Numerical methods for stochastic simulation: when stochastic integration meets geometric numerical integration

Assyr Abdulle

Abstract In this paper we discuss a framework recently introduced to construct and analyse accurate stochastic integrators for the computation of expectation of functionals of a stochastic process for both finite time or long-time dynamics. Such accurate integrators are of interest for many applications in biology, chemistry or physics and are also often needed in multiscale stochastic simulations. We describe how ideas originating from geometric numerical integration or structure preserving methods for deterministic differential equations can help to design new integrators for weak approximation of stochastic differential equations or for long time simulation of ergodic stochastic systems.

1 Introduction

In this paper we review a recently developed framework to construct and analyse efficient numerical methods to approximate expectation of a functional of stochastic processes. This is a basic problems for many applications in biology, chemistry or physics [15]. For example in molecular dynamics where a fundamental issue is the computation of macroscopic quantities, typically functionals of some variables of the system with respect to a given probability measure often given by the Boltzmann-Gibbs density. The associated numerical problem consists in solving high-dimensional integrals that are most often approximated through ergodic averages of stochastic dynamics obtained from solutions of stochastic differential equations (SDEs), e.g., Langevin SDEs [24]. The approximation of functionals of a stochastic process also arise in multiscale stochastic systems. In the SDE context, one would like to solve for

Section of Mathematics, Swiss Federal Institute of Technology (EPFL), Station 8, CH-1015, Lausanne, Switzerland. *E-mail address:* assyr.abdulle@epfl.ch

example systems of the type [1]

$$dX = f(X,Y)dt,$$
$$dY = \frac{1}{\varepsilon}g(X,Y)dt + \frac{1}{\sqrt{\varepsilon}}\sigma(X,Y)dW(t),$$

where $W(t)$ is a Wiener process and ε is a parameter $\varepsilon \ll 1$. Classical stochastic solvers will need a time-step that resolves the fast dynamics resulting in a large computational cost. For classes of such multi-scale problems, averaging or homogenization techniques [33] can lead to more efficient numerical integrators. This is the case for example when an effective macro dynamics exists, i.e., when the existence of an equation of the form $d\bar{X} = \lim_{\varepsilon\to 0}\int f(\bar{X},Y)d\mu^{\varepsilon}_{\bar{X}}(dY)dt = F(\bar{X})dt$, where $\mu^{\varepsilon}_{\bar{X}}$ is an invariant measure for the fast dynamics of the above system with $X = \bar{X}$ fixed can be established. In this situation, one can implement a multiscale scheme that consists in sampling the fast variable while the slow variable X_n fixed at time t_n, perform an averaging to recover an approximation $\tilde{F}(X_n)$ of the force $F(X_n)$ and use a macroscopic solver to advance the effective dynamics to time t_{n+1}, e.g., $X_{n+1} = X_n + h\tilde{F}(X_n)$ (see [39, 13, 2]). Such multiscale methods have also been developed for stochastic partial differential equations [3, 8]. One of the main issue for such fast-slow numerical techniques is the computational cost of the repeated computation of the effective forces, relying on the approximation of the invariant measure of the fast process. Accurate approximation of invariant measures is a central issue in such stochastic computations [13, 26]. We note that an algorithm based on the solution of a Poisson problem obtained from the generator of the fast system is also promising for classes of SDEs (or SPDEs) with multiple scales [4]. For the numerical approximation of ergodic SDEs, one faces several important questions:

1. does the numerical method have an invariant measure ?
2. how close is the numerical invariant measure to the true one ?
3. how close is the time-averaging method to the invariant measure ?

In this paper we will mainly discuss the second question. We note that many authors have discussed the first question (see [27, 38, 35, 31, 32] and the references in these papers), for the third question there are also a large body of contributions and we refer to the text book [20, 30, 17] for references.

Fast-slow processes are also ubiquitous in biology when simulating N chemical species $\{S_i\}_{i=1}^N$ interacting through M reaction channels $\{R_i\}_{i=1}^M$. The state of the system is specified by the vector $\mathbf{X}_t = (X_{1t}, \ldots, X_{Nt})^T$ that can be shown to evolve according to a Markov process. Sampling trajectories can be computed according to the stochastic simulation algorithm by updating the waiting time to the next reaction and selecting the next reaction that occurs. This numerical algorithm is called the stochastic simulation

[1] See Section 1.1 for a precise definition.

algorithm (or Gillespie algorithm see [16] for a review). In a multiscale context, when some reaction channels occur frequently on a timescale for which others will only rarely take place, a large computational effort will be needed to see the dynamics of some slower reaction channels. In this context, multiscale algorithms that share similarities with the above SDE situation have been developed. We mention the nested SSA [14], the slow-scale SSA [10] that are both based on quasi-steady approximation (the time scale separation between fast and slow processes allows for the fast process to equilibrate before significant change in the slow process occur). Here again, part of the computational effort is devoted to compute the equilibrium of fast process.

The framework presented below for constructing and analyzing stochastic integrators for the computation of expectation of functionals of stochastic processes for both finite time or long-time dynamics has been introduced in [1, 5, 6]. While this framework has been applied to SDEs, applications to discrete stochastic processes might be an interesting topic to explore in the future.

1.1 Setting and definitions

We consider a d-dimensional SDE

$$dX = f(X)dt + g(X)dW(t), \quad X(0) = X_0, \tag{1}$$

where $X_0 \in E$ is the initial condition assumed deterministic for simplicity, and $W(t)$ is a standard m-dimensional Wiener process. The maps $f : E \mapsto E$, $g : E \mapsto E^m$ are assumed to be smooth and the space E denotes either $E = \mathbb{R}^d$ or the torus $E = \mathbb{T}^d$, and is specified when needed. We will sometimes use the vector notation $g = (g^1, \dots, g^m)$ for the matrix $g(x)$, where $g^i(x) \in E$.

We consider a discrete a discrete numerical approximation of (1) given by

$$X_{n+1} = \Psi(X_n, h, \xi_n), \tag{2}$$

for $X_n \in E$ for $n \geq 0$, where $\Psi(\cdot, h, \xi_n) : E \to E$ is the discrete numerical flow, h denotes the timestep size, and ξ_n denotes a random vector.

Several concepts of convergence can be used to measure how well the numerical method (2) approximates the solution of (1). We briefly review here strong and weak convergence, as well as convergence with respect to the invariant measure of (1) provided such invariant measures exist.

Strong and weak convergence

We note by $C_P^\ell(\mathbb{R}^d, \mathbb{R})$ the space of $1 \leq \ell \leq \infty$ times continuously differentiable functions $\mathbb{R}^d \to \mathbb{R}$ with all partial derivatives with polynomial growth.

Likewise $C^\ell(\mathbb{T}^d,\mathbb{R})$, will denote the space of $1 \leq \ell \leq \infty$ times continuously differentiable functions $\mathbb{T}^d \to \mathbb{R}$. To simplify the notation, we will define $\mathcal{V}^\ell(E,\mathbb{R})$ to denote either $C^\ell_P(\mathbb{R}^d,\mathbb{R})$ or $C^\ell(\mathbb{T}^d,\mathbb{R})$ and $\mathcal{V}^\ell(E,E)$ to denote either $C^\ell_P(\mathbb{R}^d,\mathbb{R}^d)$ or $C^\ell(\mathbb{T}^d,\mathbb{T}^d)$. When needed the specific situation will be mentioned.

The numerical approximation (2), starting from the exact initial condition X_0 of (1) is said to have weak order p if for all functions $\phi \in \mathcal{V}^{2(p+1)}(E,\mathbb{R})$

$$|\mathbb{E}(\phi(X_n)) - \mathbb{E}(\phi(X(t_n)))| \leq Ch^p, \tag{3}$$

and to have strong order p if

$$\mathbb{E}(|X_n - X(t_n)|) \leq Ch^p, \tag{4}$$

for any $t_n = nh \in [0,T]$ with $T > 0$ fixed, for all h small enough, with constants C independent of h.

Remark 1. We will sometimes use the result [29] (see [30, Chap. 2.2]) of Milstein that allows to deduce global weak order from the local weak error, i.e., the weak error after one step. This results is as follows: assuming that $f, g^r \in \mathcal{V}^{2(p+1)}(E,E), r = 1,\dots,m$ are Lipschitz continuous, that for all $r \in \mathbb{N}$, the moments $\mathbb{E}(|X_n|^{2r})$ are bounded for all n, h with $0 \leq nh \leq T$ uniformly with respect to all h sufficiently small (this condition only applies for the case $E = \mathbb{R}^d$), and that for all $\phi \in \mathcal{V}^{2(p+1)}(E,\mathbb{R})$ and all initial values $X(0) = X_0$ the local weak error satisfies

$$|\mathbb{E}(\phi(X_1)) - \mathbb{E}(\phi(X(t_1)))| \leq Ch^{p+1} \tag{5}$$

for all h sufficiently small, then the global error bound (3) holds. For the strong error, conditions on local weak and strong errors are needed to infer global strong convergence (see [30] for details).

Long-time behavior: approximation of the invariant measure

The above strong and weak convergence measure finite time approximation properties of the numerical solver (2) when applied to (1). Of interest in many applications is the long-time approximation of ergodic SDEs (1), i.e., SDEs that have a unique invariant measure μ satisfying for each smooth integrable function ϕ and for any deterministic initial condition X_0,

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T \phi(X(s))ds = \int_E \phi(y)d\mu(y), \qquad \text{almost surely.} \tag{6}$$

In this paper we will assume that the numerical method (2) is ergodic, i.e., that it has a unique invariant probability law μ^h with finite moments of any order and

$$\lim_{N\to\infty} \frac{1}{N+1} \sum_{n=0}^{N} \phi(X_n) = \int_E \phi(y) d\mu^h(y), \qquad \text{almost surely,} \tag{7}$$

for all deterministic initial condition X_0 and all smooth test functions ϕ. To quantify the second question we will say that the numerical method (2) has order $p \geq 1$ with respect to the invariant measure if

$$|e(\phi, h)| \leq Ch^p \quad \text{with} \quad e(\phi, h) := \lim_{N\to\infty} \frac{1}{N+1} \sum_{n=0}^{N} \phi(X_n) - \int_E \phi(y) d\mu(y), \tag{8}$$

where h is small enough and C is independent of h and X_0 and the above limit holds with probability one.

2 Tools inspired from geometric numerical integration

In this section we recall some classical tools of geometric numerical integration for ODEs, developed in particular for the analysis of the long time dynamics of symplectic integrators. We refer to [18] for a comprehensive presentation and detailed proofs of the claims made below. Consider an ordinary differential equation written in autonomous form for simplicity

$$\begin{aligned} \frac{dX}{dt} &= f(X) \\ X(0) &= X_0, \end{aligned} \tag{9}$$

where $t \geq 0$, $f : E \longrightarrow E$. We denote by $\varphi_t(x)$ the exact flow of this differential equation. We consider a one-step numerical method with constant time-step h

$$X_{n+1} = \psi_h(X_n; f), \tag{10}$$

where $\psi_h(\cdot; f) : E \longrightarrow E$ is the discrete numerical flow and the second argument of ψ_h just emphasizes the differential equation the numerical integrator is applied to. We recall that the numerical method (10) is said to have order p if

$$\psi_h(x; f) - \varphi_h(x) = \mathcal{O}(h^{p+1}), \tag{11}$$

for all sufficiently smooth differential equations (9).

Backward error analysis

We assume that the numerical method can be expanded as

$$\psi_h(x;f) = x + hf(x) + h^2 d_1(x) + h^3 d_2(x) + \dots \tag{12}$$

We note that the form of the two first terms in the above expansion is required for a consistent method. The idea of backward error analysis is to find a modified differential equation

$$\begin{aligned} \frac{d\widetilde{X}}{dt} &= \tilde{f}_{h,s}(\widetilde{X}), \\ \widetilde{X}(0) &= X_0, \end{aligned} \tag{13}$$

where

$$\tilde{f}_{h,s}(x) = f(x) + h\tilde{f}_1(x) + h^2\tilde{f}_2(x) + \dots + h^s\tilde{f}_s(x), \tag{14}$$

and to construct $\tilde{f}_i(x)$ such that $\widetilde{X}(h) - \psi_h(x;f) = \mathcal{O}(h^{s+2})$ for $h \to 0$. It then follows that $\widetilde{X}(t_n) = X_n + \mathcal{O}(h^{s+1})$ for $h \to 0$ and bounded times $t_n = nh \le T$. We denote by $\tilde{\varphi}_t(x)$ the flow of the differential equation (13). [2] To compute the function $\tilde{f}_2, \tilde{f}_3, \dots$ one expand the solution of (13) into a Taylor series and compare the power of h with (12). In this respect, we have the following lemma

Lemma 1. *If the numerical method* (10) *is of order p with $\psi_h(x;f) - \varphi_h(x) = h^{p+1}\delta_{p+1}(x) + \mathcal{O}(h^{p+2})$ then (assuming $s \ge p$) the function $\tilde{f}_{h,s}$ given by* (14) *satisfies*

$$\tilde{f}_{h,s}(x) = f(x) + h^p\tilde{f}_p(x) + \dots h^s\tilde{f}_s(x), \tag{15}$$

where $\tilde{f}_p(x) = \delta_{p+1}(x)$.

We next rewrite the exact and modified flows in terms of differential operators. If we define the Lie derivative by $\mathcal{L}_D = f \cdot \nabla$ then, provided that f is $M+1$ times continuously differentiable, the flow of (9) satisfies

$$\phi(\varphi_t(x)) = \sum_{k=0}^{M} \frac{t^k}{k!}(\mathcal{L}_D^k\phi)(x) + \mathcal{O}(h^{M+1}), \tag{16}$$

for all smooth yest functions ϕ. Liekewise for the modified equation (13) we have

$$\phi(\tilde{\varphi}_t(x)) = \sum_{k=0}^{M} \frac{t^k}{k!}(\widetilde{\mathcal{L}}_D^k\phi)(x) + \mathcal{O}(h^{M+1}), \tag{17}$$

[2] Formally in backward error analysis one consider an infinite series $\tilde{f}_h(x) = f(x) + h\tilde{f}_1(x) + h^2\tilde{f}_2(x) + \dots$ such that $\widetilde{X}(t_n) = X_n$ at $t_n = nh$. But the infinite series for $\tilde{f}_h$ is usually not convergent and one needs therefore to consider an appropriate truncation.

where $\widetilde{\mathcal{L}}_D = \tilde{f}_{h,s} \cdot \nabla$. We note that if f is analytic in a complex neighborhood of x and the one-step integrator (10) is a Runge-Kutta method, then the coefficients $d_j(x)$ of (12) and the functions $\tilde{f}_j(x)$ of the modified equation (13) are also analytic [18, Chap. 9]. The expressions (16) and (17) then read $\phi(\varphi_t(x)) = e^{t\mathcal{L}_D}\phi(x)$ and $\phi(\tilde{\varphi}_t(x)) = e^{t\widetilde{\mathcal{L}}_D}\phi(x)$, respectively. If a numerical method as order p then it holds

$$e^{h\mathcal{L}_D}\phi(x) - \phi(\psi_h(x; f)) = \mathcal{O}(h^{p+1})$$

and by definition of the modified equation (13) we have

$$e^{h\widetilde{\mathcal{L}}_D}\phi(x) - \phi(\psi_h(x; f)) = \mathcal{O}(h^{s+2}).$$

Integrator based on modified equations

The approach of modifying integrators that first appeared in [11] consists in using a modified differential equations in order to construct higher order numerical integrators while preserving geometric properties like symplecticity. Consider an integrator (10) of order p for the problem (9). The idea to derive a more precise numerical integrator (still for the problem (9)) is to construct a modified differential equation for $s \geq p$

$$\begin{aligned} \frac{d\hat{X}}{dt} &= \hat{f}_{h,s}(\hat{X}) \\ \hat{X}(0) &= X_0, \end{aligned} \tag{18}$$

where

$$\hat{f}_{h,s}(x) = f(x) + h^p \hat{f}_p x) + \ldots + h^s \hat{f}_s(x) \tag{19}$$

such that when applying the integrator (10) to (18) it has order $s+1$ for the original equation (9) that is,

$$\psi_h(x; \hat{f}_{h,s}) - \varphi_h(x) = \mathcal{O}(h^{s+2}),$$

or in terms of Lie derivatives

$$e^{h\mathcal{L}_D}\phi(x) - \phi(\psi_h(x; \hat{f}_{h,s})) = \mathcal{O}(h^{s+2}).$$

The functions $\hat{f}_i(x)$ can be uniquely defined by the condition that $X(h) = \psi_h(x; \hat{f}_{h,s}) + \mathcal{O}(h^{s+2})$ for $h \to 0$ and it then follows that $X(t_n) = \psi_h(X_{n-1}; \hat{f}_{h,s}) + \mathcal{O}(h^{s+1})$ for $h \to 0$ and bounded times $t_n = nh \leq T$.

3 A framework for constructing and analysing stochastic integrators

We recall that associated to the SDE (1), there exists a differential operator $\mathcal{L}$, called the generator of the SDE, defined by

$$\mathcal{L} := f \cdot \nabla + \frac{1}{2} g g^T : \nabla^2, \tag{20}$$

where $\nabla^2 \phi$ denotes the Hessian of a function ϕ (scalar product on matrices are denoted $A : B = \text{trace}(A^T B)$). Then the function $u(t,x) = \mathbb{E}\left(\phi(X(t))|X_0 = x\right)$ is the solution of the partial differential equation, called the backward Kolmogorov equation, given by

$$\frac{\partial u}{\partial t} = \mathcal{L} u, \qquad u(x,0) = \phi(x), \tag{21}$$

where $\phi \in \mathcal{V}^\infty(E, \mathbb{R})$. Using a Taylor expansion together with equation (21) gives a series for $u(t,x)$ in terms of the generator of the SDE [41, 12]

$$u(x,h) - \phi(x) = \sum_{j=1}^{l} \frac{h^j}{j!} \mathcal{L}^j \phi(x) + h^{l+1} r_l(f,g,\phi)(x), \tag{22}$$

where under appropriate smoothness on f, g, ϕ the remainder $r_l(f,g,\phi)$ has polynomial growths ($E = \mathbb{R}^d$) or can be bounded ($E = \mathbb{T}^d$). We next consider for the numerical method (2) the function

$$U(x,h) = \mathbb{E}(\phi(X_1)|X_0 = x), \tag{23}$$

and assume that it can be expanded as

$$U(x,h) = \phi(x) + h A_0(f,g)\phi(x) + h^2 A_1(f,g)\phi(x) + \dots, \tag{24}$$

where $A_i(f,g),\ i = 0,1,2,\dots$ are linear differential operators that depend on the choice of the integrator with coefficients depending on f, g, and their derivatives. Assume also that for all $i = 0,1,2,\dots$

$$A_i(f + \eta\hat{f}, g + \eta\hat{g}) = A_i(f,g) + \eta A_i(f,\hat{f},g,\hat{g}) + \mathcal{O}(\eta^2), \tag{25}$$

where $A_i(f,\hat{f},g,\hat{g}) + \mathcal{O}(\eta^2)$ is again a differential operator. As (see (5)) we have $|\mathbb{E}(\phi(X_1)) - \mathbb{E}(\phi(X(t_1)))| = |U(x,h) - u(x,h)|$ for all methods of weak order $p \geq 1$ we must have $A_0 = \mathcal{L}$ (consistency condition). Furthermore, a method of weak local order $p \geq 1$ satisfies

$$|\mathbb{E}(\phi(X_1)) - \mathbb{E}(\phi(X(t_1)))| = h^{p+1} \left(A_p - \frac{\mathcal{L}^{p+1}}{(p+1)!} \right) \phi(x) + \mathcal{O}(h^{p+2}). \tag{26}$$

A global weak order result can also be expressed in terms of the above differential operators [30, Chap. 2.2, 2.3]. Indeed we have

Theorem 1. *Assume that f, g in (1) are C^∞ with bounded derivatives up to any order and consider a numerical integrator (2) on $[0,T]$ with an expansion of the form (24) with bounded moments $\mathbb{E}|X_n|^\kappa$, $\kappa \in \mathbb{N}$ sufficiently large. Assume that the numerical integrator has weak local p with a constant $C = C(x)$ with polynomial growth. Then, we have the following expansion of the global error, for all $\phi \in \mathcal{V}^\infty(E, \mathbb{R})$,*

$$\mathbb{E}(\phi(X(T))) - \mathbb{E}(\phi(X_N)) = h^p \int_0^T \mathbb{E}(\psi_e(X(s), s))ds + \mathcal{O}(h^{p+1}), \tag{27}$$

where $Nh = T$ and $\psi_e(x,t)$ satisfies

$$\psi_e(x,t) = \left(A_p - \frac{1}{(p+1)!}\mathcal{L}^{p+1} \right) u(x,t), \tag{28}$$

and $u(x,t)$ is the solution of (21).

3.1 High weak order method based on modified equations

We observe that an expansion of the type (24) holds for many numerical integrators and hence (24) is not a restrictive assumption.

Example 1. We start with an example and consider the SDE (1) with $d = m = 1$ and define for $X_0 = x$ the (semi-implicit) θ-Milstein method by

$$\begin{aligned} X_{n+1} = X_n &+ (1-\theta)hf(X_n) + \theta h f(X_{n+1}) + g(X_n)\Delta W_n \\ &+ \frac{1}{2} g'(X_n) g(X_n)((\Delta W_n)^2 - h), \end{aligned} \tag{29}$$

where the Wiener increment ΔW_n are given by independent $\mathcal{N}(0,h)$ random variables. Taylor expansion of (23) reveal that $U(x,h) = \phi(x) + hA_0(f,g)\phi(x) + h^2 A_1(f,g)\phi(x) + \mathcal{O}(h^3)$ with

$$\begin{aligned} A_1(f,g)\phi(x) = &\theta \left[f'(x)f(x) + \frac{1}{2} f''(x) g^2(x) \right] \phi'(x) \\ &+ \frac{1}{2} \left[f^2(x) + 2\theta f'(x) g^2(x) + \frac{1}{2}(g'(x)g(x))^2 \right] \phi''(x) \qquad (30) \\ &+ \frac{1}{2} \left[g'(x) g^3(x) + g^2 f(x) \right] \phi'''(x) + \frac{h^2}{8} g^4(x) \phi^{(4)}(x). \qquad (31) \end{aligned}$$

An easy computation then shows that

$$\left(\frac{1}{2}\mathcal{L}^2\phi - A_1(f,g)\phi\right)(x) = \left(\frac{1}{2}-\theta\right)\left(f'(x)f(x) + \frac{1}{2}f''(x)g^2(x)\right)\phi'(x)$$
$$+\left(\left(\frac{1}{2}-\theta\right)f'(x)g(x) + \frac{1}{2}g'(x)f(x) + \frac{1}{4}g^2(x)g''(x)\right)g(x)\phi''(x).$$

In view of Theorem 1 we can deduce that this integrator has weak order one and applying the integrator to a simple particular SDE (e.g. linear) reveals that it will not be second order in general. Note that for $\theta = 0$ we recover the Euler-Maruyama (EM) method $X_{n+1} = X_n + hf(X_n) + g(X_n)\Delta W_n$ that has also weak order one. However, the more involved θ-Milstein method has better strong and mean-square stability behavior for $\theta > 0$ [19].

Construction of higher order weak methods based on a basic classical integrator such as for example (29) can be achieved by using the framework of modified equations summarized in Section 2. Consider the following modified SDE based on (1)

$$d\hat{X} = f_{h,s}(\hat{X})dt + g_{h,s}(\hat{X})dW(t), \quad \hat{X}(0) = X_0, \tag{32}$$

where

$$f_{h,s}(x) = f(x) + hf_1(x) + \ldots + h^s f_s(x), \tag{33}$$
$$g_{h,s}(x) = g(x) + hg_1(x) + \ldots + h^s g_s(x). \tag{34}$$

Inserting the expansion for $f_{h,s}, g_{h,s}$ into the generator of the SDE (32) given by

$$\hat{\mathcal{L}}\phi := f_{h,s}\cdot\nabla_x\phi + \frac{1}{2}(g_{h,s}g_{h,s}^T) : \nabla_x^2\phi, \tag{35}$$

yields

$$\hat{\mathcal{L}} = \mathcal{L} + h\mathcal{L}_1 + h^2\mathcal{L}_2 + \ldots + h^s\mathcal{L}_s + \mathcal{O}(h^{s+1}), \tag{36}$$

where for $j = 1, 2, \ldots s$ the operators $\mathcal{L}_j$ is given by

$$\mathcal{L}_j = f_j\cdot\nabla_x + \frac{1}{2}\sum_{k=0}^{j}(g_k g_{j-k}^T) : \nabla_x^2, \tag{37}$$

using the the notations $f_0 := f$ and $g_0 := g$. We will sometimes write $\mathcal{L}_j(f_j, g, g_1, \ldots, g_j)$ to emphasize on which functions the coefficients of the operator depend. Going back to the Example 1, we set $p = s = 1$ in the modified equations (33) and we obtain using Assumption (25) for the θ-Milstein method applied to (32) the expansion

$$\begin{aligned} U(x,h) &= \phi(x) + h\hat{\mathcal{L}}\phi(x) + h^2 A_1(f_{h,1}, g_{h,1})\phi(x) + \ldots, \\ &= \phi(x) + h\mathcal{L} + h^2(\mathcal{L}_1 + A_1(f,g))\phi(x) + \mathcal{O}(h^3). \end{aligned}$$

Hence to obtain a second order weak integrator for the SDE (1) we must in view of Theorem 1 define f_1, g_1 in the modified equations such that

$$\mathcal{L}_1 = (\frac{1}{2}\mathcal{L}^2 - A_1),$$

and in view of (30) we easily find that

$$f_1(x) = \left(\frac{1}{2} - \theta\right) f'(x)f(x) + \frac{1}{2}\left(\frac{1}{2} - \theta\right) f''(x)g^2(x),$$
$$g_1(x) = \left(\frac{1}{2} - \theta\right) f'(x)g(x) + \frac{1}{2}g'(x)f(x) + \frac{1}{4}g^2(x)g''(x).$$

hence the integrator

$$X_{n+1} = X_n + (1-\theta)hf_{h,1}(X_n) + \theta h f_{h,1}(X_{n+1}) + g_{h,1}(X_n)\Delta W_n + \frac{1}{2}g'(X_n)g(X_n)((\Delta W_n)^2 - h). \tag{38}$$

has weak order two when applied to a one-dimensional SDE (1). A priori the last term of the above integrator should read $\frac{1}{2}g'_{h,1}(X_n)g_{h,1}(X_n)((\Delta W_n)^2-h)$ but it can be shown (direct calculation) that

$$\mathbb{E}\left(g'_{h,1}(X_n)g_{h,1}(X_n)((\Delta W_n)^2 - h)\right) = \mathbb{E}(g'(X_n)g(X_n)((\Delta W_n)^2-h))+\mathcal{O}(h^3), \tag{39}$$

and hence we can substitute $g_{h,1}$ with g in the last term of the modified integrator without altering the weak order two of the method.

The scheme (39) can easily be generalized to multidimensional SDEs (see [1]). We note that in general, constructing second order weak methods require to solve many order conditions (system of equations) if relying on Itô-Taylor expansion [36]. The integrator (38) derived in [1] belongs to a general class of weak second order integrators derived by Milstein [29]. For $\theta = 0$ Talay proved second order convergence in [37] and for $\theta = 1/2$, the scheme was shown to have favorable stability properties for scalar SDEs with additive noise [19]. For $\theta = 1$, the method seems to have first appeared in [1] where it is also shown that it possesses good mean-square stability properties for scalar SDEs with multiplicative noise.

A general results for high order weak integrator based on modified equations can be obtained recursively as follows: consider a numerical integrator (2) of order p for the SDE (1) and assume that a modified equation (33) with $s = p + r - 2$ has beed obtained so that (2) applied to (32) is an integrator of weak order $p + r - 1$ for the original SDE (1), i.e.,

$$\begin{aligned} U(x,h) &= \phi(x) + hA_0(f_{h,p+r-2}, g_{h,p+r-2}) + \dots \\ &\quad + h^{p+r-1}A_{p+r-2}(f_{h,p+r-2}, g_{h,p+r-2}) + \mathcal{O}(h^{p+r+1}) \\ &= \phi(x) + h\mathcal{L} + \ldots + \frac{h^{p+r-1}}{(p+r-1)!}\mathcal{L}^{p+r-1} + h^{p+r}R(f,g). \end{aligned}$$

Consider now the SDE (1) and a modified equations with $s = p + r - 1$

$$\begin{aligned} f_{h,p+r-1}(x) &= f_{h,p+r-2}(x) + h^{p+r-1}f_{p+r-1}(x), \\ g_{h,p+r-1}(x) &= g_{h,p+r-2}(x) + h^{p+r-1}g_{p+r-1}(x), \end{aligned}$$

then in view of the above expansion for $U(x,h)$ we obtain

$$\begin{aligned} U(x,h) &= \phi(x) + hA_0(f_{h,p+r-1}, g_{h,p+r-1}) + \ldots + h^{p+r}A_{p+r-1}(f_{h,p+r-1}, g_{h,p+r-1}) \\ &\quad + \mathcal{O}(h^{p+r+1}) \\ &= \phi(x) + h\mathcal{L} + \ldots + \frac{h^{p+r-1}}{(p+r-1)!}\mathcal{L}^{p+r-1} \\ &\quad + h^{p+r}(\mathcal{L}_{p+r-1}(f_{p+r-1}, g, g_1, \ldots, g_{p+r-1}) + R(f,g)) + \mathcal{O}(h^{p+r+1}), \end{aligned}$$

where we used Assumption (25), the equality $A_0(f_{h,p+r-1}, g_{h,p+r-1}) = \hat{\mathcal{L}}(f_{h,p+r-1}, g_{h,p+r-1})$ the expansion (36) and the fact that integrator is of order $p+r-1$ for the original SDE (1). We see that if we can find f_{p+r-1}, g_{p+r-1} such that the differential operator $\mathcal{L}_{p+r-1}$ satisfies

$$\mathcal{L}_{p+r-1}(f_{p+r-1}, g, g_1, \ldots, g_{p+r-1}) = \frac{1}{(p+r)!}\mathcal{L}^{p+r} - R(f,g)$$

then according to Theorem 1, the integrator (2) will be of order $p+r$ for the original SDE (1).

The framework presented above and introduced in [1] has been further used to construct new high weak order methods that are mean square stable and new high weak order invariant preserving stochastic methods. In Sections 3.2,4.1,4.2 we will further explain how this framework can be used to construct new high order methods for the approximation of invariant measure of ergodic SDEs following [5, 6].

3.2 High order numerical approximation of the invariant measure of ergodic SDEs

We now explain how ideas from backward error analysis and modified equations can lead to efficient approximations of the invariant measure or ergodic SDEs. We assume now that the SDE (1) is ergodic (see assumptions 3.1 or 3.2) and the that its unique invariant measure μ has a density function ρ_∞.

We recall ρ_∞ is then the unique solution of the Fokker-Planck equation

$$\mathcal{L}^*\rho_\infty = 0, \tag{40}$$

where $\mathcal{L}^*\phi = -\nabla\cdot(\phi f) + \frac{1}{2}gg^T : \nabla^2\phi$ is the L^2-adjoint of the generator $\mathcal{L}$ defined in (20). We consider a numerical integrator (2) that we assume to be ergodic and to have an expansion of the form (24) satisfying (25). To motivate our approach, consider a numerical integrator of weak order p. Passing to the limit $T\to\infty$ in (27) we obtain for all $\phi\in\mathcal{V}^\infty(E,\mathbb{R})$ and $h\to 0$,

$$e(\phi,h) = \lambda_p h^p + \mathcal{O}(h^{p+1}) \tag{41}$$

where $e(\phi,h)$ in defined in (8), for any deterministic initial condition, with λ_p defined as

$$\lambda_p = \int_0^{+\infty}\int_{\mathbb{R}^d}\left(A_p - \frac{1}{(p+1)!}\mathcal{L}^{p+1}\right)u(y,t)\rho_\infty(y)dydt \tag{42}$$

where $e(\phi,h)$ is defined in (8) and $u(x,t)$ is the solution of (21). Hence an ergodic numerical integrator of weak order p is also of order p with respect to the invariant measure of (1). Next using the L^2-adjoint of the differential operator $\frac{1}{(p+1)!}\mathcal{L}^{p+1} - A_p$ reveals

$$\lambda_p = -\int_0^{+\infty}\int_{\mathbb{R}^d} u(y,t)\left(A_p^* - \frac{1}{(p+1)!}(\mathcal{L}^*)^{p+1}\right)\rho_\infty(y)dydt.$$

Now from (40), $(\mathcal{L}^*)^{p+1}\rho_\infty = 0$, hence if in addition to have weak order p the numerical integrator also satisfies $A_p^*\rho_\infty = 0$ then it will have order $p+1$ with respect to the invariant measure. Of course if it happens that the integrator is of weak order $p+1$, then necessarily $A_p^*\rho_\infty = 0$, since in that case $\mathcal{L}^{p+1}/(p+1)! = A_p$. It turns out that that there exist integrators of weak order p but not $p+1$ that still fulfill $A_p^*\rho_\infty = 0$.

We now discuss a generalization of the above result and derive order conditions for constructing numerical integrators that approximate the invariant measure of ergodic SDEs with high order. The starting point is again the expansion (24) for the numerical method. We also derive a modified generator

$$\widetilde{\mathcal{L}} = \mathcal{L} + \sum_{i\geq 1} h^i L_i, \tag{43}$$

but we now require, following the framework of backward error analysis, that its expansion match (formally) the expansion of the numerical method, i.e.,

$$U(x,h) - \phi(x) = \sum_{j\geq 1}\frac{h^j}{j!}\widetilde{\mathcal{L}}^j\phi(x).$$

Using the the above equality and matching equal power of h on obtains using (24) [41, 12]

$$L_n = A_n - \frac{1}{2}(\mathcal{L}L_{n-1} + L_{n-1}\mathcal{L} + \ldots) - \cdots - \frac{1}{(n+1)!}\mathcal{L}^{n+1}. \qquad (44)$$

Compared to the framework for ODEs recalled in Section (2) we face an additional difficulty when trying to apply the ideas of backward error analysis in the SDE context. Indeed, in view of (13), one would be tempted to truncate the formal expansion for $\widetilde{\mathcal{L}}$, say $\widetilde{\mathcal{L}}_N = \mathcal{L} + \sum_{i=1}^N h^i L_i$, and to consider the backward Kolmogorov equation

$$\frac{\partial \tilde{u}_N}{\partial t} = \widetilde{\mathcal{L}}_N \tilde{u}_N.$$

However, the existence of a solution to the above PDE is not clear as $\widetilde{\mathcal{L}}_N$ is no longer a second order operator in general. The proper definition of the equation (43) it based on the following result. Under appropriate assumptions on f, g assume further there exists a constant λ and for all integer $k \geq 0$ constants C_k, κ_k such that for all $t \geq 0$

$$\|u(t,\cdot) - \int_{E^d} \phi(y)\rho_\infty(y)dy\|_{\mathcal{C}^k} \leq C_k(1 + t^{\kappa_k})e^{-\lambda t}\|\phi\|_{\mathcal{C}^k}, \qquad (45)$$

where $\|v(t,\cdot)\|_{\mathcal{C}^k}$ denotes the sup norm of the function $v(x,t)$ and its derivatives with respect to x up to order k. Then is has been shown in [12] ($E = \mathbb{T}$) and in [22] ($E = \mathbb{R}$) that for all $\ell \in \mathbb{N}$, there exists smooth functions $\tilde{u}_\ell(t,x)$ defined for all $t \geq 0$ that satisfies for all $N \in \mathbb{N}$

$$\frac{\partial \tilde{u}_N}{\partial t} - L\tilde{u}_N = \sum_{i=1}^N L_i v_{N-1}.$$

This result is used in the following lemma that is central in our numerical framework. We will consider two distinct situations either

Assumption 3.1 *$E = \mathbb{T}^d$ and*

- *f, g are C^∞ functions on the torus $\mathbb{T}^d$;*
- *the generator $\mathcal{L}$ is elliptic or hypo-elliptic;*
- *in the case where $\mathcal{L}$ is hypo-elliptic, we further assume the uniqueness of the invariant measure of (1).*

or

Assumption 3.2 *$E = \mathbb{R}^d$ and*

1. *f, g are of class C^∞, with bounded derivatives of any order, and g is bounded;*

2. *the generator* $\mathcal{L}$ *in* (20) *is a uniformly elliptic operator,* i.e. *there exists* $\alpha > 0$ *such that for all* $x, \xi \in \mathbb{R}^d$, $x^T g(\xi) g(\xi)^T x \geq \alpha x^T x$;
3. *there exist* $C, \beta > 0$ *such that for all* $x \in \mathbb{R}^d$, $x^T f(x) \leq -\beta x^T x + C$.

We note that under either of the above assumptions, the SDE (1) has a unique invariant measure. For the case $E = \mathbb{R}^d$, (3.2) also implies that the density function of ρ_∞ of the invariant measure has bounded moments of any order, i.e., for all $n \geq 0$

$$\int_{\mathbb{R}^d} |x|^n \rho_\infty(x) dx < \infty. \tag{46}$$

The following lemma is valid either for the torus $E = \mathbb{T}^d$ [12] under Assumption 3.1 or for $E = \mathbb{R}^d$ under Assumption 3.2 [22].

Lemma 2. *Consider* L_n *the operators defined in* (44). *Then there exists a sequence of functions* $(\rho_n(x))_{n\geq 0}$ *such that* $\rho_0 = \rho_\infty$ *and for all* $n \geq 1$, $\int_E \rho_n(x) dx = 0$ *and*

$$\mathcal{L}^* \rho_n = -\sum_{l=1}^{n} (L_l)^* \rho_{n-l}. \tag{47}$$

For any positive integer N, *setting*

$$\rho_N^h(x) = \rho_\infty(x) + \sum_{n=1}^{N} h^n \rho_n(x), \tag{48}$$

then there exists a constant $C(N, \phi)$ *such that for all* $\phi \in \mathcal{V}^\infty(E, \mathbb{R})$

$$\left| \int_{E^d} \phi(x) d\mu^h(x) - \int_{E^d} \phi(x) \rho_N^h(x) dx \right| \leq C(N, \phi) h^{N+1}, \tag{49}$$

where $C(N, \phi)$ *is independent of* h.

We have seen in the beginning of this section if in addition to have weak order p the numerical integrator also satisfies $A_p^* \rho_\infty = 0$ then it can achieve weak order $p+1$ with respect to the invariant measure. We now assume that a numerical integrator has an expansion (24) with differential operators A_j satisfying

$$A_j^* \rho_\infty = 0, \quad \text{for} \quad j = 1, \cdots r-1. \tag{50}$$

If (50) holds, then using Lemma 2 we can show that the numerical method has order r with respect to the invariant measure. For example assume $A_1^* \rho_\infty = 0$ then by Lemma 2 $\mathcal{L}^* \rho_1 = -L_1^* \rho_\infty = (A_1^* - \frac{1}{2}(\mathcal{L}^*)^2)\rho_\infty = 0$ and using (48) with $N = 1$ we obtain $\rho_1^h(x) = \rho_\infty(x) + \mathcal{O}(h^2)$ and using (49) we see that we obtain a numerical method of order 2 for the invariant measure. By induction, we have the following theorem [5].

Theorem 2. *Suppose that the SDE* (1) *satisfies Assumptions 3.1 or 3.2. Consider an ergodic numerical method satisfying assumptions* (24) *and* (25).

Assume that (50) *holds. Then the numerical integrator has (at least) order* r *in* (8) *for the invariant measure, i.e., for all* $\phi \in \mathcal{V}^\infty(E,\mathbb{R})$

$$\begin{aligned} e(\phi,h) &= h^r \int_0^\infty \int_{\mathbb{T}^d} A_r u(x,t)\rho_\infty(x)dxdt + \mathcal{O}(h^{r+1}) \\ &= -h^r \int_0^\infty \int_{\mathbb{T}^d} u(x,t)A_r^*\rho_\infty(x)dxdt + \mathcal{O}(h^{r+1}), \end{aligned}$$

where $u(x,t)$ *solves the backward Kolmogorov equation* (21).

Construction of high order numerical approximation of invariant measure

Following Theorem 2, the task is now to construct a numerical method (24) such that (50) holds. Of course a sufficient condition to fulfill (50) is by choosing a method of weak order r. But this is not necessary as we will show here for a class of SDEs with invariant measures of the form

$$\rho_\infty(x) = Ze^{-V(x)} \tag{51}$$

where $Z = (\int_{E^d} e^{-V(x)}dx)^{-1}$ is a normalization constant, and V is a smooth function of class C^∞. The construction is based on modified equations.

Theorem 3. *Consider an ergodic system of SDEs* (1) *with an invariant measure of the form* (51) *satisfying Assumptions 3.1 or* (3.2). *Consider a numerical method satisfying Assumptions* (24)*and* (25) *of order* p *for the invariant measure. Then, for all fixed* $m \geq 1$*, there exists a modified SDE of the form*

$$dX = (f + h^p f_p + \ldots + h^{p+m-1} f_{p+m-1})dt + gdW \tag{52}$$

such that the numerical method applied to this modified SDE satisfies

$$A_j^*(f + h^p f_p + \ldots + h^{p+m-1} f_{p+m-1}, g)\rho_\infty = 0 \quad j = p, \ldots, p+m-1. \tag{53}$$

Furthermore, if the numerical method applied to the modified SDE is ergodic, then it yields a method of order (at least) $r = p + m$ *in* (8) *for the invariant measure of* (1).

We observe in the above theorem, that a modified equation (32) involving only the drift term (33) is used. This theorem can be proved by induction. We sketch the ideas of the construction of the modified equation and refer to [6] for details. Assume that $f_j, j < k < p + m$ have been constructed and consider the scheme obtained by applying the numerical method to the modified SDE

$$dX = (f + \ldots + h^{k-1} f_{k-1})dt + gdW$$

so that the numerical integrator (24) applied to this modified SDE satisfies (53) for $j < k$. Using integration by part and the form of the differential

operator A_j it can be shown that

$$\int_{E^d} (A_k\phi)\rho_\infty dx = \int_{E^d} (\widetilde{A}_k\phi)\rho_\infty dx, \quad \text{for all } \phi \in \mathcal{V}^\infty(E,\mathbb{R}), \tag{54}$$

where $A_k = (f + \ldots + h^{k-1}f_{k-1}, g)$, $\widetilde{A}_k$ is of the form $\widetilde{A}_k = -F \cdot \nabla$, for a certain F. From (54), we deduce

$$A_k^*(f + \ldots + h^{k-1}f_{k-1}, g)\rho_\infty = \mathrm{div}(F\rho_\infty). \tag{55}$$

We set $f_k := F$. Since $A_0 = \mathcal{L}$ we have

$$A_0^*(f+\ldots+h^{k-1}f_{k-1}+h^k f_k, g)\phi = A_0^*(f+\ldots+h^{k-1}f_{k-1}, g)\phi - h^k \mathrm{div}(f_k\phi),$$

using (24), (25) and (55) we obtain

$$\begin{aligned} &A_k^*(f + \ldots + h^{k-1}f_{k-1} + h^k f_k, g)\rho_\infty \\ &= A_k^*(f + \ldots + h^{k-1}f_{k-1}, g)\rho_\infty - \mathrm{div}(f_k\rho_\infty) = 0. \end{aligned}$$

Together with an induction argument, this shows (53) and using Theorem 2, we conclude that the scheme applied to the modified SDE (52) has order $p + m$ for the invariant measure.

4 Construction of high order numerical methods for ergodic dynamical systems

In this section we discuss the actual construction of high order numerical methods for ergodic dynamical systems. We will focus on special yet important classes of SDEs, namely Brownian dynamics that describe the motion of a particle in a potential subject to thermal noise and Langevin dynamics that models the motion of a particle in a potential subject to linear friction and molecular diffusion [34]. In both cases, we will use $E = \mathbb{R}^d$. We mention the recent work [40] that introduces post-processing techniques for SDEs combined with ideas of modified equations [5] that also allows to construct higher order methods for ergodic SDEs (see [9] for an extension to SPDEs).

4.1 Brownian dynamics

The Brownian dynamics is described by the following SDE

$$dX(t) = -\nabla V(X(t))dt + \sigma dW(t), \tag{56}$$

where $V : \mathbb{R}^d \to \mathbb{R}$ is a smooth potential, $\sigma > 0$ is a constant, and $W = (W_1, \ldots, W_d)^T$ is a standard d-dimensional Wiener process. The invariant measure of this SDE, assuming ergodicity, is given by the Gibbs density function

$$\rho_\infty(x) = Ze^{-2V(x)/\sigma^2}, \tag{57}$$

where Z is a normalization constant. To illustrate the above theory, we first consider the special case $d = m = 1$ and will mention the generalization later. As basic integrator, we consider the θ-method

$$X_{n+1} = X_n + (1-\theta)hf(X_n) + \theta hf(X_{n+1}) + \sqrt{h}\sigma\xi_n, \tag{58}$$

where $\xi_n \sim \mathcal{N}(0,1)$ are independent dimensional Gaussian random variables. This method as weak order 1 when $\theta \neq 1/2$ and weak order two for $\theta = 1/2$. In particular for $\theta = 0$ it collapse to the well-known Euler-Maruyama (EM) method. Note that for the SDE (56), $f = -V'(x)$. We now construct a method or order two for the invariant measure that is only of order weak order one in general. For the method (58), a direct calculation reveal that $A_0 = \mathcal{L}$ (due the weak order one) and

$$A_1\phi = \frac{1}{2}f^2\phi'' + \frac{\sigma^2}{2}f\phi''' + \frac{\sigma^4}{8}\phi^{(4)} + \theta(f'f\phi' + \frac{\sigma^2}{2}f''\phi' + \sigma^2 f'\phi'').$$

Integration by part in (54) shows that

$$\widetilde{A}_1\phi = (-(1-2\theta)\Big(\frac{1}{2}f'f + \frac{\sigma^2}{4}f''\Big)\phi'$$

hence according to Theorem 3 defining $f_1 = -(1-2\theta)\Big(\frac{1}{2}f'f + \frac{\sigma^2}{4}f''\big)$ and applying (58) to the modified equation $dX = (f + hf_1)dt + \sigma dW$ will produce a second order method for the approximation of the invariant measure of (56). We note that for linear one-dimensional problems (56) the method (58) with $\theta = 1/2$ samples exactly the invariant measure (see [26]), hence it is not surprising that $f_1 = 0$ in that case.

For the multi-dimensional case (56), we can go through the same derivation by setting this time $f = -\nabla V(x)$. One obtains $f_1 = -(1-2\theta)\Big(\frac{1}{2}f'f + \frac{\sigma^2}{4}\Delta f\big)$ and the scheme

$$X_{n+1} = X_n + (1-\theta)h(f + hf_1)(X_n) + \theta h(f + hf_1)(X_{n+1}) + \sqrt{h}\sigma\xi_n, \tag{59}$$

with $\xi_{n,i} \simeq \mathcal{N}(0,1), i = 1, 2, \ldots, d$ will be of second order for the invariant measure of (56).

Removing derivatives in integrators based on modified equations: Runge-Kutta formulation

In general numerical integrator based on modified equations need to evaluate derivatives of drift or diffusion functions. Sometimes such derivatives are cheap to compute [11]. When such a computation is not convenient, the derivatives appearing in these integrator can also be aproximated by "finite differences" introducing internal stages. For example consider the modified Euler-Maruyama scheme obtained from the above modified θ method with $\theta = 0$ (multi-dimensional case)

$$X_{n+1} = X_n + h\left(f(X_n) - h\frac{1}{2}f'f(X_n) + \frac{\sigma^2}{4}\Delta f(X_n)\right) + \sqrt{h}\sigma\xi_n. \qquad (60)$$

One can check that the derrivative free version

$$\begin{aligned} Y_1 &= X_n + \sqrt{2}\sigma\sqrt{h}\xi_n \\ Y_2 &= X_n - \frac{3}{8}hf(Y_1) + \frac{\sqrt{2}}{4}\sigma\sqrt{h}\xi_n \\ X_{n+1} &= X_n - \frac{1}{3}hf(Y_1) + \frac{4}{3}hf(Y_2) + \sigma\sqrt{h}\xi_n, \end{aligned} \qquad (61)$$

$\xi_{n,i} \simeq \mathcal{N}(0,1), i = 1, 2, \ldots, d$ is also of weak order 1, while approximating with second order the invariant measure of (56). This can be seen by checking that the same operators A_0, A_1 appear in the expansion (24) of both schemes.

We close this section by a numerical experiments with the above derivative free second order method for the invariant measure. We consider a Brownian dynamics (56) with a two dimensional quartic potential $V(x) = (1 - x_1^2)^2 + (1 - x_2^2)^2 + \frac{x_1 x_2}{2} + \frac{x_2}{5}$. We emphasise that doing so we depart from the case of Lipschitz vector fields for which our theory apply. We will see numerically that we still get the right order of the invariant measure under these weaker assumptions. The Gibbs invariant density function is depicted in Figure (2) (left picture). We consider the Euler-Maruyama method ((58) with $\theta = 0$) and the second order modified Euler Maruyama method (61). We see in Figure (2) (right picture) that the modified method captures the invariant measure with the rate predicted by Theorem 3.

4.2 Langevin dynamics

We consider here the Langevin equation given by a second order stochastic differential equation of the form

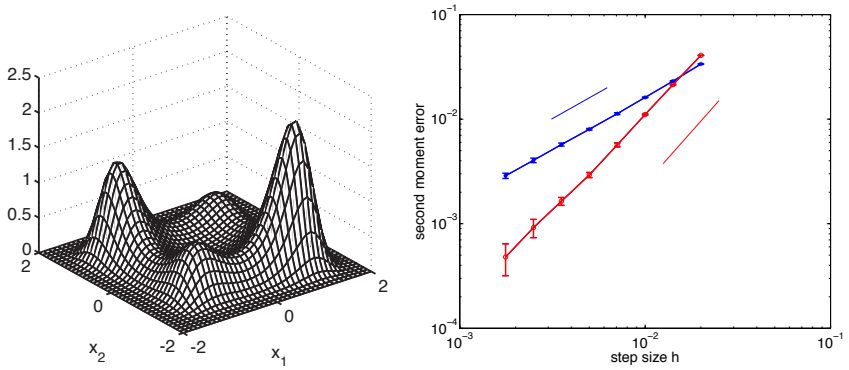

Fig. 1 Figures adapted from [5]. Left picture: Brownian dynamics with quartic potential. Gibbs density function. Right picture Error $e(\phi, h)$ defined in (8) for $\phi(x) = x_1^2 + x_2^2$ for the EM method (blue line with star symbol) and the modified EM method (red line with circle symbol). Standard deviation is depicted by the vertical line obtained from the Monte-Carlo error originating from 10 trajectories. The straight lines without symbols represent first and second order slopes.

$$dq(t) = M^{-1}p(t)dt, \tag{62a}$$
$$dp(t) = (-\nabla V(q(t)) - \gamma p(t))dt + \sqrt{2\beta^{-1}\gamma}M^{1/2}dW(t), \tag{62b}$$

where $p(t), q(t) \in \mathbb{R}^d$, $W(t)$ denotes a standard d-dimensional Wiener process, $V : \mathbb{R}^d \to \mathbb{R}$ is a potential, M is a symmetric positive definite mass matrix (taken as the identity in what follows for simplicity), and the positive scalar parameter γ, β are related respectively to friction and temperature. The dynamcis generated by (62) is ergodic provided suitable smoothness and growth assumptions on the Hamiltonian energy (see e.g. [28, 21]),

$$H(p, q) = \frac{1}{2}p^T p + V(q), \tag{63}$$

with invariant measure given by the Gibbs density function (see e.g., (57))

$$\rho_\infty(p, q) = Ze^{-\beta H(p,q)}, \tag{64}$$

where Z is the normalization constant. Efficient numerical integrators for Langevin equations are based on the Lie-Trotter splitting [7, 23, 25, 6]

$$X_{n+1} = \Phi_h \circ \Theta_{h,n}(X_n), \tag{65}$$

where $X_n = (p_n, q_n)^T$. The integrator Φ_h approximates the exact flow of the deterministic Hamiltonian part

$$dq(t) = p(t)dt, \qquad dp(t) = -\nabla V(q(t))dt, \tag{66}$$

while $\Theta_{h,n}$ is an integrator for the stochastic part given by

$$q_{n+1} = q_n, \qquad p_{n+1} = e^{-\gamma h} p_n + \sqrt{\beta^{-1}(1 - e^{-2\gamma h})}\xi_n,$$

where $\xi_n \sim \mathcal{N}(0, I)$ are independent d-dimensional Gaussian random variables. This integrator has the same law of probability as the exact solution of the stochastic part that is given by the variation of constants formula,

$$q(t_{n+1}) = q(t_n), \qquad p(t_{n+1}) = e^{-\gamma h} p(t_n) + \sqrt{2\beta^{-1}\gamma} \int_{t_n}^{t_{n+1}} e^{-\gamma(h-s)} dW(s).$$

For simplicity we assume that the potential V is a $\mathcal{C}^\infty$ function where ∇V has bounded derivatives of any order and satisfies standard growth condition $q^T \nabla V(q) \geq C_1 q^T q - C_2$, for all $q \in \mathbb{R}^d$ for $C_1, C_2 > 0$, which guaranties that (62) is ergodic and we assume that numerical flow Φ_h is globally Lipschitz in $\mathbb{R}^{2d}$. In the non-globally Lipschitz case, one could resort to implicit deterministic integrator to avoid exploding trajectories as studied in [21] or apply the theory presented below for explicit integrators while rejecting exploding trajectories. This latter procedure that has been rigorously analyzed in [31] and applied to ergodic Langevin problems in [32].

Consider the numerical integrator $X_{n+1} = \Phi_h \circ \Theta_{h,n}(X_n)$ with an expansion (24). According to Theorem 2, if (50) holds, then the numerical integrator will approximate the invariant measure of (62) with order r. To construct accurate integrator with respect to the invariant measure we will as in Section 3.2 consider a modified equation. As here the only integrator to be considered is a deterministic integrator Φ_h, we consider the modified deterministic ODE (13) $s = r$, where $f(t) = (p(t), -\nabla V(q(t)))^T$. The question now is how the condition $A_j^* \rho_\infty = 0$ of Theorem 2 translate into a condition on the $\hat{f}_j$ in (19) ?

The connection can be revealed by using the semi-group property of the Markov process

$$\mathbb{E}(\phi(X_1)|X_0 = x) = \mathbb{E}(\phi(\Phi_h \circ \Theta_{h,n})(X_0)|X_0 = x) = e^{h\mathcal{L}_S}(\phi \circ \Phi_h)(x), \quad (67)$$

for a smooth test function ϕ and $x \in \mathbb{R}^{2d}$, where $e^{h\mathcal{L}_S}\phi$ denotes the exact flow of the Kolmogorov backward equation corresponding to the stochastic part of (62) with generator $\mathcal{L}_S$ given by

$$\mathcal{L}_S := -\gamma p \cdot \nabla_p + \beta^{-1}\gamma \Delta_p.$$

We then have in view of (16) with $M = r$

$$\mathbb{E}(\phi(X_1)|X_0=x) = \left(\sum_{k=0}^{r}\frac{h^k\mathcal{L}_S^k}{k!}\right)\left(\sum_{k=0}^{r}\frac{h^k\widetilde{\mathcal{L}}_D^k}{k!}\right)\phi(x)+\mathcal{O}(h^{r+1})$$
$$= \phi(x)+h\mathcal{L}\phi(x)+\sum_{k=1}^{r}h^{k+1}A_k\phi(x)+\mathcal{O}(h^{r+1}),$$

where we recall that $\widetilde{\mathcal{L}}_D = \tilde{f}_{h,r}\cdot\nabla := F_0 + hF_1 + \cdots + F_r$, where $F_j = f_j\cdot\nabla,\ j=1,2,\dots r$ and $f_0=f$. Developing the sums in the above equality and identyfying power of h allows to find an expression for A_k in terms of power of $\mathcal{L}_S$ and product of operators F_j. It is then deduced that if

$$\mathrm{div}(f_j\rho_\infty)=0, \qquad j=1,\dots,r-1, \tag{68}$$

then $A_j^*\rho_\infty = 0,\ j=1,\cdots,r-1$ and $A_r^*\rho_\infty = \mathrm{div}(f_r\rho_\infty)$. Using Theorem 2 we find

Theorem 4. *Assume that Φ_h is a consistent method for* (66) *with Lipschitz continuous flow. If the vector fields in* (13) *($s=r$) satisfy* (68) *with $r\geq 1$, then assuming ergodicity, the Lie-Trotter splitting* (65) *has order r of accuracy for the invariant measure of* (62)*, i.e.,*

$$e(\phi,h) = -h^r\int_0^\infty\int_{\mathbb{R}^d\times\mathbb{R}^d} u(p,q,t)\mathrm{div}(f_r(p,q)\rho_\infty(p,q))dpdqdt+\mathcal{O}(h^{r+1}), \tag{69}$$

for all $\phi\in\mathcal{C}_P^\infty(\mathbb{R}^d,\mathbb{R})$ and $h\to 0$, where $e(\phi,h)$ is defined in (8) *and $u(x,t)$ is the solution of the Backward Kolmogorov equation* (21) *for* (62)*.*

The condition (68) can be rewritten

$$\mathrm{div}(f_j\rho_\infty) = \big(\mathrm{div}(f_j)-\beta f_j\cdot\nabla H\big)\rho_\infty.$$

Observe that $\mathrm{div}(f_j)=0$ for all $1\leq j\leq r-1$ is equivalent to the fact that the deterministic integrator Φ_h is volume preservation up to to order r, i.e., $\det(\partial\Phi_h(y)/\partial y) = 1+\mathcal{O}(h^{r+1})$. Also $f_j\cdot\nabla H = 0$ for all $1\leq j\leq r-1$ is equivalent to the fact that the deterministic integrator Φ_h is energy preserving up to order r, i.e., $H(\Phi_h(y)) = H(y)+\mathcal{O}(h^{r+1})$. Notice that any deterministic method of order r will fulfil both conditions and hence will produce a Lie-Trotter splitting method of order r for the invariant measure of (62) according to Theorem 4. In [7] it was shown that sufficient conditions to preserve the invariant measure of (62) up to order r is to consider a symplectic integrator in the Lie-Trotter splitting preserving the energy with order r. The condition (68), first given in [6], is thus a weaker characterization of high order Lie-Trotter splitting for Langevin dynamics. We also mention the work [23, 25] where efficient non Markovian schemes with second order accuracy for the invariant measure of (62) have been constructed.

As an illustration of the above theory we consider three different deterministic numerical integrators namely, the explicit Euler method,

$$p_{n+1} = p_n - h\nabla V(q_n), \qquad q_{n+1} = q_n + hp_n,$$

the symplectic Euler method,

$$p_{n+1} = p_n - h\nabla V(q_n), \qquad q_{n+1} = q_n + hp_{n+1},$$

and the Heun method, a second order explicit method given by

$$p_{n+1} = p_n - h\nabla V\left(q_n + \frac{h}{2}p_n\right), \qquad q_{n+1} = q_n + h\left(p_n - \frac{h}{2}\nabla V(q_n)\right).$$

Departing from the situation of globally Lipschitz vector fields (for which the theory has been derived) we consider the Langevin dynamics (62) with a quartic potential $V(q) = (1-q^2)^2 - \frac{1}{2}q$. As for the previous case of Brownian dynamics we observe numerically that the theory still apply and Figure 2 corroborate the results of Theorem 4.

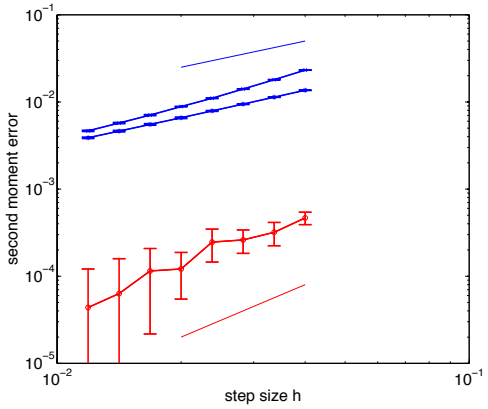

Fig. 2 Figures adapted from [6]. Langevin dynamics with quartic potential : error $e(\phi, h)$ defined in (8) for $\phi(p,q) = p^2 + q^2$ and with deterministic flow in the Lie-Trotter splitting given by the EM method (blue line with cross symbol), the symplectic Euler method (blue line with cross symbol) and the Heun method (red line with circle symbol). Standard deviation is depicted by the vertical line obtained from the Monte-Carlo error originating from 10 trajectories. The straight lines without symbols represent first and second order slopes.

References

1. A. Abdulle, D. Cohen, G. Vilmart, and K. C. Zygalakis. High order weak methods for stochastic differential equations based on modified equations. *SIAM J. Sci. Comput.*, 34(3):1800–1823, 2012.

2. A. Abdulle, W. E, B. Engquist, and E. Vanden-Eijnden. The heterogeneous multiscale method. *Acta Numer.*, 21:1–87, 2012.
3. A. Abdulle and G. Pavliotis. Numerical methods for stochastic partial differential equations with multiple scales. *J. Comput. Phys.*, 231(6):2482–2497, 2012.
4. A. Abdulle, G. Pavliotis, and U. Vaes. Spectral methods for multiscale stochastic differential equations. *Preprint, submitted for publication*, 2016.
5. A. Abdulle, G. Vilmart, and K. C. Zygalakis. High order numerical approximation of the invariant measure of ergodic SDEs. *SIAM J. Numer. Anal.*, 52(4):1600–1622, 2014.
6. A. Abdulle, G. Vilmart, and K. C. Zygalakis. Long time accuracy of Lie-Trotter splitting methods for Langevin dynamics. *SIAM J. Numer. Anal.*, 53(1):1–16, 2015.
7. N. Bou-Rabee and H. Owhadi. Long-run accuracy of variational integrators in the stochastic context. *SIAM Journal on Numerical Analysis*, 48(1):278–297, 2010.
8. C.-E. Bréhier. Analysis of an HMM time-discretization scheme for a system of stochastic PDEs. *SIAM J. Numer. Anal.*, 51(2):1185–1210, 2013.
9. C.-E. Bréhier and G. Vilmart. High-order integrator for sampling the invariant distribution of a class of parabolic stochastic pdes with additive space-time noise. *SIAM J. Sci. Comput.*, 38:A2283–A2306, 2016.
10. Y. Cao, D. T. Gillespie, and L. Petzold. The slow scale stochastic simulation algorithm. *J. Chem. Phys.*, 122:014116–014116–17, 2005.
11. P. Chartier, E. Hairer, and G. Vilmart. Numerical integrators based on modified differential equations. *Math. Comp.*, 76(260):1941–1953 (electronic), 2007.
12. A. Debussche and E. Faou. Weak backward error analysis for SDEs. *SIAM J. Numer. Anal.*, 50(3):1735–1752, 2012.
13. W. E, D. Liu, and E. Vanden-Eijnden. Analysis of multiscale methods for stochastic differential equations. *Comm. Pure Appl. Math.*, 58(11):1544–1585, 2005.
14. W. E, D. Liu, and E. Vanden-Eijnden. Nested stochastic simulation algorithms for chemical kinetic systems with multiple time scales. *J. Comput. Phys.*, 221(1):158–180, 2007.
15. C. W. Gardiner. *Handbook of stochastic methods*. Springer-Verlag, Berlin, second edition, 1985. For physics, chemistry and the natural sciences.
16. D. Gillespie. Stochastic simulation of chemical kinetics. *Annu. Rev. Phys. Chem.*, 58, 2007.
17. C. Graham and D. Talay. *Stochastic simulation and Monte Carlo methods*, volume 68 of *Stochastic Modelling and Applied Probability*. Springer, Heidelberg, 2013. Mathematical foundations of stochastic simulation.
18. E. Hairer, C. Lubich, and G. Wanner. *Geometric Numerical Integration. Structure-Preserving Algorithms for Ordinary Differential Equations*. Springer Series in Computational Mathematics 31. Springer-Verlag, Berlin, second edition, 2006.
19. D. J. Higham. A-stability and stochastic stability mean-square stability. *BIT*, 40:404–409, 2000.
20. P. Kloeden and E. Platen. *Numerical solution of stochastic differential equations*. Springer-Verlag, Berlin and New York, 1992.
21. M. Kopec. Weak backward error analysis for Langevin process. *BIT*, 55(4):1057–1103, 2015.
22. M. Kopec. Weak backward error analysis for overdamped Langevin processes. *IMA J. Numer. Anal.*, 35(2):583–614, 2015.
23. B. Leimkuhler and C. Matthews. Rational construction of stochastic numerical methods for molecular sampling. *Appl. Math. Res. Express*, pages 34–56, 2013.
24. B. Leimkuhler and C. Matthews. *Molecular dynamics*, volume 39 of *Interdisciplinary Applied Mathematics*. Springer, Cham, 2015. With deterministic and stochastic numerical methods.
25. B. Leimkuhler, C. Matthews, and G. Stoltz. The computation of averages from equilibrium and nonequilibrium Langevin molecular dynamics. *IMA J. Numer. Anal.*, 36(1):13–79, 2016.

26. T. Li, A. Abdulle, and W. E. Effectiveness of implicit methods for stiff stochastic differential equations. *Commun. Comput. Phys.*, 3(2):295–307, 2008.
27. 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.
28. J. C. Mattingly, A. M. Stuart, and M. V. Tretyakov. Convergence of numerical time-averaging and stationary measures via poisson equations. *SIAM Journal on Numerical Analysis*, 48(2):552–577, 2010.
29. G. Milstein. Weak approximation of solutions of systems of stochastic differential equations. *Theory Probab. Appl.*, 30(4):750–766, 1986.
30. G. Milstein and M. Tretyakov. *Stochastic numerics for mathematical physics*. Scientific Computing. Springer-Verlag, Berlin and New York, 2004.
31. G. N. Milstein and M. V. Tretyakov. Numerical integration of stochastic differential equations with nonglobally Lipschitz coefficients. *SIAM J. Numer. Anal.*, 43(3):1139–1154 (electronic), 2005.
32. G. N. Milstein and M. V. Tretyakov. Computing ergodic limits for Langevin equations. *Phys. D*, 229(1):81–95, 2007.
33. G. Pavliotis and A. Stuart. *Multiscale Methods: Averaging and Homogenization*, volume 53 of *Text in Applied Mathematics*. Springer-Verlag, New York, 2008.
34. H. Risken. *The Fokker-Planck equation*, volume 18 of *Springer Series in Synergetics*. Springer-Verlag, Berlin, 1989.
35. G. O. Roberts and R. L. Tweedie. Exponential convergence of Langevin distributions and their discrete approximations. 2(4):pp. 341–363, 1996.
36. A. Rößler. Second order Runge-Kutta methods for Itô stochastic differential equations. *SIAM J. Numer. Anal.*, 47(3):1713–1738, 2009.
37. D. Talay. Efficient numerical schemes for the approximation of expectations of functionals of the solution of a SDE and applications. *Lecture Notes in Control and Inform. Sci., Springer*, 61:294–313, 1984.
38. D. Talay. Second order discretization schemes of stochastic differential systems for the com- putation of the invariant law. *Stochastics Stochastics Rep.*, 29(1):13–36, 1990.
39. E. Vanden-Eijnden. Numerical techniques for multiscale dynamical system with stochastic effects. *Commun. Math. Sci.*, 1:385–391, 2003.
40. G. Vilmart. Postprocessed integrators for the high order integration of ergodic SDEs. *SIAM J. Sci. Comput.*, 37(1):A201–A220, 2015.
41. K. C. Zygalakis. On the existence and the applications of modified equations for stochastic differential equations. *SIAM J. Sci. Comput.*, 33(1):102–130, 2011.