), p(t)\big)$ in T^*Q and its arbitrary variation $\big(\delta q(t), \delta p(t)\big)$, we define the corresponding variation of the action functional as

$$\delta\mathcal{B}\big[q(\cdot), p(\cdot)\big] \equiv \frac{d}{d\epsilon}\bigg|_{\epsilon=0} \mathcal{B}\big[q(\cdot) + \epsilon\delta q(\cdot), p(\cdot) + \epsilon\delta p(\cdot)\big]. \tag{2.5}$$

Theorem 2.1 (**Stochastic Lagrange-d'Alembert Principle in Phase Space**). *Suppose that $H(q,p)$, $F(q,p)$, and $h_i(q,p)$, $f_i(q,p)$ for $i = 1, \ldots, m$ satisfy conditions (H1)-(H3). If the curve $\big(q(t), p(t)\big)$ in T^*Q satisfies the stochastic forced Hamiltonian system* (1.1) *for $t \in [t_a, t_b]$, where $t_b \geq t_a > 0$, then it also satisfies the integral equation*

$$\delta\mathcal{B}\big[q(\cdot), p(\cdot)\big] - \int_{t_a}^{t_b} F\big(q(t), p(t)\big) \cdot \delta q(t)\, dt - \sum_{i=1}^{m} \int_{t_a}^{t_b} f_i\big(q(t), p(t)\big) \cdot \delta q(t) \circ dW^i(t) = 0\,, \tag{2.6}$$

almost surely for all variations $\big(\delta q(\cdot), \delta p(\cdot)\big) \in C([t_a, t_b])$ such that almost surely $\delta q(t_a) = 0$ and $\delta p(t_b) = 0$.

Proof. Let the curve $\big(q(t), p(t)\big)$ in T^*Q satisfy (1.1) for $t \in [t_a, t_b]$. It then follows that the stochastic processes $q(t)$ and $p(t)$ are almost surely continuous, $\mathcal{F}_t$-adapted semimartingales, that is, $\big(q(\cdot), p(\cdot)\big) \in C([t_a, t_b])$ (see [9], [75]). We calculate the variation (2.5) as

$$\delta\mathcal{B}\big[q(\cdot),p(\cdot)\big] = p(t_b)\delta q(t_b) - \int_{t_a}^{t_b} p(t)\circ d_t\delta q(t) - \int_{t_a}^{t_b}\delta p(t)\circ d_t q(t)$$
$$+\int_{t_a}^{t_b}\left[\frac{\partial H}{\partial q}\big(q(t),p(t)\big)\,\delta q(t) + \frac{\partial H}{\partial p}\big(q(t),p(t)\big)\,\delta p(t)\right]dt$$
$$+\sum_{i=1}^{m}\int_{t_a}^{t_b}\left[\frac{\partial h_i}{\partial q}\big(q(t),p(t)\big)\,\delta q(t) + \frac{\partial h_i}{\partial p}\big(q(t),p(t)\big)\,\delta p(t)\right]\circ dW^i(t), \tag{2.7}$$

where we have used the end point condition, $\delta p(t_b) = 0$. Since the Hamiltonians are C^2 and the processes $q(t)$, $p(t)$ are almost surely continuous, in the last two lines we have used a dominated convergence argument to interchange differentiation with respect to ϵ and integration with respect to t and $W(t)$. Upon applying the integration by parts formula for semimartingales (see [75]), we find

$$\int_{t_a}^{t_b} p(t)\circ d_t\delta q(t) = p(t_b)\delta q(t_b) - p(t_a)\delta q(t_a) - \int_{t_a}^{t_b}\delta q(t)\circ d_t p(t). \tag{2.8}$$

Substituting and rearranging terms produces,

$$\delta\mathcal{B}\big[q(\cdot),p(\cdot)\big] = \int_{t_a}^{t_b}\delta q(t)\left[\circ d_t p(t) + \frac{\partial H}{\partial q}\big(q(t),p(t)\big)\,dt + \sum_{i=1}^{m}\frac{\partial h_i}{\partial q}\big(q(t),p(t)\big)\circ dW^i(t)\right]$$
$$-\int_{t_a}^{t_b}\delta p(t)\left[\circ d_t q(t) - \frac{\partial H}{\partial p}\big(q(t),p(t)\big)\,dt - \sum_{i=1}^{m}\frac{\partial h_i}{\partial p}\big(q(t),p(t)\big)\circ dW^i(t)\right], \tag{2.9}$$

where we have used $\delta q(t_a) = 0$. Therefore, we have

$$\delta\mathcal{B}\big[q(\cdot),p(\cdot)\big] - \int_{t_a}^{t_b} F\big(q(t),p(t)\big)\cdot\delta q(t)\,dt - \sum_{i=1}^{m}\int_{t_a}^{t_b} f_i\big(q(t),p(t)\big)\cdot\delta q(t)\circ dW^i(t)$$
$$= \underbrace{\int_{t_a}^{t_b}\delta q(t)\left[\circ d_t p(t) + \left(\frac{\partial H}{\partial q}\big(q(t),p(t)\big) - F\big(q(t),p(t)\big)\right)dt + \sum_{i=1}^{m}\left(\frac{\partial h_i}{\partial q}\big(q(t),p(t)\big) - f_i\big(q(t),p(t)\big)\right)\circ dW^i(t)\right]}_{A}$$
$$\underbrace{-\int_{t_a}^{t_b}\delta p(t)\left[\circ d_t q(t) - \frac{\partial H}{\partial p}\big(q(t),p(t)\big)\,dt - \sum_{i=1}^{m}\frac{\partial h_i}{\partial p}\big(q(t),p(t)\big)\circ dW^i(t)\right]}_{B}. \tag{2.10}$$

Since $\big(q(t),p(t)\big)$ satisfy (1.1), then by definition we have that almost surely for all $t\in[t_a,t_b]$,

$$q(t) = q(t_a) + \underbrace{\int_{t_a}^{t}\frac{\partial H}{\partial p}(q(s),p(s))\,ds}_{M_0(t)} + \sum_{i=1}^{m}\underbrace{\int_{t_a}^{t}\frac{\partial h_i}{\partial p}(q(s),p(s))\circ dW^i(s)}_{M_i(t)}, \tag{2.11}$$

that is, $q(t)$ can be represented as the sum of the semi-martingales $M_i(t)$ for $i = 0,\dots,m$, where the sample paths of the process $M_0(t)$ are almost surely continuously differentiable. Let us calculate

$$\int_{t_a}^{t_b} \delta p(t) \circ d_t q(t) = \int_{t_a}^{t_b} \delta p(t) \circ d_t\Big(q(t_a) + M_0(t) + \sum_{i=1}^{m} M_i(t)\Big)$$
$$= \int_{t_a}^{t_b} \delta p(t) \circ d_t M_0(t) + \sum_{i=1}^{m} \int_{t_a}^{t_b} \delta p(t) \circ d_t M_i(t)$$
$$= \int_{t_a}^{t_b} \delta p(t) \frac{\partial H}{\partial p}(q(t), p(t))\, dt + \sum_{i=1}^{m} \int_{t_a}^{t_b} \delta p(t) \frac{\partial h_i}{\partial p}(q(t), p(t)) \circ dW^i(t), \quad (2.12)$$

where in the last equality we have used the standard property of the Riemann-Stieltjes integral for the first term, as $M_0(t)$ is almost surely differentiable, and the associativity property of the Stratonovich integral for the second term (see [75], [38]). Substituting (2.12) in the term B of (2.10), we show that $B = 0$. By a similar argument we also prove that $A = 0$. Therefore, the left-hand side of (2.10) is equal to zero, almost surely.

□

Remark: It is natural to expect that the converse theorem, that is, if $\big(q(\cdot), p(\cdot)\big)$ satisfy the integral principle (2.6), then the curve $\big(q(t), p(t)\big)$ is a solution to (1.1), should also hold, although a larger class of variations $(\delta q, \delta p)$ may be necessary. Variants of such a theorem for systems without forcing have been proved in Lázaro-Camí & Ortega [52] and Bou-Rabee & Owhadi [14]. We leave this as an open question. Here, we will use the action functional (2.4) and the Lagrange-d'Alembert principle (2.6) to construct numerical schemes, and we will directly verify that these numerical schemes converge to solutions of (1.1).

2.2 Stochastic type-II generating function and forcing

When the functions $H(q,p)$, $F(q,p)$, $h_i(q,p)$, and $f_i(q,p)$ satisfy standard measurability and regularity conditions (e.g., (H1)-(H3)), then the system (1.1) possesses a pathwise unique stochastic flow $F_{t,t_0} : \Omega \times T^*Q \longrightarrow T^*Q$. It can be proved that for fixed t, t_0 this flow is mean-square differentiable with respect to the q, p arguments, and is also almost surely a diffeomorphism (see [9], [38], [44], [49]). We will show below that the action functional (2.4) can be used to construct a type II generating function for F_{t,t_0}. Let $(\bar{q}(t), \bar{p}(t))$ be a particular solution of (1.1) on $[t_a, t_b]$. Suppose that for almost all $\omega \in \Omega$ there is an open neighborhood $\mathcal{U}(\omega) \subset Q$ of $\bar{q}(\omega, t_a)$, an open neighborhood $\mathcal{V}(\omega) \subset Q^*$ of $\bar{p}(\omega, t_b)$, and an open neighborhood $\mathcal{W}(\omega) \subset T^*Q$ of the curve $(\bar{q}(\omega,t), \bar{p}(\omega,t))$ such that for all $q_a \in \mathcal{U}(\omega)$ and $p_b \in \mathcal{V}(\omega)$ there exists a pathwise unique solution $(\bar{q}(\omega,t;q_a,p_b), \bar{p}(\omega,t;q_a,p_b))$ of (1.1) which satisfies $\bar{q}(\omega,t_a;q_a,p_b) = q_a$, $\bar{p}(\omega,t_b;q_a,p_b) = p_b$, and $(\bar{q}(\omega,t;q_a,p_b), \bar{p}(\omega,t;q_a,p_b)) \in \mathcal{W}(\omega)$ for $t_a \le t \le t_b$. (As in the deterministic case, for t_b sufficiently close to t_a one can argue that such neighborhoods exist; see [58].) Define the function $S : \mathcal{Y} \longrightarrow \mathbb{R}$ as

$$S(q_a, p_b) = \mathcal{B}\big[\bar{q}(\cdot; q_a, p_b), \bar{p}(\cdot; q_a, p_b)\big], \quad (2.13)$$

where the domain $\mathcal{Y} \subset \Omega \times Q \times Q^*$ is given by $\mathcal{Y} = \bigcup_{\omega \in \Omega} \{\omega\} \times \mathcal{U}(\omega) \times \mathcal{V}(\omega)$. Define further the two functions $F^{\pm} : \mathcal{Y} \longrightarrow \mathbb{R}^N$ as

$$F^{-}(q_a,p_b)=\int_{t_a}^{t_b}\left(\frac{\partial\bar{q}(t;q_a,p_b)}{\partial q_a}\right)^T\Big[F\big(\bar{q}(t;q_a,p_b),\bar{p}(t;q_a,p_b)\big)\,dt+\sum_{i=1}^{m}f_i\big(\bar{q}(t;q_a,p_b),\bar{p}(t;q_a,p_b)\big)\circ dW^i(t)\Big],$$
$$F^{+}(q_a,p_b)=\int_{t_a}^{t_b}\left(\frac{\partial\bar{q}(t;q_a,p_b)}{\partial p_b}\right)^T\Big[F\big(\bar{q}(t;q_a,p_b),\bar{p}(t;q_a,p_b)\big)\,dt+\sum_{i=1}^{m}f_i\big(\bar{q}(t;q_a,p_b),\bar{p}(t;q_a,p_b)\big)\circ dW^i(t)\Big]. \tag{2.14}$$

Below we prove that the functions S and $F^{\pm}$ generate[1] the stochastic flow F_{t_b,t_a}.

Theorem 2.2. *The function $S(q_a,p_b)$ is a type-II stochastic generating function and the functions $F^{\pm}(q_a,p_b)$ are type-II stochastic exact discrete forces for the stochastic mapping F_{t_b,t_a}, that is, $F_{t_b,t_a}:(q_a,p_a)\longrightarrow(q_b,p_b)$ is implicitly given by the equations*

$$q_b=D_2S(q_a,p_b)-F^{+}(q_a,p_b),\qquad\qquad p_a=D_1S(q_a,p_b)-F^{-}(q_a,p_b), \tag{2.15}$$

where the derivatives are understood in the mean-square sense.

Proof. Under appropriate regularity assumptions on the Hamiltonians and forces (e.g., (H1)-(H3)), the solutions $\bar{q}(t;q_a,p_b)$ and $\bar{p}(t;q_a,p_b)$ are mean-square differentiable with respect to the parameters q_a and p_b, and the partial derivatives are semimartingales (see [9]). We calculate the derivative of S as

$$\begin{aligned}\frac{\partial S}{\partial q_a}(q_a,p_b)&=\left(\frac{\partial\bar{q}(t_b)}{\partial q_a}\right)^T p_b-\int_{t_a}^{t_b}\left(\frac{\partial\bar{p}(t)}{\partial q_a}\right)^T\circ d_t\bar{q}(t)-\int_{t_a}^{t_b}d_t\left(\frac{\partial\bar{q}(t)}{\partial q_a}\right)^T\circ\bar{p}(t)\\&\quad+\int_{t_a}^{t_b}\left[\left(\frac{\partial\bar{q}(t)}{\partial q_a}\right)^T\frac{\partial H}{\partial q}\big(\bar{q}(t),\bar{p}(t)\big)+\left(\frac{\partial\bar{p}(t)}{\partial q_a}\right)^T\frac{\partial H}{\partial p}\big(\bar{q}(t),\bar{p}(t)\big)\right]dt\\&\quad+\sum_{i=1}^{m}\int_{t_a}^{t_b}\left[\left(\frac{\partial\bar{q}(t)}{\partial q_a}\right)^T\frac{\partial h_i}{\partial q}\big(\bar{q}(t),\bar{p}(t)\big)+\left(\frac{\partial\bar{p}(t)}{\partial q_a}\right)^T\frac{\partial h_i}{\partial p}\big(\bar{q}(t),\bar{p}(t)\big)\right]\circ dW^i(t),\end{aligned} \tag{2.16}$$

where for notational convenience we have omitted writing q_a and p_b explicitly as arguments of $\bar{q}(t)$ and $\bar{p}(t)$. Applying the integration by parts formula for semimartingales (see [75]), we find

$$\int_{t_a}^{t_b}d_t\left(\frac{\partial\bar{q}(t)}{\partial q_a}\right)^T\circ\bar{p}(t)=\left(\frac{\partial\bar{q}(t_b)}{\partial q_a}\right)^T p_b-\bar{p}(t_a)-\int_{t_a}^{t_b}\left(\frac{\partial\bar{q}(t)}{\partial q_a}\right)^T\circ d_t\bar{p}(t), \tag{2.17}$$

where the left-hand side integral is understood as a column vector with the components given by

$$\sum_{j=1}^{N}\int_{t_a}^{t_b}\bar{p}^j(t)\circ d_t\frac{\partial\bar{q}^j(t)}{\partial q_a^i}, \tag{2.18}$$

for each $i=1,\dots,N$. Substituting and rearranging terms, we obtain

[1] A generating function for the transformation $(q_a,p_a)\longrightarrow(q_b,p_b)$ is a function of one of the variables (q_a,p_a) and one of the variables (q_b,p_b). Therefore, there are four basic types of generating functions: $S=S_1(q_a,q_b)$, $S=S_2(q_a,p_b)$, $S=S_3(p_a,q_b)$, and $S=S_4(p_a,p_b)$. In this work we use the type-II generating function $S=S_2(q_a,p_b)$.

$$\begin{aligned}\frac{\partial S}{\partial q_a}(q_a,p_b) &= \bar{p}(t_a) + \int_{t_a}^{t_b}\left(\frac{\partial \bar{q}(t)}{\partial q_a}\right)^T\left[\circ\, d_t\bar{p} + \frac{\partial H}{\partial q}\big(\bar{q}(t),\bar{p}(t)\big)\,dt + \sum_{i=1}^{m}\frac{\partial h_i}{\partial q}\big(\bar{q}(t),\bar{p}(t)\big)\circ dW^i(t)\right]\\ &\quad + \int_{t_a}^{t_b}\left(\frac{\partial \bar{p}(t)}{\partial q_a}\right)^T\left[\circ\, d_t\bar{q} - \frac{\partial H}{\partial p}\big(\bar{q}(t),\bar{p}(t)\big)\,dt - \sum_{i=1}^{m}\frac{\partial h_i}{\partial p}\big(\bar{q}(t),\bar{p}(t)\big)\circ dW^i(t)\right]\\ &= \bar{p}(t_a) + \int_{t_a}^{t_b}\left(\frac{\partial \bar{q}(t)}{\partial q_a}\right)^T\Big[F\big(\bar{q}(t),\bar{p}(t)\big)\,dt + \sum_{i=1}^{m} f_i\big(\bar{q}(t),\bar{p}(t)\big)\circ dW^i(t)\Big],\\ &= \bar{p}(t_a) + F^-(q_a,p_b), \end{aligned} \tag{2.19}$$

since $(\bar{q}(t),\bar{p}(t))$ is a solution of (1.1). After performing similar manipulations for $\partial S/\partial p_b(q_a,p_b)$, together we obtain the result

$$\bar{q}(t_b) = D_2S(q_a,p_b) - F^+(q_a,p_b), \qquad \bar{p}(t_a) = D_1S(q_a,p_b) - F^-(q_a,p_b). \tag{2.20}$$

By definition of the flow, then $F_{t_b,t_a}(q_a,\bar{p}(t_a)) = (\bar{q}(t_b),p_b)$.

□

2.3 Stochastic forced Noether's theorem

Let a Lie group G act on Q by the left action $\Phi : G\times Q \longrightarrow Q$. The Lie group G then acts on TQ and T^*Q by the tangent $\Phi^{TQ} : G\times TQ \longrightarrow TQ$ and cotangent $\Phi^{T^*Q} : G\times T^*Q \longrightarrow T^*Q$ lift actions, respectively, given in coordinates by the formulas (see [32], [58])

$$\begin{aligned}\Phi_g^{TQ}(q,\dot{q}) &\equiv \Phi^{TQ}\big(g,(q,\dot{q})\big) = \left(\Phi_g^i(q), \frac{\partial \Phi_g^i}{\partial q^j}(q)\dot{q}^j\right),\\ \Phi_g^{T^*Q}(q,p) &\equiv \Phi^{T^*Q}\big(g,(q,p)\big) = \left(\Phi_g^i(q), p_j\frac{\partial \Phi^j_{g^{-1}}}{\partial q^i}\big(\Phi_g(q)\big)\right),\end{aligned} \tag{2.21}$$

where $i,j = 1,\dots,N$ and summation is implied over repeated indices. Let $\mathfrak{g}$ denote the Lie algebra of G and $\exp : \mathfrak{g} \longrightarrow G$ the exponential map (see [32], [58]). Each element $\xi \in \mathfrak{g}$ defines the infinitesimal generators ξ_Q, ξ_{TQ}, and ξ_{T^*Q}, which are vector fields on Q, TQ, and T^*Q, respectively, given by

$$\xi_Q(q) = \frac{d}{d\lambda}\Big|_{\lambda=0}\Phi_{\exp\lambda\xi}(q), \qquad \xi_{TQ}(q,\dot{q}) = \frac{d}{d\lambda}\Big|_{\lambda=0}\Phi^{TQ}_{\exp\lambda\xi}(q,\dot{q}), \qquad \xi_{T^*Q}(q,p) = \frac{d}{d\lambda}\Big|_{\lambda=0}\Phi^{T^*Q}_{\exp\lambda\xi}(q,p). \tag{2.22}$$

The momentum map $J : T^*Q \longrightarrow \mathfrak{g}^*$ associated with the action Φ^{T^*Q} is defined as the mapping such that for all $\xi \in \mathfrak{g}$ the function $J_\xi : T^*Q \ni (q,p) \longrightarrow \langle J(q,p),\xi\rangle \in \mathbb{R}$ is the Hamiltonian for the infinitesimal generator ξ_{T^*Q}, i.e.,

$$\xi^q_{T^*Q} = \frac{\partial J_\xi}{\partial p}, \qquad \xi^p_{T^*Q} = -\frac{\partial J_\xi}{\partial q}, \tag{2.23}$$

where $\xi_{T^*Q}(q,p) = \big(q,p,\xi^q_{T^*Q}(q,p),\xi^p_{T^*Q}(q,p)\big)$. The momentum map J can be explicitly expressed as (see [32], [58])

$$J_\xi(q,p) = p\cdot\xi_Q(q). \tag{2.24}$$

Noether's theorem for deterministic Hamiltonian systems relates symmetries of the Hamiltonian to quantities preserved by the flow of the system (see [32], [58]). When the Hamiltonian system is subject to external forces that are orthogonal to the infinitesimal generators of the symmetry group, then the corresponding momentum maps are still conserved (see [60]). It turns out that this result carries over to the stochastic case, as well. A stochastic version of Noether's theorem for systems without forcing was proved in [11], [35], and [52]. Below we state and provide a proof of Noether's theorem for stochastic forced Hamiltonian systems.

Theorem 2.3 (**Stochastic forced Noether's theorem**). *Suppose that the Hamiltonians $H : T^*Q \longrightarrow \mathbb{R}$ and $h_i : T^*Q \longrightarrow \mathbb{R}$ for $i = 1,\dots,m$ are invariant with respect to the cotangent lift action $\Phi^{T^*Q} : G\times T^*Q \longrightarrow T^*Q$ of the Lie group G, that is,*

$$H\circ\Phi^{T^*Q}_g = H, \qquad h_i\circ\Phi^{T^*Q}_g = h_i, \qquad i = 1,\dots,m, \tag{2.25}$$

for all $g\in G$. If the forcing terms are orthogonal to the infinitesimal generators of G, that is,

$$F(q,p)\cdot\xi_Q(q) = 0, \qquad f_i(q,p)\cdot\xi_Q(q) = 0, \qquad i = 1,\dots,m, \tag{2.26}$$

for all $\xi\in\mathfrak{g}$ and $(q,p)\in T^*Q$, then the cotangent lift momentum map $J : T^*Q \longrightarrow \mathfrak{g}^*$ associated with Φ^{T^*Q} is almost surely preserved along the solutions of the stochastic forced Hamiltonian system (1.1).

Proof. Equation (2.25) implies that the Hamiltonians are infinitesimally invariant with respect to the action of G, that is, for all $\xi\in\mathfrak{g}$ we have

$$dH\cdot\xi_{T^*Q} = 0, \qquad dh\cdot\xi_{T^*Q} = 0, \tag{2.27}$$

where dH and dh denote differentials with respect to the variables q and p. Let $(q(t),p(t))$ be a solution of (1.1) and consider the stochastic process $J_\xi(q(t),p(t))$, where $\xi\in\mathfrak{g}$ is arbitrary. Using the rules of Stratonovich calculus we can calculate the stochastic differential

$$\begin{aligned}
d_tJ_\xi\big(q(t),p(t)\big) &= \frac{\partial J_\xi}{\partial q}(q(t),p(t))\circ d_tq(t) + \frac{\partial J_\xi}{\partial p}(q(t),p(t))\circ d_tp(t)\\
&= \left(-\frac{\partial H}{\partial q}\xi^q_{T^*Q} - \frac{\partial H}{\partial p}\xi^p_{T^*Q} + F\cdot\xi^q_{T^*Q}\right)dt + \sum_{i=1}^m\left(-\frac{\partial h_i}{\partial q}\xi^q_{T^*Q} - \frac{\partial h_i}{\partial p}\xi^p_{T^*Q} + f_i\cdot\xi^q_{T^*Q}\right)\circ dW^i(t)\\
&= \big(-dH\cdot\xi_{T^*Q} + F\cdot\xi^q_{T^*Q}\big)\,dt + \sum_{i=1}^m\big(-dh_i\cdot\xi_{T^*Q} + f_i\cdot\xi^q_{T^*Q}\big)\circ dW^i(t)\\
&= F(q(t),p(t))\cdot\xi_Q(q(t))\,dt + \sum_{i=1}^m f_i(q(t),p(t))\cdot\xi_Q(q(t))\circ dW^i(t),
\end{aligned} \tag{2.28}$$

where we used (1.1), (2.23), (2.24), and (2.27). Therefore, if (2.26) holds, then $J_\xi\big(q(t),p(t)\big) = \text{const}$ almost surely for all $\xi \in \mathfrak{g}$, which completes the proof.

□

Remark. When the external forces are not all orthogonal to the infinitesimal generators of the symmetry group, formula (2.28) provides the rate of change of the momentum map.

2.4 Conformal symplecticity and phase space volume

The flow F_{t,t_0} for stochastic Hamiltonian systems without forcing almost surely preserves the canonical symplectic two-form

$$\Omega_{T^*Q} = dq \wedge dp = \sum_{i=1}^{N} dq^i \wedge dp^i, \tag{2.29}$$

that is, $F_{t,t_0}^* \Omega_{T^*Q} = \Omega_{T^*Q}$, where F_{t,t_0}^* denotes the pull-back by the flow F_{t,t_0} (see [69], [11], [52]). This property does not hold for the general stochastic forced Hamiltonian system (1.1). However, for certain choices of the forcing terms, the flow may be *conformally symplectic*, which means that for all $t \geq t_0$ there exists a constant (possibly random) $c_{t,t_0} \in \mathbb{R}$ such that

$$F_{t,t_0}^* \Omega_{T^*Q} = c_{t,t_0}\, \Omega_{T^*Q}. \tag{2.30}$$

Deterministic conformally symplectic systems are considered in [63]. Conformal symplecticity for the special case of (1.1) with a separable Hamiltonian, an additive noise, and the forcing terms equal to $F(q,p) = -\nu p$ with a real parameter ν, and $f_i(q,p) = 0$ for $i = 1,\dots,m$, was considered in [12] and [36]. Below we demonstrate that the property of conformal symplecticity persists for more general cases.

Theorem 2.4 (**Conformal symplecticity**). *Suppose that $H(q,p)$, $F(q,p)$, and $h_i(q,p)$, $f_i(q,p)$ for $i = 1,\dots,m$ satisfy conditions (H1)-(H3). If the forcing terms have the form*

$$F(q,p) = -\nu_0 p, \qquad f_i(q,p) = -\nu_i p, \qquad i = 1,\dots,m, \tag{2.31}$$

for real parameters ν_i, then the stochastic flow F_{t,t_0} for (1.1) is almost surely conformally symplectic with the parameter c_{t,t_0} in (2.30) given by

$$c_{t,t_0} = \exp\Big(-\nu_0(t-t_0) - \sum_{i=1}^{m} \nu_i\big(W^i(t) - W^i(t_0)\big)\Big) \tag{2.32}$$

for all $t \geq t_0$.

Proof. For fixed $(q,p) \in T^*Q$, the stochastic process $F_{t,t_0}(q,p)$ satisfies the system (1.1), which can be written as

$$d_t F_{t,t_0}(q,p) = X\big(F_{t,t_0}(q,p)\big)\, dt + \sum_{i=1}^{m} Y_i\big(F_{t,t_0}(q,p)\big) \circ dW^i(t), \tag{2.33}$$

where X and Y_i are vector fields on T^*Q, and are given by, respectively,

$$X = \frac{\partial H}{\partial p}\frac{\partial}{\partial q} + \left[-\frac{\partial H}{\partial q} + F(q,p)\right]\frac{\partial}{\partial p}, \qquad Y_i = \frac{\partial h_i}{\partial p}\frac{\partial}{\partial q} + \left[-\frac{\partial h_i}{\partial q} + f_i(q,p)\right]\frac{\partial}{\partial p}, \qquad i = 1,\dots,m. \tag{2.34}$$

Let us calculate the stochastic differential of $F^*_{t,t_0}\Omega_{T^*Q}$. Using the stochastic generalization of the dynamic definition of the Lie derivative (see Theorem 1.2 in [33]), we can write

$$d_t(F^*_{t,t_0}\Omega_{T^*Q}) = F^*_{t,t_0}(\pounds_X\Omega_{T^*Q})\,dt + \sum_{i=1}^{m} F^*_{t,t_0}(\pounds_{Y_i}\Omega_{T^*Q}) \circ dW^i(t), \tag{2.35}$$

where $\pounds_X$ and $\pounds_{Y_i}$ denote the Lie derivatives with respect to the vector fields X and Y_i, respectively. Using Cartan's magic formula (see, e.g., [3]) we have that

$$\pounds_X\Omega_{T^*Q} = di_X\Omega_{T^*Q} + i_X d\Omega_{T^*Q} = di_X\Omega_{T^*Q}, \tag{2.36}$$

since $d\Omega_{T^*Q} = 0$, where i_X denotes the interior product with the vector field X. Substituting (2.34), (2.31), and (2.29), we obtain

$$\pounds_X\Omega_{T^*Q} = -\nu_0\,\Omega_{T^*Q}, \tag{2.37}$$

since the Hamiltonian function H is C^2. In a similar fashion we show that $\pounds_{Y_i}\Omega_{T^*Q} = -\nu_i\Omega_{T^*Q}$. Plugging this in (2.35), we obtain a stochastic differential equation of the form

$$d_t(F^*_{t,t_0}\Omega_{T^*Q}) = -\nu_0(F^*_{t,t_0}\Omega_{T^*Q})\,dt - \sum_{i=1}^{m} \nu_i(F^*_{t,t_0}\Omega_{T^*Q}) \circ dW^i(t). \tag{2.38}$$

It is straightforward to verify that the solution of (2.38) that satisfies the initial condition $F^*_{t_0,t_0}\Omega_{T^*Q} = \Omega_{T^*Q}$ has the form

$$F^*_{t,t_0}\Omega_{T^*Q} = c_{t,t_0}\,\Omega_{T^*Q} \tag{2.39}$$

with c_{t,t_0} given by (2.32), which proves the conformal symplecticity of the flow F^*_{t,t_0}. It holds almost surely, since the solution of the SDE (2.38) is pathwise unique (see [9], [38], [44], [49]).

□

The evolution of stochastic Hamiltonian systems without forcing preserves volumes in phase space, that is, for the standard volume form on T^*Q defined as

$$\mu = dq^1 \wedge \ldots \wedge dq^N \wedge dp^1 \wedge \ldots dp^N \tag{2.40}$$

we have that $F^*_{t,t_0}\mu = \mu$. This is a direct consequence of the symplecticity of the flow. Phase space volume preservation does not hold for the general forced system (1.1), although for certain choices of the forcing terms the flow F^*_{t,t_0} may possess a property similar to (2.30). Such a property was proved for the special case of (1.1) with a separable Hamiltonian, an additive noise, and the forcing terms equal to $F(q,p) = -\Gamma p$ with a constant $N \times N$ matrix Γ, and $f_i(q,p) = 0$ for $i = 1,\dots,m$ (see [11], [36], [66], [67], [68], [69], [70]). Below we demonstrate that this property holds also for more general cases.

Theorem 2.5 (Phase space volume evolution). *Suppose that $H(q,p)$, $F(q,p)$, and $h_i(q,p)$, $f_i(q,p)$ for $i = 1,\dots,m$ satisfy conditions (H1)-(H3). If the forcing terms have the form*

$$F(q,p) = -\Gamma_0 p, \qquad f_i(q,p) = -\Gamma_i p, \qquad i = 1,\dots,m, \tag{2.41}$$

for constant $N \times N$ matrices Γ_i, then the phase space volume form μ for $t \geq t_0$ almost surely evolves according to the formula

$$F^*_{t,t_0}\mu = b_{t,t_0}\,\mu, \tag{2.42}$$

where

$$b_{t,t_0} = \exp\Big(-\operatorname{tr}\Gamma_0 \cdot (t-t_0) - \sum_{i=1}^{m} \operatorname{tr}\Gamma_i \cdot \big(W^i(t) - W^i(t_0)\big)\Big), \tag{2.43}$$

and F_{t,t_0} is the stochastic flow for (1.1).

Proof. Similar to (2.35), we can write

$$d_t(F^*_{t,t_0}\mu) = F^*_{t,t_0}(\pounds_X \mu)\,dt + \sum_{i=1}^{m} F^*_{t,t_0}(\pounds_{Y_i}\mu) \circ dW^i(t). \tag{2.44}$$

Using the property of the divergence operator (see, e.g., [3]), we calculate

$$\pounds_X \mu = (\operatorname{div} X)\cdot \mu = -(\operatorname{tr}\Gamma_0)\cdot\mu, \tag{2.45}$$

where we have used (2.34) and (2.41), and the fact that the Hamiltonian function H is C^2. In a similar way we show that $\pounds_{Y_i}\mu = -(\operatorname{tr}\Gamma_i)\cdot\mu$. Therefore, we obtain the SDE of the form

$$d_t(F^*_{t,t_0}\mu) = -(\operatorname{tr}\Gamma_0)\cdot(F^*_{t,t_0}\mu)\,dt - \sum_{i=1}^{m}(\operatorname{tr}\Gamma_i)\cdot(F^*_{t,t_0}\mu)\circ dW^i(t). \tag{2.46}$$

It is straightforward to verify that the solution that satisfies the initial condition $F^*_{t_0,t_0}\mu = \mu$ is given by (2.42) with b_{t,t_0} as in (2.43). The formula (2.42) holds almost surely, because the solution of the SDE is pathwise unique (see [9], [38], [44], [49]).

□

3 Stochastic Lagrange-d'Alembert variational integrators

Suppose we would like to solve (1.1) on the interval $[0,T]$ with the initial conditions $(q_0,p_0) \in T^*Q$. Consider the discrete set of times $t_k = k\cdot\Delta t$ for $k = 0,1,\dots,K$, where $\Delta t = T/K$ is the time step. In order to determine the discrete curve $\{(q_k,p_k)\}_{k=0,\dots,K}$ that approximates the exact solution of (1.1) at times t_k we need to construct an approximation of the exact stochastic flow F_{t_{k+1},t_k} on each interval $[t_k,t_{k+1}]$, so that $(q_{k+1},p_{k+1}) \approx F_{t_{k+1},t_k}(q_k,p_k)$. A numerical method respecting the underlying Lagrange-d'Alembert principle (2.6) can be constructed by approximating the generating function and forcing terms in (2.15). Let the discrete Hamiltonian function $H^+_d(q_a,p_b;t_a,t_b)$ be an approximation of the generating function (2.13), and let the discrete forces $F^\pm_d(q_a,p_b;t_a,t_b)$ be

approximations of the forcing terms (2.14). The approximate numerical flow $F^+_{t_{k+1},t_k} : (q_k, p_k) \longrightarrow (q_{k+1}, p_{k+1})$ is now generated as in (2.20):

$$\begin{aligned} q_{k+1} &= D_2 H_d^+(q_k, p_{k+1}; t_k, t_{k+1}) - F_d^+(q_k, p_{k+1}; t_k, t_{k+1}), \\ p_k &= D_1 H_d^+(q_k, p_{k+1}; t_k, t_{k+1}) - F_d^-(q_k, p_{k+1}; t_k, t_{k+1}). \end{aligned} \tag{3.1}$$

If there is no risk of confusion, we will omit writing the time arguments of H_d^+ and $F_d^\pm$. We will refer to the scheme (3.1) as a stochastic Lagrange-d'Alembert variational integrator.

3.1 Discrete stochastic Lagrange-d'Alembert principle

The advantage of the integrator (3.1) is that it follows from a discrete version of the stochastic Lagrange-d'Alembert principle (2.6). The discrete Lagrange-d'Alembert principle for deterministic Lagrangian systems was proposed in [42]; see also [60]. Below we generalize it to the stochastic case in the setting of Hamiltonian systems defined on the phase space T^*Q. Define the discrete random curve space C_d as

$$C_d = \Big\{ \big\{(q_k, p_k)\big\}_{k=0,\dots,K} \,\big|\, (q_k, p_k) : \Omega \longrightarrow T^*Q \text{ are random variables for each } k = 0, \dots, K \Big\}. \tag{3.2}$$

On that space define the discrete action functional, $\mathcal{B}_d : \Omega \times C_d \longrightarrow \mathbb{R}$,

$$\mathcal{B}_d\big[\{(q_k, p_k)\}_{k=0,\dots,K}\big] = p_K q_K - \sum_{k=0}^{K-1} \big(p_{k+1} q_{k+1} - H_d^+(q_k, p_{k+1}; t_k, t_{k+1}) \big). \tag{3.3}$$

Note that $\mathcal{B}_d$ is an approximation of the stochastic action functional (2.4) on the interval $[0, T]$.

Theorem 3.1 (**Discrete stochastic Lagrange-d'Alembert Principle in Phase Space**). *Suppose the discrete Hamiltonian H_d^+ is almost surely continuously differentiable, and the discrete forces $F_d^\pm$ are almost surely continuous with respect to their arguments. The discrete random curve $\{(q_k, p_k)\}_{k=0,\dots,K}$ satisfies the set of equations*

$$\begin{aligned} q_k &= D_2 H_d^+(q_{k-1}, p_k; t_{k-1}, t_k) - F_d^+(q_{k-1}, p_k; t_{k-1}, t_k), \\ p_k &= D_1 H_d^+(q_k, p_{k+1}; t_k, t_{k+1}) - F_d^-(q_k, p_{k+1}; t_k, t_{k+1}), \end{aligned} \tag{3.4}$$

almost surely for $k = 1, \dots, K-1$, if and only if it almost surely satisfies the variational equation

$$\delta \mathcal{B}_d - \sum_{k=0}^{K-1} \big(F_d^-(q_k, p_{k+1}) \delta q_k + F_d^+(q_k, p_{k+1}) \delta p_{k+1} \big) = 0 \tag{3.5}$$

for all variations $\{(\delta q_k, \delta p_k)\}_{k=0,\dots,K}$ such that $\delta q_0 = 0$ and $\delta p_K = 0$ almost surely.

Proof. Consider an arbitrary random curve $\{(q_k, p_k)\}_{k=0,\dots,K}$. Let us calculate the variation $\delta \mathcal{B}_d$ corresponding to the arbitrary variation $\{(\delta q_k, \delta p_k)\}_{k=0,\dots,K}$ with $\delta q_0 = 0$ and $\delta p_K = 0$ (almost surely). We have

$$\delta\mathcal{B}_d = p_K\delta q_K - \sum_{k=0}^{K-1}\Big(\delta p_{k+1}q_{k+1} + p_{k+1}\delta q_{k+1} - D_1H_d^+(q_k,p_{k+1};t_k,t_{k+1})\delta q_k - D_2H_d^+(q_k,p_{k+1};t_k,t_{k+1})\delta p_{k+1}\Big)$$
$$= -\sum_{k=0}^{K-1}\big(q_{k+1} - D_2H_d^+(q_k,p_{k+1};t_k,t_{k+1})\big)\delta p_{k+1} - \sum_{k=0}^{K-1}\big(p_k - D_1H_d^+(q_k,p_{k+1};t_k,t_{k+1})\big)\delta q_k, \tag{3.6}$$

where in the second equality we have shifted the summation index in the δq_{k+1} term and used the fact that $\delta q_0 = 0$. It is now straightforward to see that if the set of equations (3.4) is satisfied, then the variational equation (3.5) holds almost surely. Conversely, if the variational equation (3.5) holds for all variations $\{(\delta q_k, \delta p_k)\}_{k=0,\dots,K}$ with $\delta q_0 = 0$ and $\delta p_K = 0$, then the set of equations (3.4) has to be satisfied almost surely.

□

3.2 Discrete stochastic forced Noether's theorem

Another advantage of the integrator (3.1) is that one can prove the discrete counterpart of stochastic forced Noether's theorem. If the discrete system inherits the symmetries of the continuous problem, then the evolution of the momentum maps will be accurately captured by the numerical solution. Discrete Noether's theorem for systems described by a type-II generating function was first proved for deterministic systems in [54], and later generalized to the stochastic case in [35]. Discrete forced Noether's theorem for deterministic Lagrangian systems with forcing was first proposed in [60]. Below we combine these ideas and formulate a version of discrete Noether's theorem applicable to discrete systems described by (3.1). Let $R_d : \Omega\times Q\times T^*Q \longrightarrow \mathbb{R}$ be the generalized discrete stochastic Lagrangian defined as

$$R_d(q_k, q_{k+1}, p_{k+1}) = p_{k+1}q_{k+1} - H_d^+(q_k, p_{k+1}). \tag{3.7}$$

Consider the action of the Lie group G on $Q\times T^*Q$ given by

$$\Phi_g^{Q\times T^*Q}(q_k, q_{k+1}, p_{k+1}) = \big(\Phi_g(q_k), \Phi_g^{T^*Q}(q_{k+1}, p_{k+1})\big). \tag{3.8}$$

For any $\xi\in\mathfrak{g}$ the corresponding infinitesimal generator on $Q\times T^*Q$ is then given by

$$\xi_{Q\times T^*Q}(q_k, q_{k+1}, p_{k+1}) = \big(\xi_Q(q_k), \xi_{T^*Q}(q_{k+1}, p_{k+1})\big) = \big(\xi_Q(q_k), \xi^q_{T^*Q}(q_{k+1}, p_{k+1}), \xi^p_{T^*Q}(q_{k+1}, p_{k+1})\big). \tag{3.9}$$

Theorem 3.2 (Discrete stochastic forced Noether's theorem). *Suppose the generalized discrete stochastic Lagrangian $R_d : \Omega\times Q\times T^*Q \longrightarrow \mathbb{R}$ is invariant under the action of the Lie group G, that is,*

$$R_d\Big(\Phi_g^{Q\times T^*Q}(q_k, q_{k+1}, p_{k+1})\Big) = R_d(q_k, q_{k+1}, p_{k+1}), \qquad \text{for all } g\in G. \tag{3.10}$$

If the discrete forces $F_d^\pm$ satisfy the condition

$$F_d^-(q_k, p_{k+1})\cdot\xi_Q(q_k) + F_d^+(q_k, p_{k+1})\cdot\xi^p_{T^*Q}(q_{k+1}, p_{k+1}) = 0 \tag{3.11}$$

for all $(q_k, q_{k+1}, p_{k+1}) \in Q \times T^*Q$*, then the cotangent lift momentum map* J *associated with* Φ^{T^*Q} *is almost surely preserved along the solutions of the discrete equations* (3.1)*, i.e., a.s.* $J(q_{k+1}, p_{k+1}) = J(q_k, p_k)$*.*

Proof. Since the generalized discrete Lagrangian R_d is invariant with respect to the actions of G, for an arbitrary $\xi \in \mathfrak{g}$ we have

$$\begin{aligned} 0 &= \frac{d}{d\lambda}\bigg|_{\lambda=0} R_d\Big(\Phi^{Q\times T^*Q}_{\exp\lambda\xi}(q_k, q_{k+1}, p_{k+1})\Big) = dR_d \cdot \xi_{Q\times T^*Q}(q_k, q_{k+1}, p_{k+1}) \\ &= -D_1H_d^+(q_k, p_{k+1}) \cdot \xi_Q(q_k) + p_{k+1} \cdot \xi_Q(q_{k+1}) + \big(q_{k+1} - D_1H_d^+(q_k, p_{k+1})\big) \cdot \xi^p_{T^*Q}(q_{k+1}, p_{k+1}), \quad (3.12) \end{aligned}$$

where we have used the fact that $\xi^q_{T^*Q}(q_{k+1}, p_{k+1}) = \xi_Q(q_{k+1})$. Assume that q_k, q_{k+1}, and p_{k+1} satisfy the discrete evolution equation (3.1). By substituting (3.1) in (3.12), we obtain

$$0 = (-p_k - F_d^-(q_k, p_{k+1})) \cdot \xi_Q(q_k) + p_{k+1} \cdot \xi_Q(q_{k+1}) - F_d^+(q_k, p_{k+1}) \cdot \xi^p_{T^*Q}(q_{k+1}, p_{k+1}). \quad (3.13)$$

This can be rewritten as

$$J_\xi(q_{k+1}, p_{k+1}) - J_\xi(q_k, p_k) = F_d^-(q_k, p_{k+1}) \cdot \xi_Q(q_k) + F_d^+(q_k, p_{k+1}) \cdot \xi^p_{T^*Q}(q_{k+1}, p_{k+1}), \quad (3.14)$$

where we have used the definition of the cotangent lift momentum map (2.24). If the condition (3.11) holds, then we have $J_\xi(q_{k+1}, p_{k+1}) = J_\xi(q_k, p_k)$. The result holds almost surely, because equation (3.1) is satisfied almost surely. □

Remark. When the discrete forces do not satisfy the condition (3.11), equation (3.14) provides the rate of change of the momentum map, which mimicks formula (2.28) in the continuous case.

3.3 Mean-square Lagrange-d'Alembert partitioned Runge-Kutta methods

3.3.1 Construction

Partitioned Runge-Kutta methods for deterministic forced Hamiltonian systems have been proposed in [41] and [60]. A general class of stochastic mean-square Runge-Kutta methods for Stratonovich ordinary differential equations was introduced and analyzed in [16], [17], and [18]. These ideas were later used by Ma & Ding & Ding [56] and Ma & Ding [57] to construct symplectic Runge-Kutta methods for stochastic Hamiltonian systems without forcing; see also [35]. Below we combine these ideas and introduce mean-square Lagrange-d'Alembert partitioned Runge-Kutta methods for stochastic forced Hamiltonian systems of the form (1.1).

Definition 3.3. *An* s*-stage mean-square Lagrange-d'Alembert partitioned Runge-Kutta method for the system* (1.1) *is given by*

$$Q_i = q_k + \Delta t \sum_{j=1}^{s} a_{ij} \frac{\partial H}{\partial p}(Q_j, P_j) + \sum_{r=1}^{m} \Delta W^r \sum_{j=1}^{s} b_{ij} \frac{\partial h_r}{\partial p}(Q_j, P_j), \qquad i = 1, \ldots, s, \tag{3.15a}$$

$$\begin{aligned} P_i = p_k &- \Delta t \sum_{j=1}^{s} \bar{a}_{ij} \frac{\partial H}{\partial q}(Q_j, P_j) - \sum_{r=1}^{m} \Delta W^r \sum_{j=1}^{s} \bar{b}_{ij} \frac{\partial h_r}{\partial q}(Q_j, P_j) \\ &+ \Delta t \sum_{j=1}^{s} \hat{a}_{ij} F(Q_j, P_j) + \sum_{r=1}^{m} \Delta W^r \sum_{j=1}^{s} \hat{b}_{ij} f_r(Q_j, P_j), \qquad i = 1, \ldots, s, \end{aligned} \tag{3.15b}$$

$$q_{k+1} = q_k + \Delta t \sum_{i=1}^{s} \alpha_i \frac{\partial H}{\partial p}(Q_i, P_i) + \sum_{r=1}^{m} \Delta W^r \sum_{i=1}^{s} \beta_i \frac{\partial h_r}{\partial p}(Q_i, P_i), \tag{3.15c}$$

$$\begin{aligned} p_{k+1} = p_k &- \Delta t \sum_{i=1}^{s} \alpha_i \frac{\partial H}{\partial q}(Q_i, P_i) - \sum_{r=1}^{m} \Delta W^r \sum_{i=1}^{s} \beta_i \frac{\partial h_r}{\partial q}(Q_i, P_i) \\ &+ \Delta t \sum_{i=1}^{s} \hat{\alpha}_i F(Q_i, P_i) + \sum_{r=1}^{m} \Delta W^r \sum_{i=1}^{s} \hat{\beta}_i f_r(Q_i, P_i), \end{aligned} \tag{3.15d}$$

where Δt is the time step, $\Delta W = (\Delta W^1, \ldots, \Delta W^m)$ are the increments of the Wiener process, Q_i and P_i for $i = 1, \ldots, s$ are the position and momentum internal stages, respectively, and the coefficients of the method a_{ij}, $\bar{a}_{ij}$, $\hat{a}_{ij}$, b_{ij}, $\bar{b}_{ij}$, $\hat{b}_{ij}$, α_i, $\hat{\alpha}_i$, β_i, and $\hat{\beta}_i$ satisfy the conditions

$$\alpha_i \bar{a}_{ij} + \alpha_j a_{ji} = \alpha_i \alpha_j, \tag{3.16a}$$
$$\beta_i \bar{b}_{ij} + \beta_j b_{ji} = \beta_i \beta_j, \tag{3.16b}$$
$$\beta_i \bar{a}_{ij} + \alpha_j b_{ji} = \beta_i \alpha_j, \tag{3.16c}$$
$$\alpha_i \bar{b}_{ij} + \beta_j a_{ji} = \alpha_i \beta_j, \tag{3.16d}$$
$$\alpha_i \hat{a}_{ij} + \hat{\alpha}_j a_{ji} = \alpha_i \hat{\alpha}_j, \tag{3.16e}$$
$$\alpha_i \hat{b}_{ij} + \hat{\beta}_j a_{ji} = \alpha_i \hat{\beta}_j, \tag{3.16f}$$
$$\beta_i \hat{a}_{ij} + \hat{\alpha}_j b_{ji} = \beta_i \hat{\alpha}_j, \tag{3.16g}$$
$$\beta_i \hat{b}_{ij} + \hat{\beta}_j b_{ji} = \beta_i \hat{\beta}_j, \tag{3.16h}$$

for $i, j = 1, \ldots, s$.

The partitioned Runge-Kutta method (3.15) can be represented by the tableau

	a	$\bar{a}$	$\hat{a}$	b	$\bar{b}$	$\hat{b}$
	α^T	α^T	$\hat{\alpha}^T$	β^T	β^T	$\hat{\beta}^T$

(3.17)

where $a = (a_{ij})_{i,j=1\ldots s}$, $\alpha = (\alpha_i)_{i=1\ldots s}$, etc. The set of equations (3.15) forms a one-step numerical scheme. Knowing q_k and p_k at time t_k, one can solve Equations (3.15a)-(3.15b) for the internal stages Q_i and P_i, and then use (3.15c)-(3.15d) to determine q_{k+1} and p_{k+1} at time t_{k+1}. If given q_k and p_{k+1} instead, one can also solve (3.15) for the remaining variables Q_i, P_i, q_{k+1} and p_k. Note that since we have only used $\Delta W^r = \int_{t_k}^{t_{k+1}} dW^r(t)$ in (3.15), we can in general expect mean-square convergence of order 1.0 at most. To obtain mean-square convergence of higher order we would also need to include higher-order multiple Stratonovich integrals, e.g., to achieve convergence of order

1.5 we would need to include terms involving $\Delta Z^r = \int_{t_k}^{t_{k+1}} \int_{t_k}^{t} dW^r(\xi)\, dt$ (see [18], [68], [69]). Below we prove that the Runge-Kutta method (3.15) with the conditions (3.16) is indeed a stochastic Lagrange-d'Alembert method of the form (3.1).

Theorem 3.4. *The s-stage mean-square partitioned Runge-Kutta method* (3.15) *with the conditions* (3.16) *is a stochastic Lagrange-d'Alembert variational integrator of the form* (3.1) *with the discrete Hamiltonian*

$$H_d^+(q_k, p_{k+1}) = p_{k+1}q_{k+1} - \Delta t \sum_{i=1}^{s} \alpha_i \left(P_i \frac{\partial H}{\partial p}(Q_i, P_i) - H(Q_i, P_i) \right) - \sum_{r=1}^{m} \Delta W^r \sum_{i=1}^{s} \beta_i \left(P_i \frac{\partial h_r}{\partial p}(Q_i, P_i) - h_r(Q_i, P_i) \right), \tag{3.18}$$

and the discrete forces

$$\begin{aligned} F_d^-(q_k, p_{k+1}) &= \Delta t \sum_{i=1}^{s} \hat{\alpha}_i \left(\frac{\partial Q_i}{\partial q_k} \right)^T F(Q_i, P_i) + \sum_{r=1}^{m} \Delta W^r \sum_{i=1}^{s} \hat{\beta}_i \left(\frac{\partial Q_i}{\partial q_k} \right)^T f_r(Q_i, P_i), \\ F_d^+(q_k, p_{k+1}) &= \Delta t \sum_{i=1}^{s} \hat{\alpha}_i \left(\frac{\partial Q_i}{\partial p_{k+1}} \right)^T F(Q_i, P_i) + \sum_{r=1}^{m} \Delta W^r \sum_{i=1}^{s} \hat{\beta}_i \left(\frac{\partial Q_i}{\partial p_{k+1}} \right)^T f_r(Q_i, P_i), \end{aligned} \tag{3.19}$$

where q_{k+1}, p_k, Q_i, and P_i satisfy the system of equations (3.15) *and are understood as functions of q_k and p_{k+1}.*

Proof. The proof involves straightforward, although rather lengthy and tedious algebraic manipulations. Therefore, for the clarity and brevity of the exposition, we only consider the one-dimensional noise case $m = 1$ and point out the key steps of the derivations. Let us introduce the following shorthand notation:

$$\begin{aligned} \dot{Q}_i &\equiv \frac{\partial H}{\partial p}(Q_i, P_i), \quad \dot{P}_i \equiv -\frac{\partial H}{\partial q}(Q_i, P_i), \quad F_i \equiv F(Q_i, P_i), \\ \dot{K}_i &\equiv \frac{\partial h}{\partial p}(Q_i, P_i), \quad \dot{G}_i \equiv -\frac{\partial h}{\partial q}(Q_i, P_i), \quad f_i \equiv f(Q_i, P_i). \end{aligned} \tag{3.20}$$

Differentiate each of the equations (3.15) with respect to q_k and p_{k+1} to express the Jacobians $\partial Q_i/\partial q_k$, $\partial P_i/\partial q_k$, $\partial q_{k+1}/\partial q_k$, $\partial p_k/\partial q_k$, and analogous Jacobians with respect to p_{k+1}, in terms of the derivatives of the terms (3.20). For instance, we have

$$\frac{\partial P_i}{\partial p_{k+1}} = I + \Delta t \sum_{j=1}^{s} (\bar{a}_{ij} - \alpha_j) \frac{\partial \dot{P}_j}{\partial p_{k+1}} + \Delta W \sum_{j=1}^{s} (\bar{b}_{ij} - \beta_j) \frac{\partial \dot{K}_j}{\partial p_{k+1}} + \Delta t \sum_{j=1}^{s} (\hat{a}_{ij} - \hat{\alpha}_j) \frac{\partial F_j}{\partial p_{k+1}} + \Delta W \sum_{j=1}^{s} (\hat{b}_{ij} - \hat{\beta}_j) \frac{\partial f_j}{\partial p_{k+1}}, \tag{3.21}$$

where I denotes the $N \times N$ identity matrix. Let us now calculate the derivative of the discrete Hamiltonian (3.18) with respect to p_{k+1}. After substituting the Jacobians (3.21) and using (3.15d) to replace p_{k+1}, we obtain the expression

$$
\begin{aligned}
D_2H_d^+(q_k,p_{k+1}) = q_{k+1} &+ \Delta t^2 \sum_{i,j=1}^{s} \alpha_i\hat{\alpha}_j\left(\frac{\partial \dot{Q}_i}{\partial p_{k+1}}\right)^T F_j + \Delta t\Delta W \sum_{i,j=1}^{s} \alpha_i\hat{\beta}_j\left(\frac{\partial \dot{Q}_i}{\partial p_{k+1}}\right)^T f_j \\
&+ \Delta t\Delta W \sum_{i,j=1}^{s} \beta_i\hat{\alpha}_j\left(\frac{\partial \dot{K}_i}{\partial p_{k+1}}\right)^T F_j + \Delta W^2 \sum_{i,j=1}^{s} \beta_i\hat{\beta}_j\left(\frac{\partial \dot{K}_i}{\partial p_{k+1}}\right)^T f_j \\
&+ \Delta t \sum_{i=1}^{s} \alpha_i\left(\frac{\partial \dot{Q}_i}{\partial p_{k+1}}\right)^T (p_k - P_i) + \Delta W \sum_{i=1}^{s} \beta_i\left(\frac{\partial \dot{K}_i}{\partial p_{k+1}}\right)^T (p_k - P_i) \\
&+ \Delta t^2 \sum_{i,j=1}^{s} (\alpha_i\alpha_j - \alpha_j a_{ji})\left(\frac{\partial \dot{Q}_i}{\partial p_{k+1}}\right)^T \dot{P}_j + \Delta t\Delta W \sum_{i,j=1}^{s} (\alpha_i\beta_j - \beta_j a_{ji})\left(\frac{\partial \dot{Q}_i}{\partial p_{k+1}}\right)^T \dot{G}_j \\
&+ \Delta t\Delta W \sum_{i,j=1}^{s} (\beta_i\alpha_j - \alpha_j b_{ji})\left(\frac{\partial \dot{K}_i}{\partial p_{k+1}}\right)^T \dot{P}_j + \Delta W^2 \sum_{i,j=1}^{s} (\beta_i\beta_j - \beta_j b_{ji})\left(\frac{\partial \dot{K}_i}{\partial p_{k+1}}\right)^T \dot{G}_j.
\end{aligned} \tag{3.22}
$$

After using (3.16a)-(3.16d) in the last four terms (e.g., $\alpha_i\alpha_j - \alpha_j a_{ji} = \alpha_i\bar{a}_{ij}$), and substituting (3.15b) for P_i, we get

$$
\begin{aligned}
D_2H_d^+(q_k,p_{k+1}) = q_{k+1} &+ \Delta t^2 \sum_{i,j=1}^{s} (\alpha_i\hat{\alpha}_j - \alpha_i\hat{a}_{ij})\left(\frac{\partial \dot{Q}_i}{\partial p_{k+1}}\right)^T F_j + \Delta t\Delta W \sum_{i,j=1}^{s} (\alpha_i\hat{\beta}_j - \alpha_i\hat{b}_{ij})\left(\frac{\partial \dot{Q}_i}{\partial p_{k+1}}\right)^T f_j \\
&+ \Delta t\Delta W \sum_{i,j=1}^{s} (\beta_i\hat{\alpha}_j - \beta_i\hat{a}_{ij})\left(\frac{\partial \dot{K}_i}{\partial p_{k+1}}\right)^T F_j + \Delta W^2 \sum_{i,j=1}^{s} (\beta_i\hat{\beta}_j - \beta_i\hat{b}_{ij})\left(\frac{\partial \dot{K}_i}{\partial p_{k+1}}\right)^T f_j.
\end{aligned} \tag{3.23}
$$

By using the conditions (3.16e)-(3.16h) and collecting terms, we finally arrive at

$$
D_2H_d^+(q_k,p_{k+1}) = q_{k+1} + \Delta t \sum_{i=1}^{s} \hat{\alpha}_i\left(\frac{\partial Q_i}{\partial p_{k+1}}\right)^T F_i + \Delta W \sum_{i=1}^{s} \hat{\beta}_i\left(\frac{\partial Q_i}{\partial p_{k+1}}\right)^T f_i = q_{k+1} + F_d^+(q_k,p_{k+1}). \tag{3.24}
$$

In a similar fashion we derive

$$
D_1H_d^+(q_k,p_{k+1}) = p_k + F_d^-(q_k,p_{k+1}). \tag{3.25}
$$

Therefore, we have proved that q_k, p_k, q_{k+1}, and p_{k+1} satisfy (3.1) if and only if they satisfy (3.15). □

3.3.2 Convergence

Mean-square convergence concentrates on pathwise approximations of the exact solutions (see [44], [65]). Let $\bar{z}(t) = (\bar{q}(t), \bar{p}(t))$ be the exact solution to (1.1) with the initial conditions q_0 and p_0, and let $z_k = (q_k, p_k)$ denote the numerical solution at time t_k obtained by applying (3.15) iteratively k

times with the constant time step Δt. The numerical solution is said to converge in the mean-square sense with global order r if there exist $\delta > 0$ and a constant $C > 0$ such that for all $\Delta t \in (0, \delta)$ we have

$$\sqrt{E(\|z_K - \bar{z}(T)\|^2)} \leq C\Delta t^r, \tag{3.26}$$

where $T = K\Delta t$, as defined before, and E denotes the expected value. In principle, in order to determine the mean-square order of convergence of the Lagrange-d'Alembert partitioned Runge-Kutta method (3.15) we need to calculate the power series expansions of q_{k+1} and p_{k+1} in terms of the powers of Δt and ΔW^i, and compare them to the Stratonovich-Taylor expansions for the exact solution $\bar{q}(t_k + \Delta t)$ and $\bar{p}(t_k + \Delta t)$ (see [18], [44], [65]). As mentioned in Section 3.3.1, the mean-square order of the method (3.15) cannot exceed 1.0. Below we provide the conditions that have to be satisfied by the coefficients of the method (3.15) in order for it to be convergent.

Theorem 3.5. *Suppose that, in addition to conditions (H1)-(H3), the functions $H(q,p)$, $F(q,p)$, and $h_i(q,p)$, $f_i(q,p)$ for $i = 1, \ldots, m$ have all the necessary partial derivatives. Let the coefficients of the method* (3.15) *satisfy the conditions*

$$\begin{aligned}
&\sum_{i=1}^{s} \alpha_i = \sum_{i=1}^{s} \hat{\alpha}_i = \sum_{i=1}^{s} \beta_i = \sum_{i=1}^{s} \hat{\beta}_i = 1, \\
&\sum_{i,j=1}^{s} \beta_i b_{ij} = \sum_{i,j=1}^{s} \beta_i \bar{b}_{ij} = \sum_{i,j=1}^{s} \beta_i \hat{b}_{ij} = \sum_{i,j=1}^{s} \hat{\beta}_i b_{ij} = \sum_{i,j=1}^{s} \hat{\beta}_i \bar{b}_{ij} = \sum_{i,j=1}^{s} \hat{\beta}_i \hat{b}_{ij} = \frac{1}{2}.
\end{aligned} \tag{3.27}$$

If the noise is commutative, that is, if the following conditions are satisfied

$$\Gamma_{ij} = \Gamma_{ji}, \qquad \Lambda_{ij} = \Lambda_{ji}, \qquad \text{for all } i, j = 1, \ldots, m, \tag{3.28}$$

where the vectors Γ_{ij} and Λ_{ij} for each $i, j = 1, \ldots, m$ are defined as

$$\begin{aligned}
\Gamma_{ij} &= \frac{\partial^2 h_j}{\partial p \partial q} \frac{\partial h_i}{\partial p} - \frac{\partial^2 h_j}{\partial p^2} \frac{\partial h_i}{\partial q} + \frac{\partial^2 h_j}{\partial p^2} f_i, \\
\Lambda_{ij} &= -\frac{\partial^2 h_j}{\partial q^2} \frac{\partial h_i}{\partial p} + \frac{\partial^2 h_j}{\partial q \partial p} \frac{\partial h_i}{\partial q} + \frac{\partial f_j}{\partial q} \frac{\partial h_i}{\partial p} - \frac{\partial f_j}{\partial p} \frac{\partial h_i}{\partial q} - \frac{\partial^2 h_j}{\partial q \partial p} f_i + \frac{\partial f_j}{\partial p} f_i,
\end{aligned} \tag{3.29}$$

then the method (3.15) *is convergent with mean-square order 1.0. If the noise is noncommutative, then the method* (3.15) *is convergent with mean-square order 0.5.*

Proof. General order conditions for stochastic non-partitioned Runge-Kutta methods have been analyzed in [17] and [18]. Conditions for mean-square convergence of order 1.0 for stochastic partitioned Runge-Kutta methods with a one-dimensional noise have been derived in [57]. However, the method (3.15) is more general, as we allow a multidimensional noise, and different coefficients are applied to the Hamiltonian and forcing terms, but the method of proof is similar to the proof of Theorem 2.1 in [57], therefore we only present the main steps. To simplify the notation, denote $\alpha = (\alpha_1, \ldots, \alpha_s)^T$, $b = (b_{ij})_{i,j=1,\ldots,s}$, and similarly for the remaining coefficients of the method. Let also $e = (1, 1, \ldots, 1)^T$ be an s-dimensional vector. Then the conditions (3.27) can be written more

compactly, e.g., $\alpha^T e = 1$ or $\beta^T b e = 1/2$. We first determine power expansions of the internal stages Q_i and P_i in terms of the powers of Δt and ΔW^i. We plug in series expansions for Q_i and P_i in Equations (3.15a)-(3.15b), and determine their coefficients by expanding the derivatives of the Hamiltonians and forcing terms into Taylor series around (q_k, p_k). Then we plug in thus found series expansions into Equations (3.15c)-(3.15d), and again expand the derivatives of the Hamiltonians and forcing terms into Taylor series around (q_k, p_k). This way we obtain the series expansions of q_{k+1} and p_{k+1} as

$$\begin{aligned}
q_{k+1} &= q_k + (\alpha^T e)\frac{\partial H}{\partial p}\Delta t + (\beta^T e)\sum_{i=1}^{m}\frac{\partial h_i}{\partial p}\Delta W^i + \frac{1}{2}\sum_{i=1}^{M}\bar{\Gamma}_{ii}(\Delta W^i)^2 + \frac{1}{2}\sum_{i=1}^{M}\sum_{\substack{j=1\\ j\neq i}}^{M}\bar{\Gamma}_{ij}\Delta W^i \Delta W^j + \ldots,\\
p_{k+1} &= p_k - (\alpha^T e)\frac{\partial H}{\partial q}\Delta t + (\hat{\alpha}^T e)F\Delta t - (\beta^T e)\sum_{i=1}^{m}\frac{\partial h_i}{\partial q}\Delta W^i + (\hat{\beta}^T e)\sum_{i=1}^{m} f_i \Delta W^i\\
&\quad + \frac{1}{2}\sum_{i=1}^{M}\bar{\Lambda}_{ii}(\Delta W^i)^2 + \frac{1}{2}\sum_{i=1}^{M}\sum_{\substack{j=1\\ j\neq i}}^{M}\bar{\Lambda}_{ij}\Delta W^i \Delta W^j + \ldots,
\end{aligned} \tag{3.30}$$

where the vectors $\bar{\Gamma}_{ij}$ and $\bar{\Lambda}_{ij}$ for each $i, j = 1, \ldots, m$ are defined as

$$\begin{aligned}
\bar{\Gamma}_{ij} &= 2(\beta^T b e)\frac{\partial^2 h_j}{\partial p \partial q}\frac{\partial h_i}{\partial p} - 2(\beta^T \bar{b} e)\frac{\partial^2 h_j}{\partial p^2}\frac{\partial h_i}{\partial q} + 2(\beta^T \hat{b} e)\frac{\partial^2 h_j}{\partial p^2} f_i,\\
\bar{\Lambda}_{ij} &= -2(\beta^T b e)\frac{\partial^2 h_j}{\partial q^2}\frac{\partial h_i}{\partial p} + 2(\beta^T \bar{b} e)\frac{\partial^2 h_j}{\partial q \partial p}\frac{\partial h_i}{\partial q} + 2(\hat{\beta}^T b e)\frac{\partial f_j}{\partial q}\frac{\partial h_i}{\partial p} - 2(\hat{\beta}^T \bar{b} e)\frac{\partial f_j}{\partial p}\frac{\partial h_i}{\partial q}\\
&\quad - 2(\beta^T \hat{b} e)\frac{\partial^2 h_j}{\partial q \partial p} f_i + 2(\hat{\beta}^T \hat{b} e)\frac{\partial f_j}{\partial p} f_i,
\end{aligned} \tag{3.31}$$

and the forcing terms and the derivatives of the Hamiltonians are evaluated at (q_k, p_k). Let $\bar{q}(t; q_k, p_k)$ and $\bar{p}(t; q_k, p_k)$ denote the exact solution of (1.1) such that $\bar{q}(t_k; q_k, p_k) = q_k$ and $\bar{p}(t_k; q_k, p_k) = p_k$. Using (1.1) we calculate the Stratonovich-Taylor expansions for $\bar{q}(t_{k+1}; q_k, p_k)$ and $\bar{p}(t_{k+1}; q_k, p_k)$ as (see [44])

$$\begin{aligned}
\bar{q}(t_{k+1}; q_k, p_k) &= q_k + \frac{\partial H}{\partial p}\Delta t + \sum_{i=1}^{m}\frac{\partial h_i}{\partial p}\Delta W^i + \frac{1}{2}\sum_{i=1}^{m}\Gamma_{ii}(\Delta W^i)^2 + \sum_{i=1}^{m}\sum_{\substack{j=1\\ j\neq i}}^{m}\Gamma_{ij} J_{ij} + \ldots,\\
\bar{p}(t_{k+1}; q_k, p_k) &= p_k + \left(-\frac{\partial H}{\partial q} + F\right)\Delta t + \sum_{i=1}^{m}\left(-\frac{\partial h_i}{\partial q} + f_i\right)\Delta W^i + \frac{1}{2}\sum_{i=1}^{m}\Lambda_{ii}(\Delta W^i)^2 + \sum_{i=1}^{m}\sum_{\substack{j=1\\ j\neq i}}^{m}\Lambda_{ij} J_{ij} + \ldots,
\end{aligned} \tag{3.32}$$

where $J_{ij} = \int_{t_k}^{t_{k+1}}\int_{t_k}^{t} dW^i(\tau) \circ dW^j(t)$ denotes a double Stratonovich integral, Γ_{ij} and Λ_{ij} have been defined in (3.29), and the forcing terms and the derivatives of the Hamiltonians are again evaluated at (q_k, p_k). Assuming the conditions (3.27) are satisfied, we have that $\bar{\Gamma}_{ij} = \Gamma_{ij}$ and $\bar{\Lambda}_{ij} = \Lambda_{ij}$, but comparing (3.30) and (3.32), we find that in the general case of noncommutative noise not all first order terms agree, and therefore we only have the local error estimates

$$E\big(q_{k+1} - \bar{q}(t_{k+1}; q_k, p_k)\big) = O(\Delta t^{\frac{3}{2}}), \qquad \sqrt{E\big(\|q_{k+1} - \bar{q}(t_{k+1}; q_k, p_k)\|^2\big)} = O(\Delta t),$$
$$E\big(p_{k+1} - \bar{p}(t_{k+1}; q_k, p_k)\big) = O(\Delta t^{\frac{3}{2}}), \qquad \sqrt{E\big(\|p_{k+1} - \bar{p}(t_{k+1}; q_k, p_k)\|^2\big)} = O(\Delta t). \tag{3.33}$$

Theorem 1.1 from [65] then implies that the method (3.15) has mean-square order 0.5. However, if the noise is commutative, then using the property $J_{ij} + J_{ji} = \Delta W^i \Delta W^j$ (see [44], [65]), one can easily show

$$\sum_{i=1}^{m} \sum_{\substack{j=1 \\ j \neq i}}^{m} \Gamma_{ij} J_{ij} = \frac{1}{2} \sum_{i=1}^{m} \sum_{\substack{j=1 \\ j \neq i}}^{m} \Gamma_{ij} \Delta W^i \Delta W^j, \qquad \sum_{i=1}^{m} \sum_{\substack{j=1 \\ j \neq i}}^{m} \Lambda_{ij} J_{ij} = \frac{1}{2} \sum_{i=1}^{m} \sum_{\substack{j=1 \\ j \neq i}}^{m} \Lambda_{ij} \Delta W^i \Delta W^j. \tag{3.34}$$

In that case all first-order terms in the expansions (3.30) and (3.32) agree, and we have the local error estimates

$$E\big(q_{k+1} - \bar{q}(t_{k+1}; q_k, p_k)\big) = O(\Delta t^2), \qquad \sqrt{E\big(\|q_{k+1} - \bar{q}(t_{k+1}; q_k, p_k)\|^2\big)} = O(\Delta t^{\frac{3}{2}}),$$
$$E\big(p_{k+1} - \bar{p}(t_{k+1}; q_k, p_k)\big) = O(\Delta t^2), \qquad \sqrt{E\big(\|p_{k+1} - \bar{p}(t_{k+1}; q_k, p_k)\|^2\big)} = O(\Delta t^{\frac{3}{2}}). \tag{3.35}$$

Theorem 1.1 from [65] then implies that the method (3.15) has mean-square order 1.0.

□

In the case of a one-dimensional noise the commutation condition (3.28) is trivially satisfied, therefore we have the following corollary.

Corollary 3.6. *Under the assumptions of Theorem 3.5, the method* (3.15) *is convergent with mean-square order 1.0 for systems driven by a one-dimensional noise.*

3.3.3 Examples

In the construction of the integrator (3.15) we may choose the number of stages s. In the deterministic case, the higher the number of stages, the higher order of convergence can be achieved (see [28], [29], [30]). In our case, however, as explained earlier, we cannot in general achieve mean-square order of convergence higher than 1.0, because we only used ΔW^r in (3.15). Since the system (3.15a)-(3.15b) requires solving $2sN$ equations for $2sN$ variables, from the computational point of view it makes sense to only consider methods with low values of s. In this work we focus on the following classical numerical integration formulas (one can easily verify that the conditions (3.16) and (3.27) are satisfied for the discussed methods).

1. *Stochastic midpoint method*
   Using the midpoint rule we obtain a one-stage non-partitioned Runge-Kutta method represented by the tableau

$$\begin{array}{c|c|c|c|c|c|c} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \hline 1 & 1 & 1 & 1 & 1 & 1 \end{array} \tag{3.36}$$

Noting that $Q_1 = (q_k + q_{k+1})/2$ and $P_1 = (p_k + p_{k+1})/2$, this method can be written as

$$\begin{aligned}
q_{k+1} &= q_k + \frac{\partial H}{\partial p}\left(\frac{q_k+q_{k+1}}{2}, \frac{p_k+p_{k+1}}{2}\right)\Delta t + \sum_{i=1}^{m} \frac{\partial h_i}{\partial p}\left(\frac{q_k+q_{k+1}}{2}, \frac{p_k+p_{k+1}}{2}\right)\Delta W^i, \\
p_{k+1} &= p_k + \left[-\frac{\partial H}{\partial q}\left(\frac{q_k+q_{k+1}}{2}, \frac{p_k+p_{k+1}}{2}\right) + F\left(\frac{q_k+q_{k+1}}{2}, \frac{p_k+p_{k+1}}{2}\right)\right]\Delta t \\
&\quad + \sum_{i=1}^{m}\left[-\frac{\partial h_i}{\partial q}\left(\frac{q_k+q_{k+1}}{2}, \frac{p_k+p_{k+1}}{2}\right) + f_i\left(\frac{q_k+q_{k+1}}{2}, \frac{p_k+p_{k+1}}{2}\right)\right]\Delta W^i.
\end{aligned} \tag{3.37}$$

It is an implicit method and in general one has to solve $2N$ equations for $2N$ unknowns. However, if the Hamiltonians are separable, that is, $H(q,p) = T_0(p) + U_0(q)$ and $h_i(q,p) = T_i(p) + U_i(q)$, then q_{k+1} from the first equation can be substituted into the second one. In that case only N nonlinear equations have to be solved for p_{k+1}.

2. *Stochastic Störmer-Verlet method*
A generalization of the classical Störmer-Verlet method can be obtained by choosing the tableau

$$\begin{array}{c|cc|cc|cc|cc|cc}
0 & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\
\frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & 0 & \frac{1}{2} & 0 & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\
\hline
\frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2}
\end{array} \tag{3.38}$$

Noting that $Q_1 = q_k$, $Q_2 = q_{k+1}$, and $P_1 = P_2$, this method can be more efficiently written as

$$\begin{aligned}
P_1 &= p_k + \frac{1}{2}\left[-\frac{\partial H}{\partial q}\big(q_k, P_1\big) + F\big(q_k, P_1\big)\right]\Delta t + \frac{1}{2}\sum_{i=1}^{m}\left[-\frac{\partial h_i}{\partial q}\big(q_k, P_1\big) + f_i\big(q_k, P_1\big)\right]\Delta W^i, \\
q_{k+1} &= q_k + \frac{1}{2}\frac{\partial H}{\partial p}\big(q_k, P_1\big)\Delta t + \frac{1}{2}\frac{\partial H}{\partial p}\big(q_{k+1}, P_1\big)\Delta t + \frac{1}{2}\sum_{i=1}^{m}\frac{\partial h_i}{\partial p}\big(q_k, P_1\big)\Delta W^i + \frac{1}{2}\sum_{i=1}^{m}\frac{\partial h_i}{\partial p}\big(q_{k+1}, P_1\big)\Delta W^i, \\
p_{k+1} &= P_1 + \frac{1}{2}\left[-\frac{\partial H}{\partial q}\big(q_{k+1}, P_1\big) + F\big(q_{k+1}, P_1\big)\right]\Delta t + \frac{1}{2}\sum_{i=1}^{m}\left[-\frac{\partial h_i}{\partial q}\big(q_{k+1}, P_1\big) + f_i\big(q_{k+1}, P_1\big)\right]\Delta W^i.
\end{aligned} \tag{3.39}$$

This method was considered in [57] in the context of symplectic integrators for stochastic Hamiltonian systems without forcing; see also [35]. It is particularly efficient, because the first equation can be solved separately from the second one, and the last equation is an explicit update. Moreover, if the Hamiltonians are separable, the second equation becomes explicit. If in addition the forcing terms F and f_i have special forms, then further improvements in efficiency are possible. For instance, if the forcing terms depend linearly on p, as is often the case in practical applications, then the first equation is a linear equation for P_1, and can be solved using linear solvers. In case the forcing terms are independent of p altogether, then the whole method becomes fully explicit.

3. *2-stage stochastic DIRK method*
In order to reduce the computational cost of solving nonlinear equations, diagonally implicit

Runge-Kutta (DIRK) methods use lower-triangular tableaus (see [28], [29], [30]). One can easily verify that the most general family of 2-stage stochastic DIRK methods that satisfy the conditions (3.16) and (3.27) has the tableau of the form

$$\begin{array}{c|cc|cc|cc|cc|cc|cc} \frac{\lambda}{2} & 0 & \frac{\lambda}{2} & 0 & \frac{\lambda}{2} & 0 & \frac{\lambda}{2} & 0 & \frac{\lambda}{2} & 0 & \frac{\lambda}{2} & 0 \\ \lambda & \frac{1-\lambda}{2} & \lambda & \frac{1-\lambda}{2} & \lambda & \frac{1-\lambda}{2} & \lambda & \frac{1-\lambda}{2} & \lambda & \frac{1-\lambda}{2} & \lambda & \frac{1-\lambda}{2} \\ \hline \lambda & 1-\lambda & \lambda & 1-\lambda & \lambda & 1-\lambda & \lambda & 1-\lambda & \lambda & 1-\lambda & \lambda & 1-\lambda \end{array} \tag{3.40}$$

where $\lambda \in \mathbb{R}$ is an arbitrary parameter. One can check that for $\lambda = 0$ and $\lambda = 1$, this method reduces to the stochastic midpoint method (3.37). For other choices of λ, one needs to solve equations (3.15a) and (3.15b), first for $i = 1$ ($2N$ equations) in order to calculate the internal stages Q_1 and P_1 ($2N$ variables), and then for $i = 2$ ($2N$ equations) to find the internal stages Q_2 and P_2 ($2N$ variables). If the Hamiltonians are separable, then equations (3.15a) can be substituted into equations (3.15b), and the problem is reduced to solving two systems of N equations each.

Note that the methods (3.37), (3.39), and (3.40) are in general implicit. One can use the Implicit Function Theorem to show that for sufficiently small Δt and $|\Delta W^i|$, the relevant nonlinear equations will have a solution. However, since the increments ΔW^i are unbounded, for some values of ΔW^i solutions might not exist. To avoid problems with numerical implementations, if necessary, one can replace ΔW^i in equations (3.37) and (3.39) with the truncated random variables $\overline{\Delta W^i}$ defined as

$$\overline{\Delta W^i} = \begin{cases} A, & \text{if } \Delta W^i > A, \\ \Delta W^i, & \text{if } |\Delta W^i| \le A, \\ -A, & \text{if } \Delta W^i < -A, \end{cases} \tag{3.41}$$

where $A > 0$ is suitably chosen for the considered problem. See [20] and [69] for more details regarding schemes with truncated random increments and their convergence.

3.4 Weak Lagrange-d'Alembert partitioned Runge-Kutta methods

3.4.1 Construction

A general class of weak stochastic Runge-Kutta methods for Stratonovich ordinary differential equations was introduced and analyzed in [78] and [79]. These ideas were later used by Wang & Hong & Xu [97] to construct weak symplectic Runge-Kutta methods for stochastic Hamiltonian systems without forcing. Below we combine these ideas and introduce weak Lagrange-d'Alembert Runge-Kutta methods for stochastic forced Hamiltonian systems of the form (1.1).

Definition 3.7. *An s-stage weak Lagrange-d'Alembert partitioned Runge-Kutta method for the system* (1.1) *is given by*

$$Q_i^{(0)} = q_k + \Delta t \sum_{j=1}^{s} a_{ij}^{(0)} \frac{\partial H}{\partial p}(Q_j^{(0)}, P_j^{(0)}) + \sum_{r=1}^{m} \hat{I}_r \sum_{j=1}^{s} b_{ij}^{(0)} \frac{\partial h_r}{\partial p}(Q_j^{(r)}, P_j^{(r)}), \qquad i = 1, \dots, s, \tag{3.42a}$$

$$\begin{aligned} P_i^{(0)} = p_k &+ \Delta t \sum_{j=1}^{s} a_{ij}^{(0)} \Big[- \frac{\partial H}{\partial q}(Q_j^{(0)}, P_j^{(0)}) + F(Q_j^{(0)}, P_j^{(0)}) \Big] \\ &+ \sum_{r=1}^{m} \hat{I}_r \sum_{j=1}^{s} b_{ij}^{(0)} \Big[- \frac{\partial h_r}{\partial q}(Q_j^{(r)}, P_j^{(r)}) + f_r(Q_j^{(r)}, P_j^{(r)}) \Big], \qquad i = 1, \dots, s, \end{aligned} \tag{3.42b}$$

$$\begin{aligned} Q_i^{(l)} = q_k &+ \Delta t \sum_{j=1}^{s} a_{ij}^{(1)} \frac{\partial H}{\partial p}(Q_j^{(0)}, P_j^{(0)}) + \hat{I}_l \sum_{j=1}^{s} b_{ij}^{(1)} \frac{\partial h_l}{\partial p}(Q_j^{(l)}, P_j^{(l)}) \\ &+ \sum_{\substack{r=1 \\ r \neq l}}^{m} \hat{I}_r \sum_{j=1}^{s} b_{ij}^{(3)} \frac{\partial h_r}{\partial p}(Q_j^{(r)}, P_j^{(r)}), \qquad i = 1, \dots, s, \quad l = 1, \dots, m, \end{aligned} \tag{3.42c}$$

$$\begin{aligned} P_i^{(l)} = p_k &+ \Delta t \sum_{j=1}^{s} a_{ij}^{(1)} \Big[- \frac{\partial H}{\partial q}(Q_j^{(0)}, P_j^{(0)}) + F(Q_j^{(0)}, P_j^{(0)}) \Big] \\ &+ \hat{I}_l \sum_{j=1}^{s} b_{ij}^{(1)} \Big[- \frac{\partial h_l}{\partial q}(Q_j^{(l)}, P_j^{(l)}) + f_l(Q_j^{(l)}, P_j^{(l)}) \Big] \\ &+ \sum_{\substack{r=1 \\ r \neq l}}^{m} \hat{I}_r \sum_{j=1}^{s} b_{ij}^{(3)} \Big[- \frac{\partial h_r}{\partial q}(Q_j^{(r)}, P_j^{(r)}) + f_r(Q_j^{(r)}, P_j^{(r)}) \Big], \qquad i = 1, \dots, s, \quad l = 1, \dots, m, \end{aligned} \tag{3.42d}$$

$$q_{k+1} = q_k + \Delta t \sum_{i=1}^{s} \alpha_i \frac{\partial H}{\partial p}(Q_i^{(0)}, P_i^{(0)}) + \sum_{r=1}^{m} \hat{I}_r \sum_{i=1}^{s} \beta_i \frac{\partial h_r}{\partial p}(Q_i^{(r)}, P_i^{(r)}), \tag{3.42e}$$

$$\begin{aligned} p_{k+1} = p_k &+ \Delta t \sum_{i=1}^{s} \alpha_i \Big[- \frac{\partial H}{\partial q}(Q_i^{(0)}, P_i^{(0)}) + F(Q_i^{(0)}, P_i^{(0)}) \Big] \\ &+ \sum_{r=1}^{m} \hat{I}_r \sum_{i=1}^{s} \beta_i \Big[- \frac{\partial h_r}{\partial q}(Q_i^{(r)}, P_i^{(r)}) + f_r(Q_i^{(r)}, P_i^{(r)}) \Big], \end{aligned} \tag{3.42f}$$

where Δt is the time step, $\hat{I}_1, \dots, \hat{I}_m$ are independent three-point distributed random variables with $P(\hat{I}_r = \pm\sqrt{3\Delta t}) = 1/6$ and $P(\hat{I}_r = 0) = 2/3$, $Q_i^{(0)}$, $Q_i^{(l)}$, $P_i^{(0)}$, and $P_i^{(l)}$ for $i = 1, \dots, s$ and $l = 1, \dots, m$ are the position and momentum internal stages, respectively, and the coefficients of the method $a_{ij}^{(0)}$, $a_{ij}^{(1)}$, $b_{ij}^{(0)}$, $b_{ij}^{(1)}$, $b_{ij}^{(3)}$, α_i, β_i satisfy the conditions

$$\alpha_i a_{ij}^{(0)} + \alpha_j a_{ji}^{(0)} = \alpha_i \alpha_j, \tag{3.43a}$$

$$\alpha_i b_{ij}^{(0)} + \beta_j a_{ji}^{(1)} = \alpha_i \beta_j, \tag{3.43b}$$

$$\beta_i b_{ij}^{(1)} + \beta_j b_{ji}^{(1)} = \beta_i \beta_j, \tag{3.43c}$$

$$\beta_i b_{ij}^{(3)} + \beta_j b_{ji}^{(3)} = \beta_i \beta_j, \tag{3.43d}$$

for $i, j = 1, \dots, s$.

The Runge-Kutta method (3.42) can be represented by the tableau

$$\begin{array}{c|c|c} & a^{(0)} & b^{(0)} & \\ \hline & a^{(1)} & b^{(1)} & b^{(3)} \\ \hline & \alpha^T & \beta^T & \end{array} \tag{3.44}$$

where $a^{(0)} = (a^{(0)}_{ij})_{i,j=1\dots s}$, $\alpha = (\alpha_i)_{i=1\dots s}$, etc. The set of equations (3.42) forms a one-step numerical scheme. Knowing q_k and p_k at time t_k, one can solve Equations (3.42a)-(3.42d) for the internal stages $Q_i^{(0)}$, $Q_i^{(l)}$, $P_i^{(0)}$ and $P_i^{(l)}$, and then use (3.42e)-(3.42f) to determine q_{k+1} and p_{k+1} at time t_{k+1}. Depending on the choice of the coefficients, the method (3.42) is in general implicit. However, since the random variables $\hat{I}_l$ are bounded, one can show that for sufficiently small Δt, the relevant nonlinear equations will have a solution. Below we prove that the Runge-Kutta method (3.42) with the conditions (3.43) is indeed a stochastic Lagrange-d'Alembert method of the form (3.1).

Theorem 3.8. *The s-stage weak Runge-Kutta method* (3.42) *with the conditions* (3.43) *is a stochastic Lagrange-d'Alembert variational integrator of the form* (3.1) *with the discrete Hamiltonian*

$$\begin{aligned} H_d^+(q_k, p_{k+1}) = p_{k+1}q_{k+1} - \Delta t \sum_{i=1}^{s} \alpha_i \left(P_i^{(0)} \frac{\partial H}{\partial p}(Q_i^{(0)}, P_i^{(0)}) - H(Q_i^{(0)}, P_i^{(0)}) \right) \\ - \sum_{r=1}^{m} \hat{I}_r \sum_{i=1}^{s} \beta_i \left(P_i^{(r)} \frac{\partial h_r}{\partial p}(Q_i^{(r)}, P_i^{(r)}) - h_r(Q_i^{(r)}, P_i^{(r)}) \right), \end{aligned} \tag{3.45}$$

and the discrete forces

$$\begin{aligned} F_d^-(q_k, p_{k+1}) &= \Delta t \sum_{i=1}^{s} \alpha_i \left(\frac{\partial Q_i^{(0)}}{\partial q_k} \right)^T F(Q_i^{(0)}, P_i^{(0)}) + \sum_{r=1}^{m} \hat{I}_r \sum_{i=1}^{s} \beta_i \left(\frac{\partial Q_i^{(r)}}{\partial q_k} \right)^T f_r(Q_i^{(r)}, P_i^{(r)}), \\ F_d^+(q_k, p_{k+1}) &= \Delta t \sum_{i=1}^{s} \alpha_i \left(\frac{\partial Q_i^{(0)}}{\partial p_{k+1}} \right)^T F(Q_i^{(0)}, P_i^{(0)}) + \sum_{r=1}^{m} \hat{I}_r \sum_{i=1}^{s} \beta_i \left(\frac{\partial Q_i^{(r)}}{\partial p_{k+1}} \right)^T f_r(Q_i^{(r)}, P_i^{(r)}), \end{aligned} \tag{3.46}$$

where q_{k+1}, p_k, $Q_i^{(0)}$, $Q_i^{(r)}$, $P_i^{(0)}$, and $P_i^{(r)}$, satisfy the system of equations (3.42) *and are understood as functions of q_k and p_{k+1}.*

Proof. The proof is analogous to the proof of Theorem 3.4. □

Remark. For stochastic Hamiltonian systems without forcing, i.e. $F \equiv 0$, $f_r \equiv 0$, the method (3.42) reduces to the weak symplectic Runge-Kutta methods introduced in [97]. Therefore, in that case Theorem 3.8 also provides a type-II generating function for such methods, and consequently an alternative proof of their symplecticity.

3.4.2 Convergence

Rather than precisely approximating each sample path, weak convergence concentrates on approximating the probability distribution and functionals of the exact solution (see [44], [65]). Let $\bar{z}(t) = (\bar{q}(t), \bar{p}(t))$ be the exact solution to (1.1) with the initial conditions q_0 and p_0, and let $z_k = (q_k, p_k)$ denote the numerical solution at time t_k obtained by applying (3.42) iteratively k

times with the constant time step Δt. The numerical solution is said to converge weakly with weak global order r if for each $\varphi \in C_P^{2(r+1)}(T^*Q, \mathbb{R})$ there exists $\delta > 0$ and a constant $C > 0$ such that for all $\Delta t \in (0, \delta)$ we have

$$\left\| E\big(\varphi(z_K)\big) - E\big(\varphi(\bar{z}(T))\big) \right\| \leq C\Delta t^r, \tag{3.47}$$

where $T = K\Delta t$, and $C_P^\alpha(T^*Q, \mathbb{R})$ denotes the space of all $\varphi \in C^\alpha(T^*Q, \mathbb{R})$ with polynomial growth, i.e., there exists a constant $A > 0$ and $\gamma \in \mathbb{N}$ such that $|\partial_z^\beta \varphi(z)| \leq A(1 + \|z\|^{2\gamma})$ for all $z \in T^*Q$ and any partial derivative of order $\beta \leq \alpha$. Weak convergence of the Runge-Kutta methods of type (3.42) has been analyzed, and the relevant order conditions for the coefficients have been derived in [79].

3.4.3 Examples

In [97] a number of weak symplectic Runge-Kutta methods for stochastic Hamiltonian systems without forcing have been proposed. Since the symplecticity conditions derived in [97] are equivalent to the conditions (3.43), these methods become Lagrange-d'Alembert integrators when applied to systems with forcing. In this work, we particularly focus on two methods, namely $SRKw1$ and $SRKw2$, as dubbed in [97].

1. *SRKw1*
   The family of 1-stage $SRKw1$ methods is defined by the tableau

$$\begin{array}{c|c|c} \frac{1}{2} & \lambda & \\ \hline 1-\lambda & \frac{1}{2} & \frac{1}{2} \\ \hline 1 & 1 & \end{array} \tag{3.48}$$

   where $\lambda \in \mathbb{R}$ is an arbitrary parameter. This method is weakly convergent with order 1.0 (see [79], [97]). Since $b^{(1)} = b^{(3)}$, equations (3.42c) and (3.42d) imply that $Q_1^{(1)} = \ldots = Q_1^{(m)}$ and $P_1^{(1)} = \ldots = P_1^{(m)}$. Therefore, in general one has to solve the system (3.42a)-(3.42d) for the $4N$ variables $Q_1^{(0)}$, $P_1^{(0)}$, $Q_1^{(1)}$, and $P_1^{(1)}$. However, for several choices of the parameter λ the computational cost can be reduced. If $\lambda = 0$, then one can first solve the $2N$ equations (3.42a)-(3.42b) for the $2N$ variables $Q_1^{(0)}$, $P_1^{(0)}$, and then the $2N$ equations (3.42c)-(3.42d) for the remaining $2N$ variables $Q_1^{(1)}$, $P_1^{(1)}$. Moreover, if the Hamiltonians are separable, that is, $H(q,p) = T_0(p) + U_0(q)$ and $h_i(q,p) = T_i(p) + U_i(q)$, then equation (3.42a) can be substituted into equation (3.42b), and equation (3.42c) can be substitted into equation (3.42d), thus reducing the complexity to solving two systems of N equations each. A similar situation occurs for $\lambda = 1$. For $\lambda = \frac{1}{2}$ we further have $Q_1^{(0)} = Q_1^{(1)} = (q_k + q_{k+1})/2$ and $P_1^{(0)} = P_1^{(1)} = (p_k + p_{k+1})/2$, and the $SRKw1$ method takes the form of the stochastic midpoint method (3.37) with ΔW^i replaced by $\hat{I}_i$.

2. *SRKw2*
   For systems driven by a single noise ($m = 1$) we can consider methods with $b^{(3)} \equiv 0$. The family of 4-stage $SRKw2$ methods is defined by the tableau

$$
\begin{array}{c|cccc|cccc}
\frac{1}{8} & 0 & 0 & 0 & \frac{5}{6}-\frac{\sqrt{3}}{3} & -\frac{1}{2} & 0 & 0 \\
\frac{1}{4} & \frac{1}{8} & 0 & 0 & -\frac{1}{6}+\frac{\sqrt{3}}{3} & \frac{1}{2} & 0 & 0 \\
\frac{1}{4} & \frac{1}{4} & \frac{1}{8} & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 \\
\frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{8} & -\frac{1}{6} & \frac{1}{2} & 0 & 0 \\
\hline
-\frac{1}{6}+\frac{\sqrt{3}}{6} & \frac{1}{3}-\frac{\sqrt{3}}{6} & 0 & \frac{1}{3} & \frac{1}{4} & \frac{1}{4}-\frac{\sqrt{3}}{6} & 0 & 0 \\
\frac{1}{2} & 0 & 0 & 0 & \frac{1}{4}+\frac{\sqrt{3}}{6} & \frac{1}{4} & 0 & 0 \\
0 & 0 & 0 & 0 & \lambda_1 & \lambda_2 & 0 & \lambda_3 \\
0 & 0 & 0 & 0 & 0 & -\frac{1}{2} & 0 & 0 \\
\hline
\frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{2} & \frac{1}{2} & 0 & 0
\end{array}
\tag{3.49}
$$

where $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}$ are arbitrary parameters. This method is weakly convergent with order 2.0 (see [79], [97]). Note that $\beta_3 = \beta_4 = 0$, so the values of the internal stages $Q_3^{(1)}$, $Q_4^{(1)}$, $P_3^{(1)}$, and $P_4^{(1)}$ are not needed in (3.42e) and (3.42f) to calculate q_{k+1} and p_{k+1}, respectively. Moreover, equations (3.42c) and (3.42d) for $i = 3, 4$ are explicit updates, therefore there is no need to solve for or calculate the values of these internal stages. In fact, the choice of the parameters λ_1, λ_2, and λ_3 has no effect on the values of q_{k+1} and p_{k+1}, therefore we can set them to zero for convenience. Consequently, the system of equations (3.42a) and (3.42b) for $i = 1, 2, 3, 4$, and equations (3.42c) and (3.42d) for $i = 1, 2$ ($12N$ equations) has to be solved for the internal stages $Q_1^{(0)}, \ldots, Q_4^{(0)}$, $P_1^{(0)}, \ldots, P_4^{(0)}$, $Q_1^{(1)}$, $Q_2^{(1)}$, $P_1^{(1)}$, and $P_2^{(1)}$ ($12N$ variables). If the Hamiltonians are separable, then equations (3.42a) and (3.42c) can be substituted into equations (3.42b) and (3.42d), and the resulting system of $6N$ equations can be solved for $P_1^{(0)}, \ldots, P_4^{(0)}$, $P_1^{(1)}$, and $P_2^{(1)}$ ($6N$ variables).

3.5 Quasi-symplecticity

The idea of quasi-symplectic integrators has been proposed in [70] as an attempt to construct numerical methods that at least to some extent emulate the special time evolution of the symplectic and volume forms, as pointed out in Theorem 2.4 and Theorem 2.5, respectively. The authors considered a special form of the stochastic forced Hamiltonian system, namely

$$
\begin{aligned}
&H(q,p) = \frac{1}{2} p^T M^{-1} p + U(q), && F(q,p) = -\Gamma p, \\
&h_i(q,p) = -\sigma_i^T q, && f_i(q,p) = 0, \qquad \text{for } i = 1, \ldots, m,
\end{aligned}
\tag{3.50}
$$

where M is an $N \times N$ constant positive definite matrix, Γ is an $N \times N$ constant matrix, and σ_i are constant vectors. The authors call a numerical integrator $F^+_{t_{k+1},t_k} : (q_k, p_k) \longrightarrow (q_{k+1}, p_{k+1})$ quasi-symplectic if it satisfies the following two conditions when applied to the system (3.50):

(QS1) it degenerates to a symplectic method when the forcing term vanishes, i.e., $\Gamma = 0$

(QS2) the Jacobian

$$
J \equiv \det DF^+_{t_{k+1},t_k} = \frac{D(q_{k+1}, p_{k+1})}{D(q_k, p_k)}
\tag{3.51}
$$

does not depend on q_k and p_k.

The condition (QS2) is natural, since the exact Jacobian (2.43) does not depend on the phase space variables. Several quasi-symplectic numerical methods have been proposed and tested in [70]; see also [66]. Below we demonstrate that the idea of quasi-symplecticity can be extended to more general systems than (3.50).

The methods presented in Section 3.3.3 and Section 3.4.3 preserve the underlying variational structure of the general system (1.1), as has been shown in Theorem 3.1. These methods also naturally reduce to symplectic methods, when the forcing terms F and f_i vanish (see [35], [56], [57], [69], [97]). Below we show that the Störmer-Verlet method satisfies the condition (QS2) for a much broader class of systems than (3.50).

Theorem 3.9. *Suppose that $H(q,p)$, $F(q,p)$, and $h_i(q,p)$, $f_i(q,p)$ for $i = 1,\dots,m$ satisfy conditions (H1)-(H3). If the Hamiltonians are separable, that is,*

$$H(q,p) = T_0(p) + U_0(q), \qquad h_i(q,p) = T_i(p) + U_i(q), \qquad i = 1,\dots,m, \tag{3.52}$$

and the forcing terms have the form

$$F(q,p) = -\Gamma_0 p, \qquad f_i(q,p) = -\Gamma_i p, \qquad i = 1,\dots,m, \tag{3.53}$$

for constant $N \times N$ matrices Γ_i, then the Jacobian J of the discrete flow $F^+_{t_{k+1},t_k} : (q_k,p_k) \longrightarrow (q_{k+1},p_{k+1})$ defined by the Störmer-Verlet method (3.39) does not depend on q_k and p_k, and is almost surely equal to

$$J = \det\left(I + \gamma\Big(I - \frac{1}{2}\gamma\Big)^{-1}\right), \tag{3.54}$$

where I is the $N \times N$ identity matrix, $\gamma = \Delta t\Gamma_0 + \sum_{i=1}^m \Delta W^i \Gamma_i$, and we assume that the matrix $I - \frac{1}{2}\gamma$ is almost surely invertible.

Proof. With the separable Hamiltonians (3.52) and the linear forcing terms (3.53), the first equation in (3.39) is linear, and P_1 can be expressed as

$$P_1 = \Big(I - \frac{1}{2}\gamma\Big)^{-1}\left(p_k - \frac{1}{2}\Delta t\frac{\partial U_0}{\partial q}(q_k) - \frac{1}{2}\sum_{i=1}^m \Delta W^i \frac{\partial U_i}{\partial q}(q_k)\right). \tag{3.55}$$

We then plug in P_1 into the second and third equations in (3.39) to obtain expressions for q_{k+1} and p_{k+1} as functions of q_k and p_k. Let us introduce the notation

$$\begin{aligned} \eta &= I - \frac{1}{2}\gamma, & A &= \Delta t\frac{\partial^2 T_0}{\partial p^2}(P_1) + \sum_{i=1}^m \Delta W^i \frac{\partial^2 T_i}{\partial p^2}(P_1), \\ B &= \Delta t\frac{\partial^2 U_0}{\partial q^2}(q_k) + \sum_{i=1}^m \Delta W^i \frac{\partial^2 U_i}{\partial q^2}(q_k), & C &= \Delta t\frac{\partial^2 U_0}{\partial q^2}(q_{k+1}) + \sum_{i=1}^m \Delta W^i \frac{\partial^2 U_i}{\partial q^2}(q_{k+1}). \end{aligned} \tag{3.56}$$

Using this notation, the Jacobian J of the mapping $(q_k,p_k) \longrightarrow (q_{k+1},p_{k+1})$ can be expressed as

$$J = \begin{vmatrix} \frac{\partial q_{k+1}}{\partial q_k} & \frac{\partial q_{k+1}}{\partial p_k} \\ \frac{\partial p_{k+1}}{\partial q_k} & \frac{\partial p_{k+1}}{\partial p_k} \end{vmatrix} = \begin{vmatrix} I - \frac{1}{2}A\eta^{-1}B & A\eta^{-1} \\ -\frac{1}{2}(I+\gamma\eta^{-1})B - \frac{1}{2}C + \frac{1}{4}CA\eta^{-1}B & I - \frac{1}{2}CA\eta^{-1} + \gamma\eta^{-1} \end{vmatrix}. \tag{3.57}$$

Let us transform this determinant into a block upper triangular form by performing basic linear manipulations on its columns and rows. First, multiply the upper and lower right blocks by $\frac{1}{2}B$ on the right, and add the results to the upper and lower left blocks, respectively. Then, multiply the upper left and right blocks by $\frac{1}{2}C$ on the left, and add the results to the lower left and right blocks, thus obtaining a block upper triangular form. Writing out these steps explicitly, we have

$$J = \begin{vmatrix} I & A\eta^{-1} \\ -\frac{1}{2}C & I - \frac{1}{2}CA\eta^{-1} + \gamma\eta^{-1} \end{vmatrix} = \begin{vmatrix} I & A\eta^{-1} \\ 0 & I + \gamma\eta^{-1} \end{vmatrix} = \det(I + \gamma\eta^{-1}), \tag{3.58}$$

which completes the proof. □

Remark. In case the matrix $\eta = I - \frac{1}{2}\gamma$ is not almost surely invertible, one can replace ΔW^i with the suitably chosen truncated increments (3.41).

4 Numerical experiments

In this section we present the results of our numerical experiments. We have tested the performance of the stochastic Lagrange-d'Alembert integrators presented in Section 3, namely the midpoint method (3.37), the Störmer-Verlet method (3.39), the DIRK method (3.40) with $\lambda = 1/2$, the $SRKw1$ method (3.48) with $\lambda = 0$, and the $SRKw2$ method (3.49), and compared it to the performance of some popular general purpose non-geometric explicit stochastic integrators, namely the mean-square Heun method ([21], [44]), the mean-square $R2$ and $E1$ methods (see [16], [17], [18], [21]), and the weak $RS1$ and $RS2$ methods ([79]). The Lagrange-d'Alembert integrators have demonstrated superior behavior in long-time simulations in all of the examples described below. In the case of the midpoint, Störmer-Verlet, and DIRK methods, we used unbounded increments ΔW^i, but observed no numerical issues. In principle, one should use truncated increments of the form (3.41), but for the chosen parameters in the examples below, the probability of encountering a singularity was negligible.

4.1 Long-time energy behavior

The Kubo oscillator is a stochastic Hamiltonian system with the Hamiltonians given by $H(q,p) = p^2/2 + q^2/2$ and $h(q,p) = \beta(p^2/2 + q^2/2)$, where β is the noise intensity (see [69]). The Kubo oscillator is used in the theory of magnetic resonance and laser physics. Here we consider the damped Kubo oscillator with the forcing terms given by $F(q,p) = -\nu p$ and $f(q,p) = -\beta\nu p$, where ν is the damping coefficient. It is straightforward to verify that the exact solution is given by

$$\begin{aligned} \bar{q}(t) &= q_0 e^{-\frac{\nu}{2}(t+\beta W(t))}\cos\omega(t+\beta W(t)) + \frac{1}{\omega}(p_0 + \frac{\nu}{2}q_0)e^{-\frac{\nu}{2}(t+\beta W(t))}\sin\omega(t+\beta W(t)), \\ \bar{p}(t) &= p_0 e^{-\frac{\nu}{2}(t+\beta W(t))}\cos\omega(t+\beta W(t)) - \frac{1}{\omega}(q_0 + \frac{\nu}{2}p_0)e^{-\frac{\nu}{2}(t+\beta W(t))}\sin\omega(t+\beta W(t)), \end{aligned} \tag{4.1}$$

where q_0 and p_0 are the initial conditions, the angular frequency is $\omega = \frac{1}{2}\sqrt{4-\nu^2}$, and we have assumed the underdamped case $0 \le \nu < 2$. Note that (4.1) is the solution of the deterministic damped harmonic oscillator with the time argument shifted by $\beta W(t)$. Given that $W(t) \sim N(0,t)$ is normally distributed, one can explicitly calculate the expected value of the Hamiltonian H as a function of time as

$$E\Big(H\big(\bar{q}(t),\bar{p}(t)\big)\Big) = ae^{-\frac{\nu(2-\beta^2\nu)}{2}t} + e^{-((2-\nu^2)\beta^2+\nu)t}\Big[b\cos\big(2(1-\beta^2\nu)\omega t\big) + c\sin\big(2(1-\beta^2\nu)\omega t\big)\Big], \quad (4.2)$$

where

$$a = \frac{2(p_0^2+q_0^2+\nu p_0 q_0)}{4-\nu^2}, \qquad b = -\frac{\nu^2(p_0^2+q_0^2)+4\nu p_0 q_0}{2(4-\nu^2)}, \qquad c = \frac{\nu(q_0^2-p_0^2)}{2\sqrt{4-\nu^2}}. \quad (4.3)$$

Simulations with the initial conditions $q_0 = 2$, $p_0 = 0$, and the parameters $\beta = 0.5$ and $\nu = 0.001$ were carried out until the time $T = 5000$ (approximately 800 periods of the oscillator in the absence of noise). In each case 50000 sample paths were generated. The numerical value of the mean Hamiltonian $E(H)$ as a function of time is depicted in Figure 4.1 and Figure 4.2 for the mean-square and weak integrators, respectively. We see that the Lagrange-d'Alembert integrators capture the exponential decay of $E(H)$ very accurately even when relatively large time steps Δt are used. The explicit Heun and $R2$ methods fail to reproduce that behavior even for the significantly smaller time step. While the explicit $E1$, $RS1$, and $RS2$ methods capture the qualitative decay of $E(H)$, still much smaller time steps would be needed to reach the level of accuracy of the Lagrange-d'Alembert integrators, thus rendering them inefficient.

4.2 Ergodic limits

In many cases of practical interest the system (1.1) is *ergodic*, which means that

(1) it possesses a unique invariant measure represented by the probability density function $\rho_\infty(\xi,\zeta)$ with $(\xi,\zeta) \in T^*Q$, i.e. a stationary solution of the corresponding Fokker-Planck equation (see [26])

(2) for any function $\varphi : T^*Q \longrightarrow \mathbb{R}$ with polynomial growth at infinity, its ergodic limit, i.e. the expected value with respect to the invariant measure, can be calculated as the limit

$$\varphi^{\text{erg}} \equiv \int\int \varphi(\xi,\zeta)\rho_\infty(\xi,\zeta)\,d\xi d\zeta = \lim_{t\to+\infty} E\Big(\varphi\big(\bar{q}(t),\bar{p}(t)\big)\Big), \quad (4.4)$$

where $(\bar{q}(t),\bar{p}(t))$ is an arbitrary solution of (1.1) with arbitrary initial conditions.

For more information about ergodic systems and ergodic numerical schemes see, e.g., [12], [15], [36], [61], [62], [66], [88]. For many applications, it is interesting to compute the mean of a given function with respect to the invariant law of the diffusion, but the explicit form of the invariant measure is often not known. If the considered system is ergodic, then the ergodic limit can be approximated as

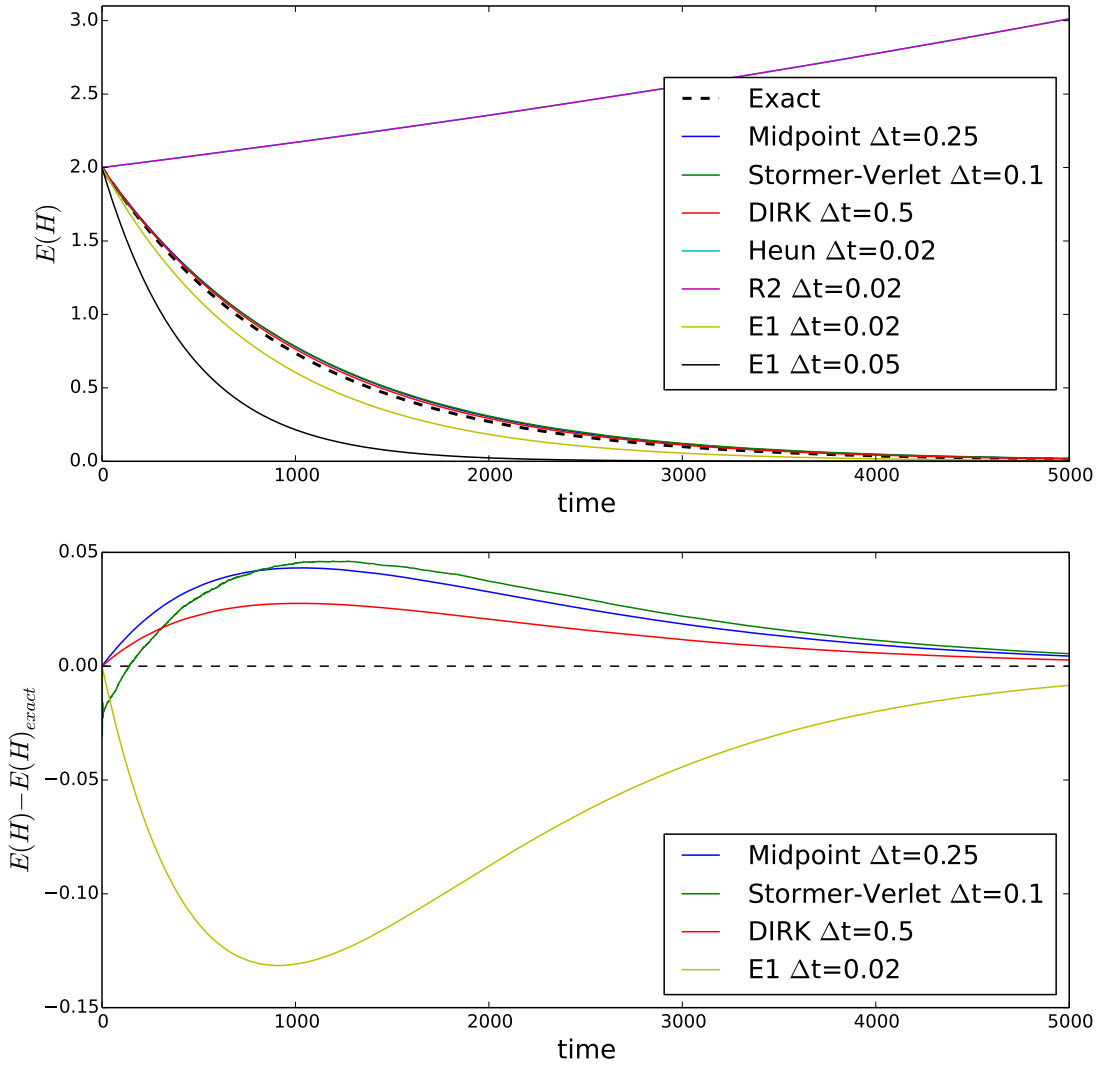

Figure 4.1: *Top:* The numerical value of the mean Hamiltonian $E(H)$ for the simulations of the damped Kubo oscillator with the initial conditions $q_0 = 2$, $p_0 = 0$, and the parameters $\beta = 0.5$ and $\nu = 0.001$ is shown for the solutions computed with the mean-square explicit Heun, $R2$, and $E1$ methods, and the mean-square Lagrange-d'Alembert methods presented in Section 3.3.3. The Lagrange-d'Alembert integrators accurately capture the exponential decay of $E(H)$, whereas the explicit methods either fail to reproduce that behavior or do so inaccurately. Note that the plots for the Heun and $R2$ methods, as well as for the midpoint and Störmer-Verlet methods, overlap very closely. *Bottom:* The difference between the numerical value of the mean Hamiltonian $E(H)$ and the exact value (4.2) is shown for the $E1$ method and the Lagrange-d'Alembert integrators. The stochastic DIRK method proves to be particularly accurate even when the time step $\Delta t = 0.5$ is used.

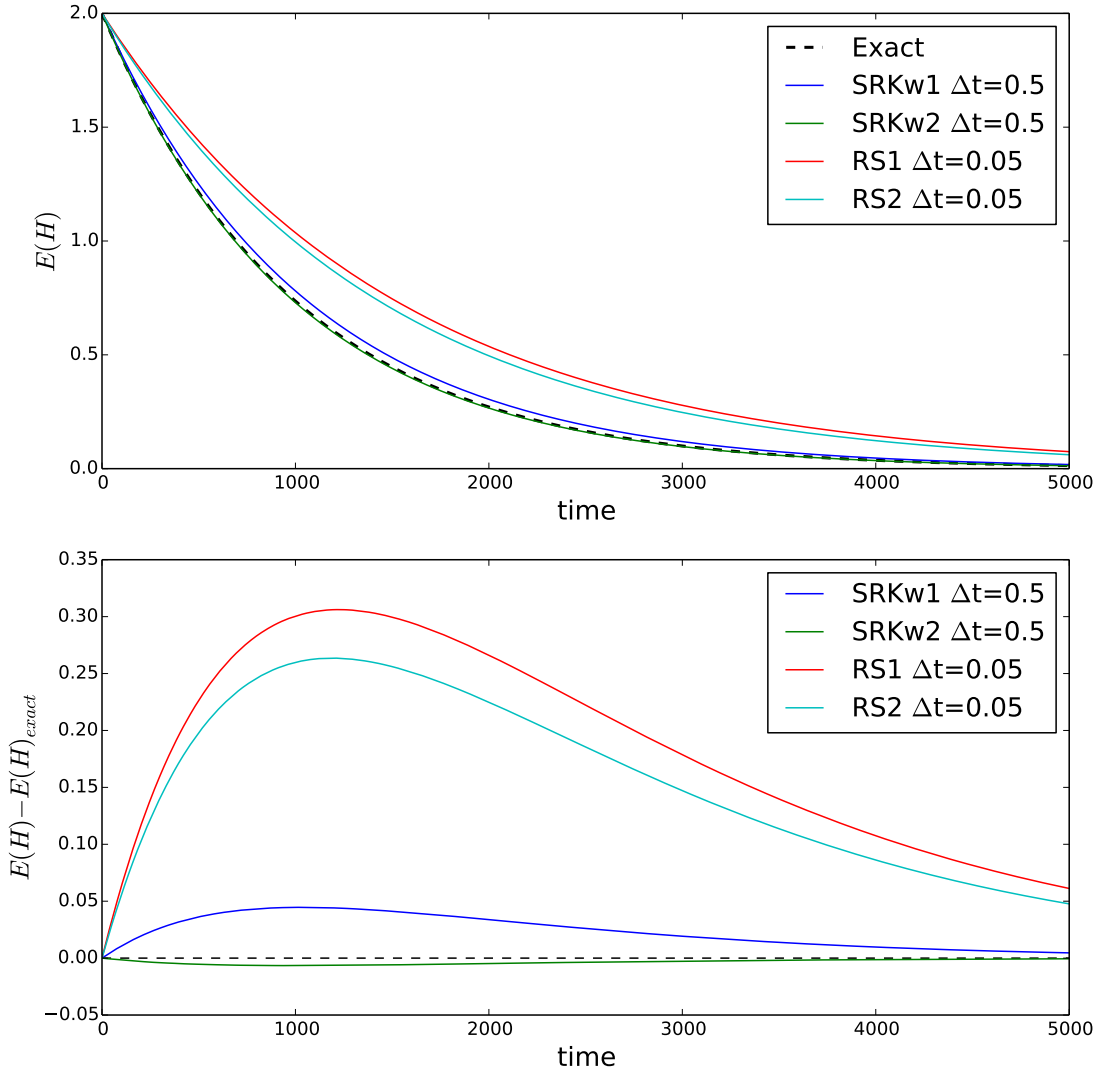

Figure 4.2: *Top:* The numerical value of the mean Hamiltonian $E(H)$ for the simulations of the damped Kubo oscillator with the initial conditions $q_0 = 2$, $p_0 = 0$, and the parameters $\beta = 0.5$ and $\nu = 0.001$ is shown for the solutions computed with the weak explicit $RS1$ and $RS2$ methods, and the weak Lagrange-d'Alembert methods presented in Section 3.4.3. The Lagrange-d'Alembert integrators capture the exponential decay of $E(H)$ much more accurately then the explicit ones, even when much larger time steps are used. *Bottom:* The difference between the numerical value of the mean Hamiltonian $E(H)$ and the exact value (4.2) is shown instead. The $SRKw2$ method proves to be significantly more accurate then the others, even when the time step $\Delta t = 0.5$ is used.

$$\varphi^{\mathrm{erg}} \approx E\Big(\varphi\big(\bar{q}(T),\bar{p}(T)\big)\Big) \tag{4.5}$$

by choosing a sufficiently large time T. One can then use numerical integrators to approximate $\bar{q}(T)$ and $\bar{p}(T)$. However, formula (4.5) requires integration of the system over comparatively long time intervals, which poses a significant computational difficulty. Below we compare the performance of the geometric integrators introduced in Section 3 with the performance of explicit schemes.

Consider Van der Pol's equation with additive noise (see [66]), which is a stochastic forced Hamiltonian system of the form (1.1) with

$$\begin{aligned} H(q,p) &= \frac{1}{2}p^2 + \frac{1}{2}q^2, & F(q,p) &= \nu(1-q^2)p, \\ h(q,p) &= -\sigma q, & f(q,p) &= 0, \end{aligned} \tag{4.6}$$

where $\nu \geq 0$ and $\sigma \geq 0$ are parameters. The explicit form of the invariant measure for this system is unknown, however, it is interesting to compute the ergodic value of the energy. Simulations with the initial conditions $q_0 = 1$, $p_0 = 1$, and the parameters $\sigma = 0.05$ and $\nu = 0.001$ were carried out until the time $T = 5000$. In each case 10^6 sample paths were generated. The numerical value of the mean Hamiltonian $E(H)$ as a function of time is depicted in Figure 4.3 for the DIRK, Heun, and $E1$ methods. As the reference value we take $H^{\mathrm{erg}} = 2.3165$, which was calculated in [66] using a second-order weak quasi-symplectic method at the time $T_{\mathrm{ref}} = 10000$ with the time step $\Delta t = 0.05$ and 4×10^6 sample paths. We see that the DIRK method accurately reproduces the reference value even with the relatively large time step $\Delta t = 0.2$, while the Heun and $E1$ methods require the much smaller time step $\Delta t = 0.02$ to reach that level of accuracy. The situation is similar for the other Lagrange-d'Alembert and explicit integrators. Figure 4.4 depicts the behavior of $E(H)$ near the reference value on the time interval $[4000, 5000]$ for each of the tested integrators.

4.3 Vlasov equation

The following two-dimensional Vlasov Fokker-Planck equation

$$\frac{\partial \rho}{\partial t} + v\frac{\partial \rho}{\partial x} - E(x)\frac{\partial \rho}{\partial v} = \nu\left(\mu\frac{\partial (v\rho)}{\partial v} + \frac{D^2}{2}\frac{\partial^2 \rho}{\partial v^2}\right) \tag{4.7}$$

has been studied in [43] and [85] as a model for collisional kinetic plasmas, where $\rho = \rho(x,v,t)$ denotes the particle distribution function in the position-velocity phase space, $E(x) = -\phi'(x)$ is the external electric field with the electrostatic potential $\phi(x)$, and $\nu > 0$, $\mu > 0$, $D > 0$ are real parameters. A stochastic split particle-in-cell (PIC) method for the numerical simulation of (4.7) has been proposed in [85], whereby the advection part is solved using the standard PIC method, and the diffusion part is modeled by a stochastic differential equation. Below we demonstrate a structure-preserving approach to solving (4.7). When ρ is interpreted as a probability density function, then (4.7) is the Fokker-Planck equation for the two-dimensional stochastic process $(X(t), V(t))$ whose evolution is governed by the stochastic differential equation (see [26], [44])

$$d_t X = V\,dt, \qquad d_t V = \big(-E(X) - \nu\mu V\big)\,dt + \sqrt{\nu}D \circ dW(t), \tag{4.8}$$

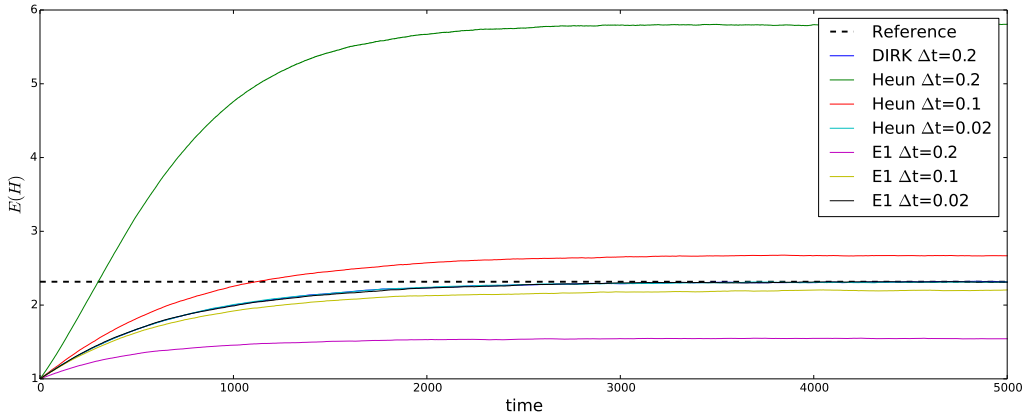

Figure 4.3: The numerical value of the mean Hamiltonian $E(H)$ as a function of time for the simulations of Van der Pol's equation with the initial conditions $q_0 = 1$, $p_0 = 1$, and the parameters $\sigma = 0.05$ and $\nu = 0.001$, is shown for the solutions computed with the DIRK, Heun, and $E1$ methods. The reference value $H^{\text{erg}} = 2.3165$ was calculated in [66]. The DIRK method accurately reproduces the reference value even with the relatively large time step $\Delta t = 0.2$, while the Heun and $E1$ methods require the much smaller time step $\Delta t = 0.02$ to reach that level of accuracy. For the clarity of the plot the other Lagrange-d'Alembert and explicit methods are not depicted, but they demonstrate similar behavior. Note that the plots for the DIRK method, and the Heun and $E1$ methods with $\Delta t = 0.02$ overlap very closely.

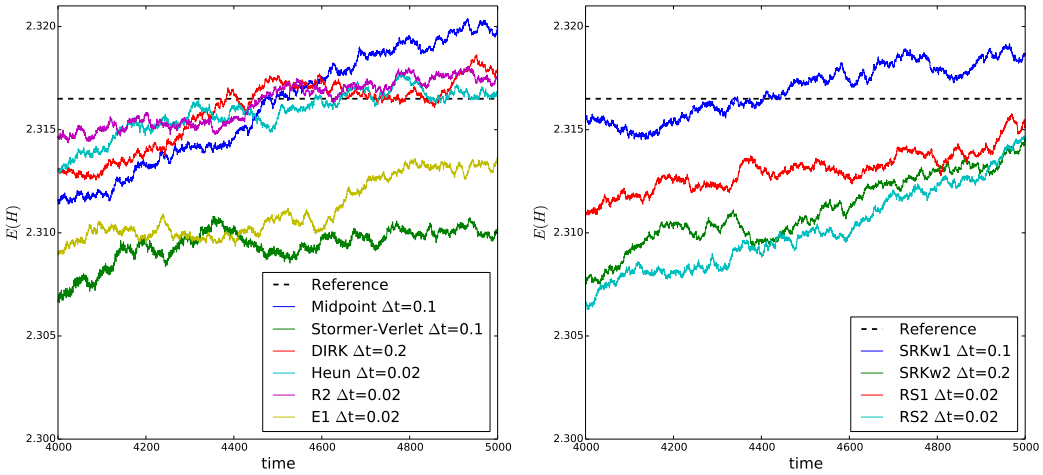

Figure 4.4: The numerical value of the mean Hamiltonian $E(H)$ for the simulations of Van der Pol's equation with the initial conditions $q_0 = 1$, $p_0 = 1$, and the parameters $\sigma = 0.05$ and $\nu = 0.001$, is shown on the time interval $[4000, 5000]$ near the reference value for the solutions computed with the mean-square (*Left*) and weak (*Right*) methods. The reference value $H^{\text{erg}} = 2.3165$ was calculated in [66].

driven by the one-dimensional Wiener process $W(t)$. This equation is a stochastic forced Hamiltonian system (1.1) with

$$H(X,V) = \frac{1}{2}V^2 - \phi(X), \qquad F(X,V) = -\nu\mu V,$$
$$h(X,V) = -\sqrt{\nu}DX, \qquad f(X,V) = 0. \tag{4.9}$$

It can be easily verified that the stationary solution of (4.7) is given by the *Gibbs measure*

$$\rho_\infty(x,v) = \frac{1}{Z}e^{-\frac{2\mu}{D^2}H(x,v)} = \frac{1}{Z}e^{\frac{2\mu}{D^2}\phi(x)}e^{-\frac{\mu}{D^2}v^2}, \tag{4.10}$$

where Z is the normalizing constant such that $\int\int \rho_\infty(x,v)\,dv\,dx = 1$. Let us consider (4.7) on the domain $(x,v) \in [0,1]\times\mathbb{R}$ with periodic boundary conditions in x, and with the electrostatic potential

$$\phi(x) = -\frac{E_0}{4\pi}\sin 4\pi x, \tag{4.11}$$

where $E_0 > 0$ is the maximum magnitude of the electric field $E(x) = -\phi'(x)$. One can check that the system (4.9) with the potential (4.11) is ergodic (see Theorem 3.2 in [61]). As the initial condition, we take the probability distribution of the form

$$\rho(x,v,0) = \rho_X(x)\rho_V(v) = (1+\epsilon\cos 2\pi x)\left(\frac{1}{1+a}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}v^2} + \frac{a}{1+a}\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{1}{2\sigma^2}(v-v_0)^2}\right), \tag{4.12}$$

where $\rho_X(x)$ for $\epsilon > 0$ describes a perturbation of the uniform distribution along the spatial direction x, and $\rho_V(v)$ for $a > 0$ is the so called *bump-on-tail* distribution in the velocity space, where the bump is centered at v_0 with the standard deviation $\sigma > 0$. Simulations with the parameters $\nu = 1$, $\mu = 1$, $D = \sqrt{2}$, $E_0 = 3$, $\epsilon = 0.25$, $a = 0.5$, $v_0 = 4$, and $\sigma = 0.5$ were carried out until the time $T = 1000$. In each case 10^7 sample paths were generated. The initial conditions X_0 and V_0 were randomly drawn from the probability distribution (4.12) using rejection sampling (see Figure 4.5). The exact ergodic value H^{erg} of the Hamiltonian can be calculated using the invariant probability density (4.10) as

$$H^{\text{erg}} = \int_0^1\int_{-\infty}^{\infty} H(x,v)\rho_\infty(x,v)\,dv\,dx \approx 0.471705. \tag{4.13}$$

The numerical value of the mean Hamiltonian $E(H)$ as a function of time is depicted in Figure 4.6 for the DIRK, Heun, and $E1$ methods. We see that the DIRK method accurately reproduces the ergodic limit even with the relatively large time step $\Delta t = 0.15$, while the $E1$ method requires the much smaller time step $\Delta t = 0.02$ to reach a comparable level of accuracy. The Heun method yields a less accurate result even for $\Delta t = 0.02$. The situation is similar for the other Lagrange-d'Alembert and explicit integrators. Figure 4.7 depicts the behavior of $E(H)$ near the exact ergodic limit on the time interval $[500, 1000]$ for each of the tested integrators. The numerical probability density at the final time $T = 1000$ calculated with each of the mean-square and weak methods is depicted in comparison to the exact invariant measure (4.10) in Figure 4.8 and Figure 4.9, respectively.

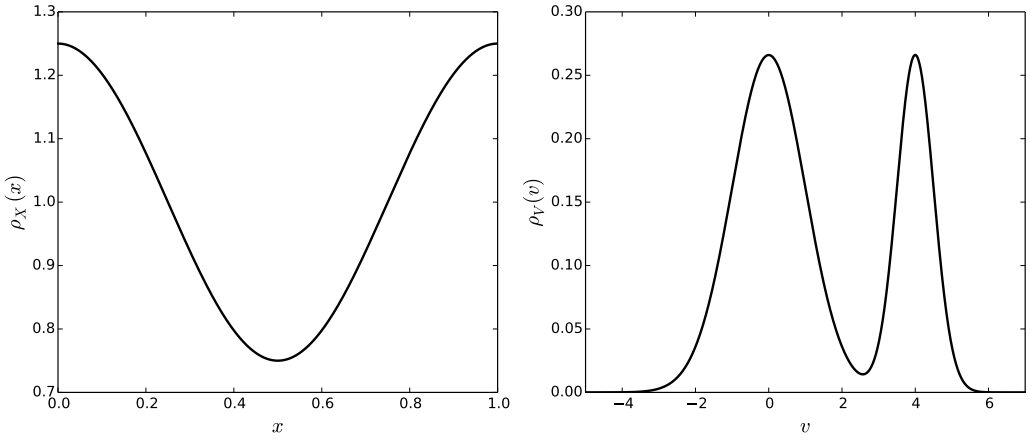

Figure 4.5: The initial probability density (4.12) for the simulations of the Vlasov equation with the parameters $\epsilon = 0.25$, $a = 0.5$, $v_0 = 4$, and $\sigma = 0.5$.

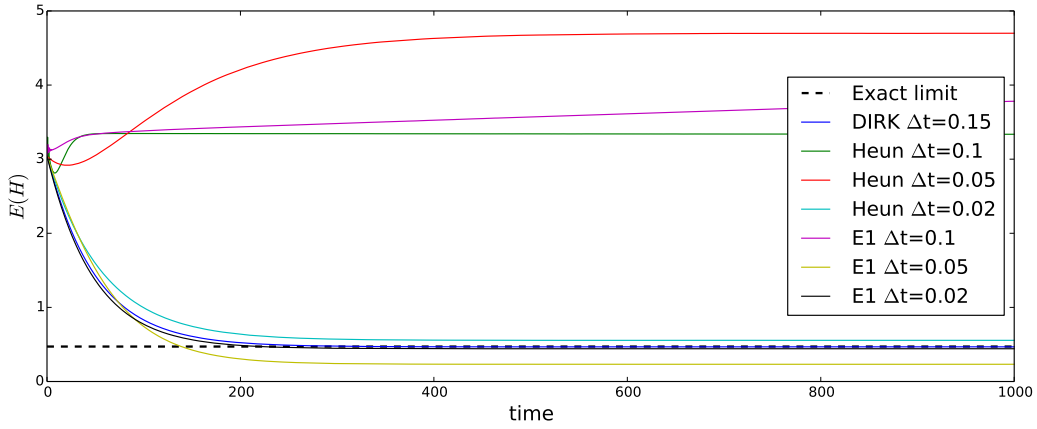

Figure 4.6: The numerical value of the mean Hamiltonian $E(H)$ as a function of time for the simulations of the Vlasov equation with the parameters $\nu = 1$, $\mu = 1$, $D = \sqrt{2}$, $E_0 = 3$, $\epsilon = 0.25$, $a = 0.5$, $v_0 = 4$, and $\sigma = 0.5$, and the initial conditions X_0 and V_0 sampled from the distribution (4.12), is shown for the solutions computed with the DIRK, Heun, and $E1$ methods. The exact ergodic limit $H^{\text{erg}} \approx 0.471705$ was calculated in (4.13). The DIRK method accurately reproduces the ergodic limit even with the relatively large time step $\Delta t = 0.15$, while the $E1$ method requires the much smaller time step $\Delta t = 0.02$ to reach a comparable level of accuracy. The Heun method yields a less accurate result even for $\Delta t = 0.02$. For the clarity of the plot the other Lagrange-d'Alembert and explicit methods are not depicted, but they demonstrate similar behavior. Note that the plots for the DIRK method and the $E1$ method with $\Delta t = 0.02$ overlap very closely.

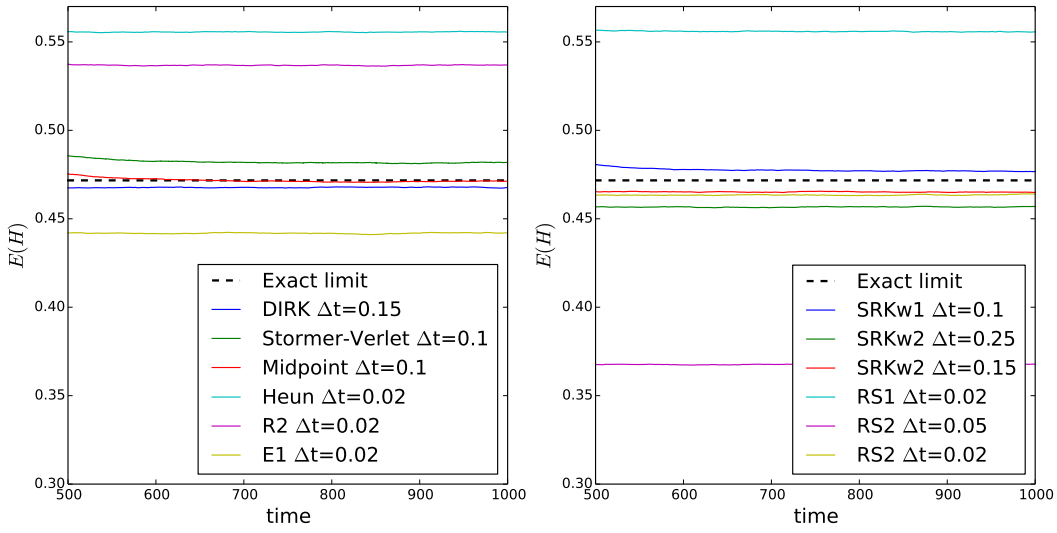

Figure 4.7: The numerical value of the mean Hamiltonian $E(H)$ for the simulations of the Vlasov equation with the parameters $\nu = 1$, $\mu = 1$, $D = \sqrt{2}$, $E_0 = 3$, $\epsilon = 0.25$, $a = 0.5$, $v_0 = 4$, and $\sigma = 0.5$, and the initial conditions X_0 and V_0 sampled from the distribution (4.12), is shown on the time interval $[500, 1000]$ near the exact ergodic limit for the solutions computed with the mean-square (*Left*) and weak (*Right*) methods. The exact ergodic limit $H^{\text{erg}} \approx 0.471705$ was calculated in (4.13).

5 Summary and future work

In this paper we have presented a general framework for constructing a new class of stochastic variational integrators for stochastic forced Hamiltonian systems. We have extended the approach taken in [35] by considering the stochastic Lagrange-d'Alembert principle and constructing the corresponding structure-preserving schemes, which we have dubbed stochastic Lagrange-d'Alembert variational integrators. We have shown that in the presence of a symmetry such integrators satisfy a discrete version of stochastic forced Noether's theorem. We have further considered certain classes of mean-square and weak Runge-Kutta methods previously known in the literature, and determined the conditions under which such methods become Lagrange-d'Alembert integrators. We have finally pointed out several examples of low-stage Runge-Kutta methods of that type, and demonstrated their superior long-time numerical performance via numerical experiments. In particular, as one of the test cases we have considered the Vlasov Fokker-Planck equation and proposed a new geometric approach to the simulation of collisional kinetic plasmas.

Our work can be extended in several ways. The mean-square partitioned Runge-Kutta methods introduced in Section 3.3 only use the increments $\Delta W^r = \int_{t_k}^{t_{k+1}} dW^r(t)$, therefore their mean-square order of convergence cannot exceed 1.0 (see [18], [68], [69]). To obtain mean-square convergence of higher order one can extend the definitions of the discrete Hamiltonian (3.18) and the discrete forces (3.19) to include higher-order multiple Stratonovich integrals, e.g., to achieve convergence of order 1.5 we would need to include terms involving $\Delta Z^r = \int_{t_k}^{t_{k+1}} \int_{t_k}^{t} dW^r(\xi)\, dt$; see [35] for an example how this can be done for unforced Hamiltonian systems. Another aspect worth a more detailed investigation is the issue of ergodicity of the Lagrange-d'Alembert methods. In Section 4.2 and Section 4.3

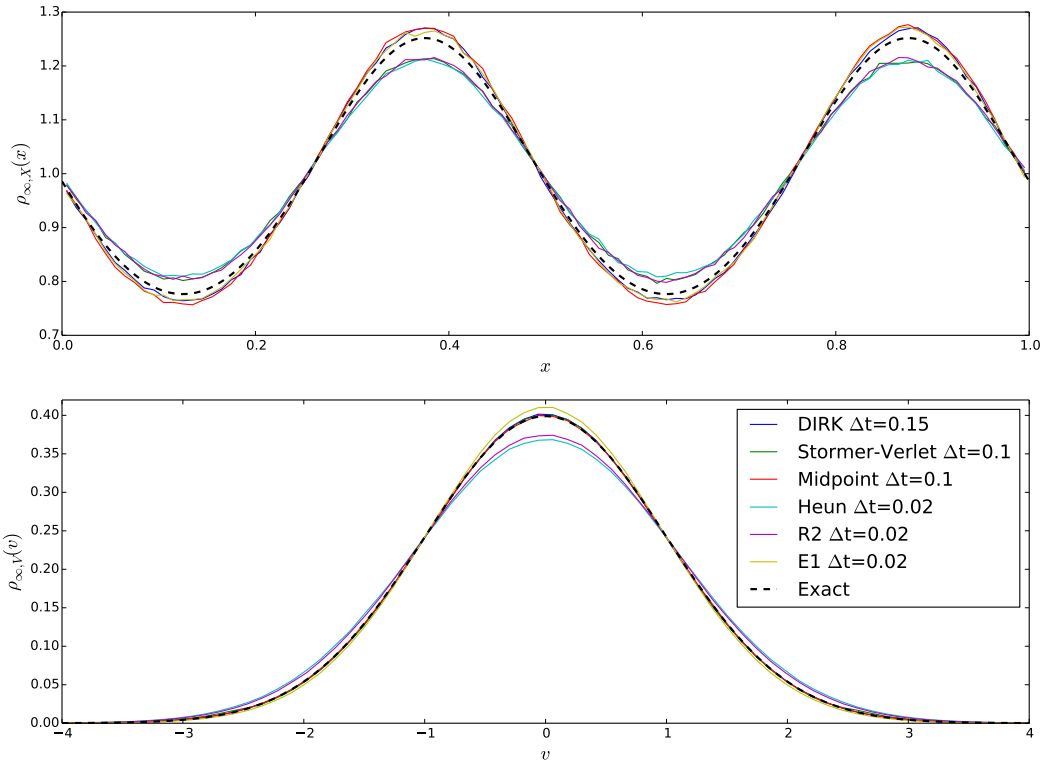

Figure 4.8: The numerical probability density at time $T = 1000$ for the simulations of the Vlasov equation with the parameters $\nu = 1$, $\mu = 1$, $D = \sqrt{2}$, $E_0 = 3$, $\epsilon = 0.25$, $a = 0.5$, $v_0 = 4$, and $\sigma = 0.5$, and the initial conditions X_0 and V_0 sampled from the distribution (4.12), is depicted for each of the mean-square integrators, and compared to the exact invariant measure (4.10).

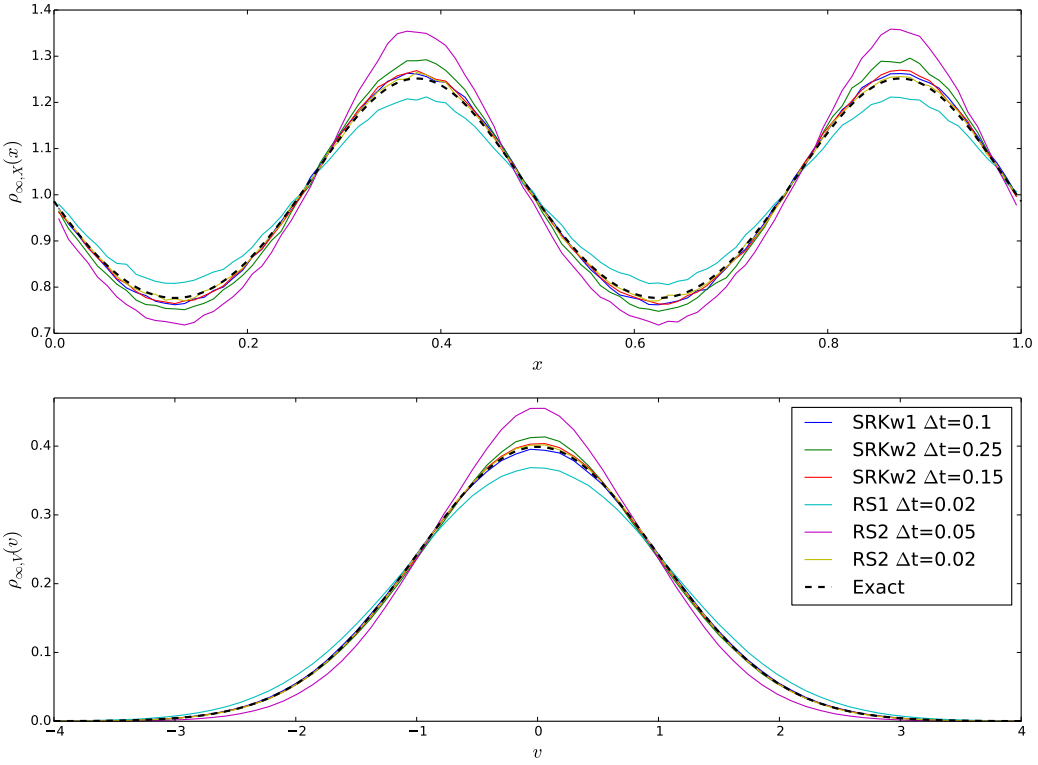

Figure 4.9: The numerical probability density at time $T = 1000$ for the simulations of the Vlasov equation with the parameters $\nu = 1$, $\mu = 1$, $D = \sqrt{2}$, $E_0 = 3$, $\epsilon = 0.25$, $a = 0.5$, $v_0 = 4$, and $\sigma = 0.5$, and the initial conditions X_0 and V_0 sampled from the distribution (4.12), is depicted for each of the weak integrators, and compared to the exact invariant measure (4.10).

we have experimentally demonstrated the usefulness of our integrators in calculating the ergodic limits, but have not formally proved their ergodicity. It would be beneficial to determine under what conditions Lagrange-d'Alembert integrators can be ergodic in the sense discussed in, e.g., [61], [62], or [88], when applied to ergodic Hamiltonian systems. It would also be interesting to extend the idea of Lagrange-d'Alembert integrators to stochastic Hamiltonian systems that are both forced and constrained. Structure-preserving numerical methods for such systems would be of great interest in molecular dynamics (see [13], [23], [92]). Yet another direction of great practical significance would be a further study of the geometric approach to collisional kinetic plasmas presented in Section 4.3 and application of more realistic collision operators that preserve the total energy and momentum, as well as an extension to the self-consistent Maxwell-Vlasov equations (see [47], [48]). Finally, one may extend the idea of variational integration to stochastic multisymplectic partial differential equations such as the stochastic Korteweg-de Vries, Camassa-Holm or Hunter-Saxton equations. Theoretical groundwork for such numerical schemes has been recently presented in [34].

Acknowledgments

We would like to thank Darryl Holm, Katharina Kormann, Omar Maj, Philip Morrison, Cesare Tronci, and Udo von Toussaint for useful comments and references. We are particularly endebted to Eric Sonnendrücker for pointing out the connections between stochastic systems and the Vlasov equation. The study is a contribution to the Reduced Complexity Models grant number ZT-I-0010 funded by the Helmholtz Association of German Research Centers.

References

[1] A. Abdulle, D. Cohen, G. Vilmart, and K. Zygalakis. High weak order methods for stochastic differential equations based on modified equations. *SIAM Journal on Scientific Computing*, 34(3):A1800–A1823, 2012.

[2] A. Abdulle, G. Vilmart, and K. Zygalakis. Long time accuracy of Lie–Trotter splitting methods for Langevin dynamics. *SIAM Journal on Numerical Analysis*, 53(1):1–16, 2015.

[3] R. Abraham, J. Marsden, and T. Ratiu. *Manifolds, Tensor Analysis, and Applications*. Applied Mathematical Sciences. Springer New York, 1993.

[4] S. Anmarkrud and A. Kværnø. Order conditions for stochastic Runge–Kutta methods preserving quadratic invariants of Stratonovich SDEs. *Journal of Computational and Applied Mathematics*, 316:40 – 46, 2017.

[5] C. Anton. Weak backward error analysis for stochastic Hamiltonian systems. *BIT Numerical Mathematics*, 2019.

[6] C. Anton, J. Deng, and Y. S. Wong. Weak symplectic schemes for stochastic Hamiltonian equations. *Electronic Transactions on Numerical Analysis*, 43:1–20, 2014.

[7] C. Anton, Y. S. Wong, and J. Deng. On global error of symplectic schemes for stochastic Hamiltonian systems. *International Journal of Numerical Analysis and Modeling, Series B*, 4(1):80–93, 2013.

[8] C. Anton, Y. S. Wong, and J. Deng. Symplectic numerical schemes for stochastic systems preserving Hamiltonian functions. In I. Dimov, I. Faragó, and L. Vulkov, editors, *Numerical Analysis and Its Applications: 5th International Conference, NAA 2012, Lozenetz, Bulgaria, June 15-20, 2012, Revised Selected Papers*, pages 166–173. Springer Berlin Heidelberg, Berlin, Heidelberg, 2013.

[9] L. Arnold. *Stochastic Differential Equations: Theory and Applications*. Dover Books on Mathematics. Dover Publications, 2013.

[10] D. A. Beard and T. Schlick. Inertial stochastic dynamics. I. Long-time-step methods for Langevin dynamics. *The Journal of Chemical Physics*, 112(17):7313–7322, 2000.

[11] J. Bismut. Mecanique aleatoire. In P. Hennequin, editor, *Ecole d'Eté de Probabilités de Saint-Flour X - 1980*, volume 929 of *Lecture Notes in Mathematics*, pages 1–100. Springer Berlin Heidelberg, 1982.

[12] N. Bou-Rabee and H. Owhadi. Boltzmann-Gibbs preserving Langevin integrators. Unpublished, arXiv:0712.4123v3, 2007.

[13] N. Bou-Rabee and H. Owhadi. Stochastic variational partitioned Runge-Kutta integrators for constrained systems. Unpublished, arXiv:0709.2222, 2007.

[14] N. Bou-Rabee and H. Owhadi. Stochastic variational integrators. *IMA Journal of Numerical Analysis*, 29(2):421–443, 2009.

[15] N. Bou-Rabee and H. Owhadi. Long-run accuracy of variational integrators in the stochastic context. *SIAM J. Numer. Anal.*, 48(1):278–297, 2010.

[16] K. Burrage and P. M. Burrage. High strong order explicit Runge-Kutta methods for stochastic ordinary differential equations. *Appl. Numer. Math.*, 22:81–101, 1996.

[17] K. Burrage and P. M. Burrage. General order conditions for stochastic Runge-Kutta methods for both commuting and non-commuting stochastic ordinary differential equation systems. *Appl. Numer. Math.*, 28:161–177, 1998.

[18] K. Burrage and P. M. Burrage. Order Conditions of Stochastic Runge–Kutta Methods by B-Series. *SIAM Journal on Numerical Analysis*, 38(5):1626–1646, 2000.

[19] K. Burrage and P. M. Burrage. Low rank Runge–Kutta methods, symplecticity and stochastic Hamiltonian problems with additive noise. *Journal of Computational and Applied Mathematics*, 236(16):3920 – 3930, 2012.

[20] K. Burrage and T. Tian. Implicit stochastic Runge–Kutta methods for stochastic differential equations. *BIT Numerical Mathematics*, 44(1):21–39, 2004.

[21] P. Burrage. *Runge-Kutta methods for stochastic differential equations*. PhD thesis, University of Queensland, 1999.

[22] P. M. Burrage and K. Burrage. Structure-preserving Runge-Kutta methods for stochastic Hamiltonian equations with additive noise. *Numerical Algorithms*, 65(3):519–532, 2014.

[23] G. Ciccotti, T. Lelièvre, and E. Vanden-Eijnden. Projection of diffusions on submanifolds: Application to mean force computation. *Communications on Pure and Applied Mathematics*, 61(3):371–408, 2008.

[24] A. Debussche and E. Faou. Weak backward error analysis for SDEs. *SIAM Journal on Numerical Analysis*, 50(3):1735–1752, 2012.

[25] J. Deng, C. Anton, and Y. S. Wong. High-order symplectic schemes for stochastic Hamiltonian systems. *Communications in Computational Physics*, 16:169–200, 2014.

[26] C. Gardiner. *Stochastic Methods: A Handbook for the Natural and Social Sciences*. Springer Series in Synergetics. Springer Berlin Heidelberg, 2009.

[27] Y. N. Gornostyrev, M. I. Katsnelson, A. V. Trefilov, and S. V. Tret'jakov. Stochastic approach to simulation of lattice vibrations in strongly anharmonic crystals: Anomalous frequency dependence of the dynamic structure factor. *Phys. Rev. B*, 54:3286–3294, Aug 1996.

[28] E. Hairer, C. Lubich, and G. Wanner. *Geometric Numerical Integration: Structure-Preserving Algorithms for Ordinary Differential Equations*. Springer Series in Computational Mathematics. Springer, New York, 2002.

[29] E. Hairer, S. Nørsett, and G. Wanner. *Solving Ordinary Differential Equations I: Nonstiff Problems*, volume 8 of *Springer Series in Computational Mathematics*. Springer, 2nd edition, 1993.

[30] E. Hairer and G. Wanner. *Solving Ordinary Differential Equations II: Stiff and Differential-Algebraic Problems*, volume 14 of *Springer Series in Computational Mathematics*. Springer, 2nd edition, 1996.

[31] J. Hall and M. Leok. Spectral variational integrators. *Numer. Math.*, 130(4):681–740, Aug. 2015.

[32] D. D. Holm, T. Schmah, and C. Stoica. *Geometric Mechanics and Symmetry: From Finite to Infinite Dimensions*. Oxford Texts in Applied and Engineering Mathematics. Oxford University Press, Oxford, 2009.

[33] D. D. Holm and T. M. Tyranowski. Variational principles for stochastic soliton dynamics. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 472(2187), 2016.

[34] D. D. Holm and T. M. Tyranowski. New variational and multisymplectic formulations of the Euler–Poincaré equation on the Virasoro–Bott group using the inverse map. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 474(2213), 2018.

[35] D. D. Holm and T. M. Tyranowski. Stochastic discrete Hamiltonian variational integrators. *BIT Numerical Mathematics*, 58(4):1009–1048, 2018.

[36] J. Hong, L. Sun, and X. Wang. High order conformal symplectic and ergodic schemes for the stochastic Langevin equation via generating functions. *SIAM Journal on Numerical Analysis*, 55(6):3006–3029, 2017.

[37] J. Hong, D. Xu, and P. Wang. Preservation of quadratic invariants of stochastic differential equations via Runge–Kutta methods. *Applied Numerical Mathematics*, 87:38 – 52, 2015.

[38] N. Ikeda and S. Watanabe. *Stochastic Differential Equations and Diffusion Processes*. Kodansha scientific books. North-Holland, 1989.

[39] J. A. Izaguirre, D. P. Catarello, J. M. Wozniak, and R. D. Skeel. Langevin stabilization of molecular dynamics. *The Journal of Chemical Physics*, 114(5):2090–2098, 2001.

[40] L. O. Jay. Structure preservation for constrained dynamics with super partitioned additive Runge–Kutta methods. *SIAM Journal on Scientific Computing*, 20(2):416–446, 1998.

[41] L. O. Jay. Lagrange-d'Alembert SPARK integrators for nonholonomic Lagrangian systems. Tech. Rep. 175, Dept. of Mathematics, Univ. of Iowa, USA, 2009.

[42] C. Kane, J. E. Marsden, M. Ortiz, and M. West. Variational integrators and the Newmark algorithm for conservative and dissipative mechanical systems. *International Journal for Numerical Methods in Engineering*, 49(10):1295–1325, 2000.

[43] R. Kleiber, R. Hatzky, A. Könies, K. Kauffmann, and P. Helander. An improved control-variate scheme for particle-in-cell simulations with collisions. *Computer Physics Communications*, 182(4):1005 – 1012, 2011.

[44] P. Kloeden and E. Platen. *Numerical Solution of Stochastic Differential Equations*. Applications of Mathematics : Stochastic Modelling and Applied Probability. Springer, 1995.

[45] M. Kopec. Weak backward error analysis for overdamped Langevin processes. *IMA Journal of Numerical Analysis*, 35(2):583–614, 2014.

[46] M. Kopec. Weak backward error analysis for Langevin process. *BIT Numerical Mathematics*, 55(4):1057–1103, 2015.

[47] M. Kraus. *Variational integrators in plasma physics*. PhD thesis, Technische Universität München, 2013.

[48] M. Kraus, K. Kormann, P. Morrison, and E. Sonnendrücker. GEMPIC: geometric electromagnetic particle-in-cell methods. *Journal of Plasma Physics*, 83(4):905830401, 2017.

[49] H. Kunita. *Stochastic Flows and Stochastic Differential Equations*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 1997.

[50] A. M. Lacasta, J. M. Sancho, A. H. Romero, I. M. Sokolov, and K. Lindenberg. From subdiffusion to superdiffusion of particles on solid surfaces. *Phys. Rev. E*, 70:051104, Nov 2004.

[51] P. S. Landa. Noise-induced transport of Brownian particles with consideration for their mass. *Phys. Rev. E*, 58:1325–1333, Aug 1998.

[52] J. A. Lázaro-Camí and J. P. Ortega. Stochastic Hamiltonian Dynamical Systems. *Reports on Mathematical Physics*, 61(1):65 – 122, 2008.

[53] M. Leok and T. Shingel. General techniques for constructing variational integrators. *Frontiers of Mathematics in China*, 7(2):273–303, 2012.

[54] M. Leok and J. Zhang. Discrete Hamiltonian variational integrators. *IMA Journal of Numerical Analysis*, 31(4):1497–1532, 2011.

[55] A. Lew, J. E. Marsden, M. Ortiz, and M. West. Asynchronous variational integrators. *Archive for Rational Mechanics and Analysis*, 167(2):85–146, 2003.

[56] Q. Ma, D. Ding, and X. Ding. Symplectic conditions and stochastic generating functions of stochastic Runge-Kutta methods for stochastic Hamiltonian systems with multiplicative noise. *Applied Mathematics and Computation*, 219(2):635–643, 2012.

[57] Q. Ma and X. Ding. Stochastic symplectic partitioned Runge-Kutta methods for stochastic Hamiltonian systems with multiplicative noise. *Appl. Math. Comput.*, 252(C):520–534, Feb. 2015.

[58] J. Marsden and T. Ratiu. *Introduction to Mechanics and Symmetry*, volume 17 of *Texts in Applied Mathematics*. Springer Verlag, 1994.

[59] J. E. Marsden, G. W. Patrick, and S. Shkoller. Multisymplectic geometry, variational integrators, and nonlinear PDEs. *Communications in Mathematical Physics*, 199(2):351–395, 1998.

[60] J. E. Marsden and M. West. Discrete mechanics and variational integrators. *Acta Numerica*, 10(1):357–514, 2001.

[61] J. Mattingly, A. Stuart, and D. Higham. Ergodicity for SDEs and approximations: locally Lipschitz vector fields and degenerate noise. *Stochastic Processes and their Applications*, 101(2):185 – 232, 2002.

[62] J. Mattingly, A. Stuart, and M. Tretyakov. Convergence of numerical time-averaging and stationary measures via Poisson equations. *SIAM Journal on Numerical Analysis*, 48(2):552–577, 2010.

[63] R. McLachlan and M. Perlmutter. Conformal Hamiltonian systems. *Journal of Geometry and Physics*, 39:276–300, 2001.

[64] R. I. McLachlan and G. R. W. Quispel. Geometric integrators for ODEs. *Journal of Physics A: Mathematical and General*, 39(19):5251–5285, 2006.

[65] G. Milstein. *Numerical Integration of Stochastic Differential Equations*. Mathematics and Its Applications. Springer Netherlands, 1995.

[66] G. Milstein and M. Tretyakov. Computing ergodic limits for Langevin equations. *Physica D: Nonlinear Phenomena*, 229(1):81 – 95, 2007.

[67] G. N. Milstein, Y. M. Repin, and M. V. Tretyakov. Mean-square symplectic methods for Hamiltonian systems with multiplicative noise. Preprint 670, Weierstraß-Institut für Angewandte Analysis und Stochastik, Berlin, Germany, 2001.

[68] G. N. Milstein, Y. M. Repin, and M. V. Tretyakov. Symplectic integration of Hamiltonian systems with additive noise. *SIAM J. Numer. Anal.*, 39(6):2066–2088, June 2001.

[69] G. N. Milstein, Y. M. Repin, and M. V. Tretyakov. Numerical methods for stochastic systems preserving symplectic structures. *SIAM J. Numer. Anal.*, 40(4):1583 – 1604, 2002.

[70] G. N. Milstein and M. V. Tretyakov. Quasi-symplectic methods for Langevin-type equations. *IMA Journal of Numerical Analysis*, 23(4):593–626, 2003.

[71] T. Misawa. Symplectic integrators to stochastic Hamiltonian dynamical systems derived from composition methods. *Mathematical Problems in Engineering*. Vol. 2010, Article ID 384937, 12 pages, 2010.

[72] S. Ober-Blöbaum. Galerkin variational integrators and modified symplectic Runge–Kutta methods. *IMA Journal of Numerical Analysis*, 37(1):375–406, 2017.

[73] S. Ober-Blöbaum and N. Saake. Construction and analysis of higher order Galerkin variational integrators. *Advances in Computational Mathematics*, 41(6):955–986, 2015.

[74] D. Pavlov, P. Mullen, Y. Tong, E. Kanso, J. E. Marsden, and M. Desbrun. Structure-preserving discretization of incompressible fluids. *Physica D: Nonlinear Phenomena*, 240(6):443–458, 2011.

[75] P. Protter. *Stochastic Integration and Differential Equations*. Stochastic Modelling and Applied Probability. Springer Berlin Heidelberg, 2005.

[76] R. Reigada, A. H. Romero, A. Sarmiento, and K. Lindenberg. One-dimensional arrays of oscillators: Energy localization in thermal equilibrium. *The Journal of Chemical Physics*, 111(4):1373–1384, 1999.

[77] M. Ripoll, M. H. Ernst, and P. Español. Large scale and mesoscopic hydrodynamics for dissipative particle dynamics. *The Journal of Chemical Physics*, 115(15):7271–7284, 2001.

[78] A. Rößler. Runge–Kutta methods for Stratonovich stochastic differential equation systems with commutative noise. *Journal of Computational and Applied Mathematics*, 164-165:613 – 627, 2004.

[79] A. Rößler. Second order Runge–Kutta methods for Stratonovich stochastic differential equations. *BIT Numerical Mathematics*, 47(3):657–680, 2007.

[80] C. W. Rowley and J. E. Marsden. Variational integrators for degenerate Lagrangians, with application to point vortices. In *Decision and Control, 2002, Proceedings of the 41st IEEE Conference on*, volume 2, pages 1521–1527. IEEE, 2002.

[81] J. M. Sanz-Serna. Symplectic integrators for Hamiltonian problems: an overview. *Acta Numerica*, 1:243–286, 1992.

[82] B. L. Sawford. Turbulent relative dispersion. *Annu. Rev. Fluid Mech.*, 33:289–317, 2001.

[83] T. Shardlow. Modified equations for stochastic differential equations. *BIT Numerical Mathematics*, 46(1):111–125, 2006.

[84] R. D. Skeel. Integration schemes for molecular dynamics and related applications. In M. Ainsworth, J. Levesley, and M. Marletta, editors, *The Graduate Student's Guide to Numerical Analysis '98: Lecture Notes from the VIII EPSRC Summer School in Numerical Analysis*, pages 119–176. Springer Berlin Heidelberg, Berlin, Heidelberg, 1999.

[85] E. Sonnendrücker, A. Wacher, R. Hatzky, and R. Kleiber. A split control variate scheme for PIC simulations with collisions. *Journal of Computational Physics*, 295:402 – 419, 2015.

[86] A. Stern, Y. Tong, M. Desbrun, and J. E. Marsden. Variational integrators for Maxwell's equations with sources. *PIERS Online*, 4(7):711–715, 2008.

[87] L. Sun and L. Wang. Stochastic symplectic methods based on the Padé approximations for linear stochastic Hamiltonian systems. *Journal of Computational and Applied Mathematics*, 2016. http://dx.doi.org/10.1016/j.cam.2016.08.011.

[88] D. Talay. Stochastic Hamiltonian systems: Exponential convergence to the invariant measure, and discretization by the implicit Euler scheme. *Markov Processes and Related Fields*, 8(2):163–198, 2002.

[89] D. J. Thomson. Criteria for the selection of stochastic models of particle trajectories in turbulent flows. *Journal of Fluid Mechanics*, 180:529–556, 1987.

[90] T. M. Tyranowski and M. Desbrun. R-adaptive multisymplectic and variational integrators. Preprint arXiv:1303.6796, 2013.

[91] T. M. Tyranowski and M. Desbrun. Variational partitioned Runge-Kutta methods for Lagrangians linear in velocities. Preprint arXiv:1401.7904, 2013.

[92] E. Vanden-Eijnden and G. Ciccotti. Second-order integrators for Langevin equations with holonomic constraints. *Chemical Physics Letters*, 429(1):310 – 316, 2006.

[93] J. Vankerschaver and M. Leok. A novel formulation of point vortex dynamics on the sphere: geometrical and numerical aspects. *J. Nonlin. Sci.*, 24(1):1–37, 2014.

[94] L. Wang. *Variational Integrators and Generating Functions for Stochastic Hamiltonian Systems*. PhD thesis, Karlsruhe Institute of Technology, 2007.

[95] L. Wang and J. Hong. Generating functions for stochastic symplectic methods. *Discrete and Continuous Dynamical Systems*, 34(3):1211–1228, 2014.

[96] L. Wang, J. Hong, and L. Sun. Modified equations for weakly convergent stochastic symplectic schemes via their generating functions. *BIT Numerical Mathematics*, 56(3):1131–1162, 2016.

[97] P. Wang, J. Hong, and D. Xu. Construction of symplectic Runge-Kutta methods for stochastic Hamiltonian systems. *Communications in Computational Physics*, 21(1):237–270, 2017.

[98] W. Zhou, J. Zhang, J. Hong, and S. Song. Stochastic symplectic Runge–Kutta methods for the strong approximation of Hamiltonian systems with additive noise. *Journal of Computational and Applied Mathematics*, 325:134 – 148, 2017.

[99] K. Zygalakis. On the existence and the applications of modified equations for stochastic differential equations. *SIAM Journal on Scientific Computing*, 33(1):102–130, 2011.