High order numerical approximation of the invariant measure of ergodic SDEs

Assyr Abdulle[1], Gilles Vilmart[2], and Konstantinos C. Zygalakis[3]

Abstract

We introduce new sufficient conditions for a numerical method to approximate with high order of accuracy the invariant measure of an ergodic system of stochastic differential equations, independently of the weak order of accuracy of the method. We then present a systematic procedure based on the framework of modified differential equations for the construction of stochastic integrators that capture the invariant measure of a wide class of ergodic SDEs (Brownian and Langevin dynamics) with an accuracy independent of the weak order of the underlying method. Numerical experiments confirm our theoretical findings.

Keywords: stochastic differential equations, invariant measure, weak convergence, modified differential equations, backward error analysis, ergodicity.

AMS subject classification (2010): 65C30, 60H35, 37M25

1 Introduction

We consider a system of (Itô) stochastic differential equations (SDEs)

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

where $X(t)$ is the solution in the space E, $X_0 \in E$ is the initial condition, $f : E \mapsto E, g : E \mapsto E^m$, and $W(t)$ is a standard m-dimensional Brownian motion. The space E denotes either $E = \mathbb{R}^d$ or the torus $E = \mathbb{T}^d$, and this is specified when needed. With the exception of some special cases, the solutions to (1) are not explicitly known, and numerical methods are needed. We consider a one step numerical integrator for the approximation of (1) at time $t = nh$ of the form

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

where h denotes the stepsize and ξ_n is a random vector. The choice behind the numerical method used to approximate (1), depends crucially on the type of the approximation that one wants to achieve. In particular, for the approximation of individual trajectories one is interested in the strong convergence properties of the numerical method, while for the approximation of the expectation of functionals of the solution, one is interested

[1]Mathematics Section, École Polytechnique Fédérale de Lausanne, Station 8, 1015 Lausanne, Switzerland, Assyr.Abdulle@epfl.ch

[2]École Normale Supérieure de Cachan, Antenne de Bretagne, INRIA Rennes, IRMAR, CNRS, UEB, av. Robert Schuman, F-35170 Bruz, France, Gilles.Vilmart@bretagne.ens-cachan.fr Present address: Université de Genève, Section de mathématiques, 2-4 rue du Lièvre, CP 64, 1211 Genève 4

[3]Mathematical Sciences, University of Southampton, Southampton SO17 1BJ, UK, k.zygalakis@soton.ac.uk

in its weak convergence properties. The numerical approximation (2), starting from the initial condition X_0 of (1) is said to have local weak order p if for all functions[1] $\phi :\in C_P^{2(p+1)}(\mathbb{R}^d, \mathbb{R})$,

$$|\mathbb{E}(\phi(X_1)) - \mathbb{E}(\phi(X(h)))| \leq C(X_0)h^{p+1}, \tag{3}$$

for all h sufficiently small, where $C(X_0)$ is independent of h. Under appropriate conditions one can infer "a global weak order p" from the local weak error [15] (see [16, Chap. 2.2]). This results will be briefly discussed in Theorem 2.6 (we need a slight reformulation of the existing results).

Strong and weak types of convergence relate to the finite time properties of (1) and its numerical approximations. We say that the process $X(t)$ is ergodic if it has a unique invariant measure μ satisfying for each $\mu-$integrable function ϕ and for any deterministic initial condition $X_0 = x$,

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T \phi(X(s))ds = \int_E \phi(y)d\mu(y), \qquad \text{almost surely.} \tag{4}$$

Before considering the different sources of error, one needs to make sure that the numerical approximation is itself ergodic. In particular, the case where the coefficients are not globally Lipschitz is particularly challenging and it is still an active research area [17, 13, 19, 20, 22, 9]. This important question is however not the focus of the present paper as we will rather assume ergodicity of the numerical method. We recall that the numerical method (2) is called ergodic if it has a unique invariant probability law μ^h with finite moments of any order and

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \phi(X_n) = \int_E \phi(y)d\mu^h(y), \qquad \text{almost surely,} \tag{5}$$

for all deterministic initial condition $X_0 = x$ and all μ^h-integrable functions ϕ.

We will say that the numerical method (2) has order $r \geq 1$ with respect to the invariant measure if

$$e(\phi, h) := \left| \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \phi(X_n) - \int_E \phi(y)d\mu(y) \right| \leq Ch^r. \tag{6}$$

In the sequel, we will assume that the ergodic measure μ has a density function ρ_∞. The study of the error $e(\phi, h)$ in approximating the invariant measure, its relation with the weak error and the construction of numerical method with high order of convergence with respect to the invariant measure is the main focus of our paper. We mention that various papers related to the study of $e(\phi, h)$ appeared in the literature. In [21] an error estimate for $e(\phi, h)$ has been established for a variety of different numerical methods. In addition, in [23] with a use of a global weak error expansion, an expansion of (6) in powers of h was derived for Euler-Maruyama and the Milstein methods. This allowed the use of extrapolation techniques to further reduce the bias in the calculation of the error $e(\phi, h)$ between the numerical time average and its true value.

The error $e(\phi, h)$ was also the subject of study of [14]. Given an ergodic integrator of weak order p for an ergodic SDE (1), it is shown that it has order $r \geq p$ for the

[1] Here and in what follows, $C_P^\ell(\mathbb{R}^d, \mathbb{R})$ denotes the space of ℓ times continuously differentiable functions $\mathbb{R}^d \to \mathbb{R}$ with all partial derivatives with polynomial growth.

invariant measure (6). In [12] an example of integrator with $r > p$ is given: for the so-called stochastic θ-method with $\theta = 1/2$ applied to the Orstein-Uhlenbeck process, we have $e(\phi, h) = 0$ despite the weak order two of the method. Related works where such a mismatch is mentioned are [3, 2, 10].

In this paper, we present two results for the numerical approximation of ergodic nonlinear systems of SDEs. Firstly, we derive new sufficient conditions for an ergodic integrator to have high order (6) for the invariant measure, possibly larger than its weak order of accuracy (3). A crucial ingredient is a new expansion of the error $e(\phi, h)$ based on the work [23], and the analysis in [4] of numerical invariant measures. Secondly, we introduce a systematic procedure to design high order integrators for the invariant measure based on modified differential equations for SDEs proposed in [1]. Our new methodology is based on modified differential equations, which is a fundamental tool for the study of geometric integrators for ODEs [6, 11]. It was recently extended to SDEs in [24, 4] for the backward error analysis of stochastic integrators and in [1] for the construction of high weak order integrators.

The paper is organized as follows. In Section 2, we present the framework and derive a new expansion of the error $e(\phi, h)$. In Section 3, we derive our main results: sufficient order conditions for the invariant measure of an ergodic integrator and a construction procedure of high order integrators based on modified differential equations. In Section 4, we apply our methodology and construct a range of new integrators based on the stochastic θ-method for Brownian dynamics. Finally in Section 5, we present various numerical investigations, that illustrate the behaviour of our new integrators and corroborate the claimed orders of convergence.

2 Preliminaries

In Section 2.1, we describe some preliminary results related to ergodicity of SDEs and their numerical approximations, using the standard framework of the Fokker-Planck and backward Kolmogorov equations. In Section 2.2, We discuss a global error expansion for both the weak error and the error with respect to the invariant measure.

2.1 Exact and numerical invariant measure for ergodic SDEs

Let us denote by $\rho(x,t)$ the probability density of $X(t)$ defined by (1) with initial condition $X_0 = x$. Then we have

$$\mathbb{E}\left(\phi(X(t))|X_0 = x\right) = \int_E \phi(y)\rho(y,t)dy, \tag{7}$$

where $\rho(y,t)$ is the solution of the Fokker-Planck equation

$$\begin{aligned} \frac{\partial \rho}{\partial t} &= \mathcal{L}^*\rho, && \text{(8a)} \\ \rho(y,0) &= \rho_0(y), && \text{(8b)} \end{aligned}$$

here $\rho_0(y) = \delta(y-x)$ (Dirac mass) for the deterministic initial condition of $X(t)$ and $\mathcal{L}^*$ is given by

$$\mathcal{L}^*\rho = -\nabla_y \cdot (f(y)\rho) + \frac{1}{2}\nabla_y \cdot \nabla_y \cdot (g^T(y)g(y)\rho). \tag{9}$$

This operator is the L^2-adjoint of

$$\mathcal{L} := f(x)\cdot\nabla_x + \frac{1}{2}g(x)g^T(x) : \nabla_x\nabla_x, \tag{10}$$

the generator of the Markov process $X(t)$ defined by (1). Recall that if ρ_∞ is the density of the invariant measure of (1) (assuming ergodicity), then ρ_∞ is the unique stationary solution of (8) and thus satisfies

$$\mathcal{L}^*\rho_\infty = 0. \tag{11}$$

Next, we consider

$$u(x,t) = \mathbb{E}\left(\phi(X(t))|X_0 = x\right), \tag{12}$$

where $X(t)$ is the solution of (1). We note that $u(x,t)$ is the solution of the backward Kolmogorov equation

$$\frac{\partial u}{\partial t} = \mathcal{L}u, \tag{13a}$$

$$u(x,0) = \phi(x). \tag{13b}$$

A formal Taylor series expansion in terms of the generator operator $\mathcal{L}$ of the Markov process is derived in [24] for u and a rigorous finite term expansion is proposed in [4] namely

$$u(x,h) - \phi(x) = \sum_{j=1}^{l}\frac{h^j}{j!}\mathcal{L}^j\phi(x) + h^{l+1}r_l(f,g,\phi)(x), \tag{14}$$

for all positive integer l, with a bound of the form $|r_l(f,g,\phi)(x)| \leq c_l(1+|x|^{\kappa_l})$.

Remark 2.1. *One way to turn* $u(x,h) = \phi(x) + h\mathcal{L}\phi + \frac{h^2}{2}\mathcal{L}^2\phi + \cdots$ *into a rigorous expansion* (14) *is to restrict* (1) *to* $E = \mathbb{T}^d$ *as it was done in [4]. Another way is to follow the approach in [23, Lemma 2] and assume that* f, g *are* C^∞ *where derivatives of any order are bounded. This together with the assumption that*

$$|\phi(x)| \leq C(1+|x|^s) \tag{15}$$

for some positive integer s *are enough to prove that the solution* u *of* (13) *has derivatives in space of any order that have a polynomial growth of the form* (15)*, with other constants* C, s *that are independent of* $t \in [0,T]$*, and this implies that* (14) *holds.*

In terms of the numerical solution (2) one can define

$$U(x,h) = \mathbb{E}(\phi(X_1)|X_0 = x)), \tag{16}$$

for the expectation at time h, where again for simplicity one assumes that the initial condition X_0 is deterministic. We make the following regularity and consistency assumption on the integrator, which is easily satisfied by any reasonable numerical method.

Assumption 2.2. *Let* f, g *be* C^∞ *with bounded derivatives of any order. We assume for all deterministic initial conditions* X_0 *that*

$$|\mathbb{E}(X_1 - X_0)| \leq C(1+|X_0|)h, \qquad |X_1 - X_0| \leq M_n(1+|X_0|)\sqrt{h}, \tag{17}$$

where C is independent of h small enough and M_n has bounded moments of all orders independent of h. We assume that (16) *has a weak Taylor series expansion of the form,*

$$U(x,h) = \phi(x) + hA_0(f,g)\phi(x) + h^2 A_1(f,g)\phi(x) + \dots, \tag{18}$$

where $A_i(f,g)$, $i = 0,1,2,\dots$ are linear differential operators with coefficients depending smoothly on the drift and diffusion functions f, g, and their derivatives (and depending on the choice of the integrator). In addition, we assume that $A_0(f,g)$ coincides with the generator $\mathcal{L}$ given in (10)*, which means that the method has (at least) weak order one,*

$$A_0(f,g) = \mathcal{L}. \tag{19}$$

Example 2.3. *Consider the stochastic θ-method [8] for* (1) *where $g = \sigma I$ and $d = m$ (additive noise case) defined as*

$$X_{n+1} = X_n + h(1-\theta)f(X_n) + \theta f(X_{n+1}) + \sigma\sqrt{h}\xi_n. \tag{20}$$

For $\theta = 0$, this scheme coincides with the explicit Euler-Maruyama method while for $\theta \neq 0$ it is implicit, i.e. it requires the resolution of a nonlinear system at each timestep. A straightforward calculation yields that the differential operator A_1 in (18) *is given by*

$$\begin{aligned} \mathcal{A}_1\phi &= \frac{1}{2}\phi''(f,f) + \frac{\sigma^2}{2}\sum_{i=1}^{d}\phi'''(e_i,e_i,f) + \frac{\sigma^4}{8}\sum_{i,j=1}^{d}\phi^{(4)}(e_i,e_i,e_j,e_j) \\ &+ \theta\phi'(f'f + \frac{\sigma^2}{2}\sum_{i=1}^{d} f''(e_i,e_i)) + \frac{\theta\sigma^2}{2}\sum_{i=1}^{d}\phi''(f'e_i,e_i), \end{aligned} \tag{21}$$

where $e_1,\dots,e_d$ denotes the canonical basis of $\mathbb{R}^d$ and $\phi'(\cdot), \phi''(\cdot,\cdot), \phi'''(\cdot,\cdot,\cdot),\dots,$ are the derivatives of ϕ which are linear, symmetric bilinear, trilinear, $\dots$, forms, respectively. In dimension $d=1$, it reduces to $\mathcal{A}_1\phi = \frac{1}{2}f^2\phi'' + \frac{\sigma^2}{2}f\phi''' + \frac{\sigma^4}{8}\phi^{(4)} + \theta(f'f\phi' + \frac{\sigma^2}{2}f''\phi' + \frac{\sigma^2}{2}f'\phi'')$.

Assumption 2.2 immediately implies that we have the rigorous expansion

$$U(x,h) = \phi(x) + \sum_{i=0}^{l} h^{i+1}A_i(f,g)\phi(x) + h^{l+2}R_l(f,g,\phi)(x) \tag{22}$$

for all positive integers l, with a remainder satisfying $|R_l(f,g,\phi)(x)| \leq C_l(1+|x|^{k_l})$. We also deduce that the moments of the numerical solution are uniformly bounded, as stated in the following result, shown in the proof of [16, Lemma 2.2, p. 102]. We observe that if the numerical solution (2) has local weak order p (see (3)) and satisfies Assumption 2.2 then

$$\mathbb{E}(\phi(X(h))) - \mathbb{E}(\phi(X_1)) = h^{p+1}\left(\frac{\mathcal{L}^{p+1}}{(p+1)!} - A_p\right)\phi(X_0) + \mathcal{O}(h^{p+2}). \tag{23}$$

Proposition 2.4. *Assume* (17)*. Then, for all positive integers k, there exist constants C_k, D_k such that*

$$\mathbb{E}\big(|X_n|^k\big) \leq C_k e^{D_k T}, \quad \textit{for all} \quad nh \leq T. \tag{24}$$

2.2 Global error expansion for the weak error

We now study how accurate the numerical invariant measure ρ_∞^h is compared to the true invariant measure ρ_∞. The first step to show this is the establishment of a global error expansion for the weak error

$$E(\phi, h, T) = |\mathbb{E}(\phi(X(T))) - \mathbb{E}(\phi(X_N))|, \tag{25}$$

where X_N denotes the numerical solution at the final time $T = Nh$ calculated with a time step h with a numerical method of weak order p.

In the sequel, we assume that solution $X(t)$ of (1) ergodic. We recall in Remark 2.5 some necessary conditions in order for $X(t)$ to be ergodic.

Remark 2.5. *Let $X(t)$ the solution of* (1)*. The following assumptions implies that* $(X(t))$ *is ergodic (see [7]),*

1. *f, g are of class C^∞, with bounded derivatives of any order, and g is a bounded function;*
2. *The generator $\mathcal{L}$ in* (10) *is a uniformly elliptic operator,* i.e. *there exists $\alpha > 0$ such that $\forall\ x, \xi \in \mathbb{R}^d, \quad x^T g(\xi) g(\xi)^T x \geq \alpha |x|^2$;*
3. *there exists $\beta > 0$ and a compact set K in $\mathbb{R}^d$ such that $\forall\ x \in \mathbb{R}^d - K, \langle x, f(x) \rangle \leq -\beta x^2$.*

Likewise, we assume that the Markov chain defined by our numerical solution is ergodic (see equation (5)). The following theorem combines results derived by Talay and Milstein. Precisely, the expression (26) has been proved in [23] for specific methods (e.g., he Euler-Maruyama or the Milstein methods), while the general procedure to infer the global weak order from the local weak order is due to Milstein [15] (see [16, Chap. 2.2]). The novelty here is the new formulation of the error function (27) in terms of the operator A_i in Assumption 2.2 and generator $\mathcal{L}$ that will be useful for our main results.

Theorem 2.6. *Assume the hypotheses on f, g in Remark 2.5. Let X_N be a numerical solution of* (1) *on* $[0, T]$ *($E = \mathbb{R}^d$) satisfying Assumption 2.2 and the local weak order p estimate* (3) *where $C(x)$ satisfies* (15)*. Then, we have the following expansion of the global error* (25)*, for all $\phi \in C_P^{2p+4}(\mathbb{R}^d, \mathbb{R})$,*

$$E(\phi, h, T) = h^p \int_0^T \mathbb{E}(\psi_e(X(s), s)) ds + \mathcal{O}(h^{p+1}) \tag{26}$$

where $\psi_e(x, t)$ satisfies

$$\psi_e(x, t) = \left(\frac{1}{(p+1)!} \mathcal{L}^{p+1} - A_p \right) v(x, t), \tag{27}$$

with $v(x, t) = \mathbb{E}(\phi(X(T)) | X(t) = x)$ satisfying

$$\frac{\partial v}{\partial t} + \mathcal{L} v = 0, \tag{28a}$$

$$v(x, T) = \phi(x). \tag{28b}$$

Proof. The proof, similar to the one found in [23] and [16, Chap. 2.2], is provided for completeness in the Appendix. □

Using Theorem 2.6 one can obtain a similar expansion to (26) for the difference between the true and the numerical ergodic averages. In particular we have the following theorem.

Theorem 2.7. *Assume that the conditions 1.,2.,3. from Remark 2.5 hold, and let* $\phi : \mathbb{R}^d \to \mathbb{R}$ *a smooth function satisfying* (15)*. Then, if a numerical method of weak order* p *is ergodic, it satisfies*

$$\lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^{N} \phi(X_i) - \int_{\mathbb{R}^d} \phi(y)\rho_\infty(y)dy = -\lambda_p h^p + \mathcal{O}(h^{p+1}) \tag{29}$$

for any deterministic initial condition, with λ_p *defined as*

$$\lambda_p = \int_0^{+\infty} \int_{\mathbb{R}^d} \psi_e(t,y)\rho_\infty(y)dydt \tag{30}$$

with ψ_e *satisfying* (27)*.*

Proof. The proof is similar to the one found in [23, Theorem 4], with the main difference being that now (26) is used as the starting point of the proof instead of the specific formula for the Euler-Maruyama method used in [23]. □

Theorem 2.7 provides an explicit expression of the first term in the error $e(\phi,h)$ in (6) for the invariant measure. It will thus be the key result in deriving integrators that have an order for the invariant measure strictly larger than the weak order of accuracy.

3 Main results: high order approximation of invariant measures

In this section, we present our methodology for constructing integrators of weak order p that approximate the ergodic averages with order $p+k$, with $k \geq 1$. In Section 3.1, we provide a characterization of numerical methods with high order invariant measure on $E = \mathbb{R}^d$ with $k = 1$ and then on $E = \mathbb{T}^d$ with arbitrary $k \geq 1$. We then introduce in Section 3.2 a framework based on modified equations to construct numerical method with high order invariant measure.

3.1 A characterization of high order numerical invariant measure

An immediate consequence of Theorem 2.7 is the following result in $\mathbb{R}^d$ which gives necessary conditions for an ergodic integrator of weak order p to have the higher order $p+1$ for the invariant measure.

Theorem 3.1. *Assume the hypothesis of Theorem 2.7. If an ergodic integrator of weak order* p *satisfies* $A_p^*\rho_\infty = 0$ *in the weak Taylor expansion* (18)*, then it has ergodic order* $r = p+1$ *in* (6)*.*

Proof. We consider the identity (27) and denote $D_p = (\frac{1}{(p+1)!}\mathcal{L}^{p+1} - A_p)$. The idea is to use the adjoint operator of D_p in (29), i.e., $(D_p v, \rho_\infty) = (v, D_p^*\rho_\infty)$. Using (11), we deduce from (29) in Theorem 2.7,

$$e(\phi,h) = h^p \left| \int_0^{+\infty} \int_{\mathbb{R}^d} v(y,t) A_p^*\rho_\infty(y)dydt \right| + \mathcal{O}(h^{p+1}),$$

where v is the solution of (28). Using the assumption $A_p^*\rho_\infty = 0$ yields the result $e(\phi, h) = \mathcal{O}(h^{p+1})$. □

We next show that sufficient conditions up to arbitrarily high order can be derived for the invariant measure error (6) on the torus $E = \mathbb{T}^d$. To this aim, we first recall a result from [4], which permits to expand the numerical invariant measure μ^h of an ergodic method in series with respect to h. The idea originating from backward error analysis is to construction a modified generator given as a formal series

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

such that $U(h, x)$ in (18) satisfies formally

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

The operators L_n can be computed recursively as

$$L_n = A_n - \frac{1}{2}(\mathcal{L}L_{n-1} + L_{n-1}\mathcal{L}) - \cdots - \frac{1}{(n+1)!}\mathcal{L}^{n+1} \tag{31}$$

where $A_i,\ i = 1, \cdots, n$ are the differential operators defined in (18). Equation (31) has been derived in [24] in the framework of modified equations and coincides with an expression used in [4] involving the Bernoulli numbers.

Lemma 3.2. *[4] Let $E = \mathbb{T}^d$ and assume Assumption 2.2. Consider L_n the operators defined in* (31)*. Then there exists a sequence of functions $(\rho_n(x))_{n>0}$ such that $\rho_0 = \rho_\infty$ and for all $n \geq 1$, $\int_{\mathbb{T}^d} \rho_n(x)dx = 0$ and*

$$\mathcal{L}^*\rho_n = -\sum_{l=1}^{n} (L_l)^*\rho_{n-l}. \tag{32}$$

For any positive integer M, if $\rho_M^h(x) = \rho_\infty(x) + \sum_{n=1}^{M} h^n \rho_n(x)$, then

$$\int_{\mathbb{T}^d} \rho_M^h(x)dx = 1, \tag{33}$$

and there exists a constant $C(M)$ such that for all smooth $\phi : \mathbb{R}^d \to \mathbb{R}$,

$$\left| \int_{\mathbb{T}^d} \phi(x)d\mu^h(x) - \int_{\mathbb{T}^d} \phi(x)\rho_M^h(x)dx \right| \leq C(M)h^{M+1}, \tag{34}$$

where $C(M)$ is independent of h and $\mu^h(x)$ is the unique invariant probability measure of the numerical method (16)*.*

We observe that Lemma 3.2 not only provides an expansion for the numerical invariant measure in powers of h, but also provides an explicit way for calculating the corrections ρ_n. For example, ρ_2 satisfies

$$\mathcal{L}^*\rho_2 = -L_1^*\rho_1 - L_2^*\rho_\infty$$

and since $\rho_1 = 0$ (assuming $A_1^* \rho_\infty = 0$) we have that

$$\mathcal{L}^* \rho_2 = A_2^* \rho_\infty - \frac{1}{2} \left(\mathcal{L}^* L_1^* + L_1^* \mathcal{L}^* \right) \rho_\infty - \frac{1}{6} (\mathcal{L}^*)^3 \rho_\infty.$$

Using (11) and $L_1^* \rho_\infty = A_1^* \rho_\infty - \frac{1}{2} L_0^2 \rho_\infty = 0$, this implies that

$$\mathcal{L}^* \rho_2 = A_2^* \rho_\infty.$$

We thus see with a similar argument as before that if a weak first order method satisfies $A_1^* \rho_\infty = A_2^* \rho_\infty = 0$ then its order of convergence for the ergodic averages is 3. Similarly, as generalized in the next theorem, we see that a sufficient condition for a numerical integrator of weak order p to have r-th order of convergence for the ergodic averages is

$$A_j^* \rho_\infty = 0, \quad \text{for} \quad j = 1, \cdots r - 1. \tag{35}$$

Of course an obvious way for achieving this is by choosing a method of weak order r (which implies $A_j^* = 0$ for all $j < r$, since $(j+1)! A_j = \mathcal{L}^{j+1}$), but as we will see in the next section for a certain class of ergodic SDEs we can achieve this by using a numerical integrator only of weak order one.

Theorem 3.3. *Consider equation* (1) *on* $\mathbb{T}^d$ *solved by an ergodic numerical method satisfying Assumption 2.2 and* (35)*. Then it has order* r *in* (6) *for the invariant measure.*

Proof. We start our proof by noticing on the one hand that since our numerical method is ergodic then

$$\lim_{N \to \infty} \frac{1}{N} \sum_{i=1}^{N} \phi(X_i) = \int_{\mathbb{T}^d} \phi(y) d\mu^h(y),$$

for all deterministic initial conditions $X_0 = x$. Thus, in order to prove the theorem one needs to bound the difference

$$\int_{\mathbb{T}^d} \phi(y) d\mu^h(y) - \int_{\mathbb{T}^d} \phi(y) \rho_\infty(y) dy. \tag{36}$$

On the other hand, Lemma 3.2 allows to expand $\rho_M^h(y)$ in powers of h and allows for an explicit characterization of each term in the expansion. Using (11), (31), and (32), we prove by induction on j that $\mathcal{L}^* \rho_j = A_j^* \rho_\infty = 0$ and $\rho_j = 0$ for $j = 1, \ldots, r-1$. Finally, using equation (34) with $M = r - 1$, observing that $\rho_\infty(y) = \rho_{r-1}^h(y)$ implies that

$$\left| \int_{\mathbb{T}^d} \phi(y) d\mu^h(y) - \int_{\mathbb{T}^d} \phi(y) \rho_\infty(y) dy \right| \leq C h^r,$$

where C depends on r but is independent of h. Thus the proof is complete. □

Remark 3.4. *One can extend to arbitrarily high order the extrapolation results described in [23] for the Euler and the Milstein methods. In particular, under the hypotheses of Theorem 2.6, a straightforward calculation shows that if one considers the Romberg extrapolation*

$$Z_n^h = \frac{2^p}{2^p - 1} \phi(X_{2n}^{h/2}) - \frac{1}{2^p - 1} \phi(X_n^h), \tag{37}$$

where X_n^h denotes the numerical solution at time $T = nh$ with stepsizes h, then Z_n^h yields an approximation of weak order $p+1$, i.e. $|\mathbb{E}(\phi(X(T))) - \mathbb{E}(Z_n^h)| \leq Ch^{p+1}$. Analogously, considering an ergodic method X_n^h of order p for the invariant measure and under the assumptions of Theorem 3.3, the Romberg extrapolation (37) yields an approximation of order $p+1$ for the invariant measure, i.e.

$$\left| \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} Z_n^h - \int_{\mathbb{T}^d} \phi(y)\rho_\infty(y)dy \right| \leq Ch^{p+1}.$$

3.2 High order numerical methods for the invariant measure based on modified equations

Our second main result is the derivation of a framework for the construction of numerical methods with high order (6) for the numerical invariant measure. We explain how Theorem 3.3 permits to construct high order integrators for the invariant measure by considering the framework of modified differential equations, an approach first considered in [24, 4] in the context of backward error analysis for the study of numerical integrators, and extended in [1] for the construction of high weak order integrators.

Precisely, given an ergodic integrator (2) with weak order p for an ergodic system of SDEs (1), we search for modified vector fields f_h and g_h of the form

$$f_h = f + h^p f_p + \ldots + h^{p+m-1} f_{p+m-1}, \qquad g_h = g + h^p g_p + \ldots + h^{p+m-1} g_{p+m-1},$$

such that the integrator (2) applied to the modified SDE

$$dX = f_h dt + g_h dW$$

has order $r = p+m$ in (6) with respect to the invariant measure. To this aim, we consider an ergodic SDE (1) and assume that it has an invariant measure whose Gibbs density function has the form

$$\rho_\infty(x) = Ze^{-V(x)} \tag{38}$$

where $Z = (\int_E e^{-V(x)}dx)^{-1}$ is a normalization constant. Again, we assume that it has bounded moments of any order, i.e. for all $n \geq 0$,

$$\int_E |x|^n \rho_\infty(x)dx < \infty$$

and we assume that the potential function $V : E \to \mathbb{R}$ is a smooth function in $C_P^\infty(E, \mathbb{R})$. Notice that the above assumptions on ρ_∞ are automatically satisfied if ρ_∞ is a smooth positive function on the torus $E = \mathbb{T}^d$. Furthermore, in the case $E = \mathbb{R}^d$, such an assumption is satisfied in the case of Brownian and Langevin dynamics (see Section 4).

Lemma 3.5. *Let $E = \mathbb{R}^d$ or $\mathbb{T}^d$. For all $\phi, w \in C_P^\infty(\mathbb{R}^d, \mathbb{R})$, consider the linear differential operator*

$$B\phi := w \frac{\partial^j \phi}{\partial x_{k_1} \cdots \partial x_{k_j}}, \tag{39}$$

where $k_i, i = 1, \ldots, j$ are indices with $1 \leq k_i \leq d$. Then, the following identity holds

$$\int_E (B\phi)\rho_\infty dx = \int_E (\widetilde{B}\phi)\rho_\infty dx, \quad \textit{for all } \phi \in C_P^\infty(\mathbb{R}^d, \mathbb{R}), \tag{40}$$

where $\widetilde{B}$ is the order one linear differential operator given by

$$\widetilde{B}\phi := \big(D_{k_2} \circ \cdots \circ D_{k_j}(w)\big)\frac{\partial \phi}{\partial x_{k_1}}$$

with D_i, $1 \leq i \leq d$ the linear differential operator defined as

$$D_i w := w\frac{\partial V}{\partial x_i} - \frac{\partial w}{\partial x_i}, \tag{41}$$

where V is the potential involved in the density (38)*.*

Proof. Integrating by parts successively with respect to $x_{k_2}, \ldots, x_{k_j}$, we obtain

$$\int_E B\phi\rho_\infty dx = \int_E \frac{\partial^j \phi}{\partial x_{k_1}\cdots\partial x_{k_j}} w\rho_\infty dx = (-1)^{j-1}\int_E \frac{\partial \phi}{\partial x_{k_1}}\frac{\partial^{j-1}(w\rho_\infty)}{\partial x_{k_2}\cdots\partial x_{k_j}}dx$$

We conclude using repeatedly the identity

$$\frac{\partial(w\rho_\infty)}{\partial x_i} = -(D_i w)\rho_\infty$$

for all w and all $i = k_2, \ldots, k_j$ (a consequence of $\frac{\partial \rho_\infty}{\partial x_i} = -\frac{\partial V}{\partial x_i}\rho_\infty$). □

The above lemma is a crucial ingredient to prove the following two theorems (for the space $E = \mathbb{R}^d$ and $E = \mathbb{T}^d$, respectively) on the construction of numerical integrators that approximate (1) with high order for the invariant measure.

Theorem 3.6. *Let $E = \mathbb{R}^d$. Consider an ergodic system of SDEs* (1) *with an invariant measure of the form* (38) *and a numerical method* (2) *or order p for the invariant measure, and satisfying Assumption 2.2. Then, there exists f_p such that if the numerical method applied to the modified SDE*

$$dX = (f + h^p f_p)dt + gdW \tag{42}$$

is ergodic then it has order $r = p + 1$ in (6) *for the invariant measure.*

Proof. By Assumption 2.2, the differential operator A_p in (18) is a sum of differential operators of the form (39), where w is an expression involving f and g and their derivatives.[2] It follows from Lemma 3.5 that there exists a smooth vector field $f_p : \mathbb{R}^d \to \mathbb{R}^d$ such that $A_p^* = \widetilde{A}_p^*\rho_\infty$, where $\widetilde{A}_p = -f_p \cdot \nabla\phi$, equivalently $A_p^*\rho_\infty = \mathrm{div}(f_p\rho_\infty)$. Using (19) and the definition (10), we deduce

$$A_p^*(f + h^p f_p, g)\phi = A_p^*(f, g)\phi - \mathrm{div}(f_p\phi) = 0.$$

Applying Theorem 3.1, we obtain that the numerical method applied to (42) yields an approximation of order $p + 1$ for the invariant measure of (1). □

[2]See for example the expression for A_1 in (21) for the θ-method.

Using Lemma 3.2 and extending the idea of the proof of Theorem (42), we derive the following result, for the construction of arbitrarily high order integrators for the invariant measure.

Theorem 3.7. *Let $E = \mathbb{T}^d$. Consider an ergodic system of SDEs* (1) *with an invariant measure of the form* (38) *and a numerical method* (2) *or order p for the invariant measure, and satisfying Assumption 2.2. Then, for all fixed $m \geq 1$, there exist a modified SDE of the form*

$$dX = (f + h^p f_p + \ldots + h^{p+m-1} f_{p+m-1})dt + gdW \tag{43}$$

such that the numerical method applied to this modified SDE satisfies

$$A_k^*(f + h^p f_p + \ldots + h^{p+m-1} f_{p+m-1}, g)\rho_\infty = 0 \quad k = p, \ldots, p+m-1. \tag{44}$$

Furthermore, if the numerical method applied to this modified SDE is ergodic, then this yields a method of order $r = p + m$ in (6) *for the invariant measure of* (1).

Proof. The construction of the vector fields $f_k, k < p+m$ is made by induction on k. Assume that $f_j, j < k$ has been constructed. Consider the scheme obtained by applying the numerical method to the modified SDE

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

and the corresponding weak expansion (22) involving the differential operators $A_j(f + \ldots + h^{k-1} f_{k-1}, g)$, $j = 1, 2, 3, \ldots$. It follows from Lemma 3.5 that for all differential operator of the form (39), we have $B^*\rho_\infty = \widetilde{B}^*\rho_\infty$ where $\widetilde{B}\phi$ is a differential operator of order one. Since by Assumption 2.2, A_k is a sum of such differential operator, we obtain that there exists a vector field f_k such that $A_k^*(f + \ldots + h^{k-1} f_{k-1}, g)\rho_\infty = \widetilde{A}_k^*\rho_\infty$ where $\widetilde{A}_k\phi = -f_k \cdot \nabla\phi$, equivalently

$$A_k^*(f + \ldots + h^{k-1} f_{k-1}, g)\rho_\infty = \mathrm{div}(f_k \rho_\infty). \tag{45}$$

Using (19) and the definition (10), we have

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

which yields

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

Using (45), this achieves the proof of (44). Applying Theorem 3.3, we conclude that the scheme applied to the modified SDE (43) has order $p + m$ for the invariant measure. $\square$

Note that the proofs of Theorems 3.6 and 3.7 not only show the existence of the vector fields f_i, but also provide an explicit way for calculating them. This is exemplified in the next section, where we discuss long time integrators for Brownian and Langevin dynamics.

4 Examples of high order integrators

We mention two wide classes of ergodic SDEs that have an invariant measure of the form (38), with a wide range of applications in different branches of physics, biology and chemistry.

The first one is the Langevin equation which describes the motion of a particle in the potential $U(q)$ subject to linear friction and molecular diffusion [18, 5]

$$dq = pdt, \qquad dp = -(\gamma p + \nabla U(q))dt + \sigma dW(t) \tag{46}$$

where $q(t), p(t) \in \mathbb{R}^d$ $U : \mathbb{R}^d \to \mathbb{R}$ is a smooth potential $\sigma > 0$ is a constant, and $W = (W_1, \ldots, W_d)^T$ is a vector of independent Wiener processes. It has the invariant measure density (38) with $V(p,q) = 2\gamma/\sigma^2\, H(p,q)$ and $H(p,q) = \frac{1}{2}p^2 + U(q)$ is the Hamiltonian.

The second one is the Brownian dynamics equation, describing the motion of a particle in a potential subject to thermal noise [18, 5]

$$dX(t) = -\nabla V(X(t))dt + \sigma dW(t), \tag{47}$$

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 vector of independent Wiener processes. Assuming ergodicity, the Gibbs density function of the invariant measure is given by

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

where Z is a renormalization constant such that $\int_{\mathbb{R}^d} \rho_\infty dx = 1$.

In this section, we shall focus on the class of ergodic SDEs (47) and construct numerical integrators that have low weak order of accuracy but high order with respect to the invariant measure (6). We emphasize that similar constructions could be obtained in the context of the Langevin equation (46).

For the nonlinear system of SDEs (47), consider the standard θ-method defined in (20) where $f = -\nabla V$. For general nonlinear systems (20), it can be checked that the weak order and the error (6) for the invariant measure coincide: it is 1 for $\theta \neq 1/2$ and 2 for $\theta = 1/2$. In this latter case, it is shown in [12] that the method samples exactly the invariant measure for linear problems (i.e. $e(\phi, h) = 0$ in (6) if V quadratic), but this is not true for nonlinear systems in general. In this section, we explain using the strategy of modified equations introduced in the previous section how the θ-method can be modified to increase the order (6) of accuracy for the invariant measure for nonlinear systems.

4.1 An illustrative example: linear case

As an example, consider first the linear scalar case where $V(x) = \gamma x^2$, corresponding to the classical Orstein-Uhlenbeck process,

$$dX = -\gamma X dt + \sigma dW. \tag{48}$$

The exact solution $X(t)$ is a Gaussian random variable satisfying $\lim_{t\to\infty} \mathbb{E}(X(t)^2) = \frac{\sigma^2}{2\gamma}$. Considering the Euler-Maruyama method, $x_{n+1} = x_n - \gamma h x_n + \sqrt{h}\sigma\xi_n$, a calculation yields

$$\lim_{n\to\infty} \mathbb{E}(x_n^2) = \frac{\sigma^2}{2\gamma(1 - \gamma h/2)}.$$

Then, applying the Euler-Maruyama method to the modified SDE

$$dX = -\tilde{\gamma}_h X dt + \sigma dW_t,$$

where $\tilde{\gamma}_h$ satisfies $\tilde{\gamma}_h(1 - \tilde{\gamma}_h h/2) = \gamma$, i.e. for all $h \leq 1/(2\gamma)$,

$$\tilde{\gamma}_h = h^{-1}(1 - \sqrt{1 - 2h\gamma}) = \gamma + \frac{h\gamma^2}{2} + \frac{h^2\gamma^3}{2} + \frac{5h^3\gamma^4}{8} + \frac{7h^4\gamma^5}{8} + \dots \tag{49}$$

yields a method which is exact for the invariant measure ($\rho_\infty^h = \rho_\infty$), i.e. the left hand side in (6) is zero, even-though the approximation has only weak order 2. Notice also that truncating (49) after the h^{p-1} term and applying the Euler-Maryuama yields a scheme of order p for the invariant measure.

4.2 Nonlinear case: modified theta method of order two for the invariant measure

Given a vector field f_1, consider the θ method applied to the modified SDE $dX = (f + hf_1)dt + \sigma dW$, i.e.,

$$X_{n+1} = X_n + (1-\theta)(f + hf_1)(X_n) + \theta(f + hf_1)(X_{n+1}) + \sqrt{h}\sigma\xi_n. \tag{50}$$

The following proposition with proof postponed to Appendix states that order two for the invariant measure can be achieved if the corrector f_1 is appropriately chosen.

Proposition 4.1. *Let $E = \mathbb{R}^d$ or $\mathbb{T}^d$. Consider the numerical method* (50) *applied to* (47)*. If*

$$f_1 = -(1 - 2\theta)\Big(\frac{1}{2}f'f + \frac{\sigma^2}{4}\Delta f\Big) \tag{51}$$

and (50) *is ergodic, then it has order $r = 2$ for the invariant measure in* (6)*.*

Remark 4.2. *In [1], a modified weak order two θ scheme was constructed for general systems of SDEs with non-commutative noise. In the context of additive noise* (47) *it has the form*

$$X_{n+1} = X_n + (1-\theta)(f - hf_1)(X_n) + \theta(f - hf_1)(X_{n+1}) + \sqrt{h}\sigma(\xi_n + h(\frac{1}{2} - \theta)f'(x_n)\xi_n).$$

It can be observed that both the drift and diffusion functions are modified in contrast to the scheme (50) *where only the drift function is modified. Notice that for $\theta = 1/2$, we have $f_1 = 0$ in* (51) *which is not surprising because in this case, the θ-method has weak order two of accuracy.*

Applying the recursive procedure of Theorem 3.7 we may next derive a modification of the θ method of order 3.

Proposition 4.3. *Let $E = \mathbb{T}^d$. Consider the Euler-Maruyama method applied to the modified SDE $dX = (f + hf_1 + h^2 f_2)dt + \sigma dW$ i.e.*

$$X_{n+1} = X_n + hf(X_n) + h^2 f_1(X_n) + h^3 f_2(X_n) + \sqrt{h}\xi_n, \tag{52}$$

where $f = -\nabla V$*,* f_1 *is defined in* (51) *with* $\theta = 0$ *and* f_2 *is defined by*

$$f_2 = -\Big(\frac{1}{2}f'f'f + \frac{1}{6}f''(f,f) + \frac{1}{3}\sigma^2\sum_i f''(e_i, f'e_i) + \frac{1}{4}\sigma^2 f'\Delta f\Big). \tag{53}$$

Assume that the obtained numerical method applied to (47) *is ergodic. Then, it has order* $r = 3$ *for the invariant measure in* (6).

The proof of Proposition 4.3 is postponed to Appendix.

Remark 4.4. *We highlight that integrators with arbitrarily higher order for the invariant measure could be constructed analogously using Theorem 3.7. The statement of Proposition 4.3 can be generalized to the* θ*-method* (20) *and yield again an order* 3 *method for the invariant measure, but the calculation becomes rather tedious. In the linear case* (48)*, the obtained scheme reduces to*

$$\begin{aligned} X_{n+1} &= x_n - (h\gamma + (1-2\theta)h^2\frac{\gamma^2}{2} + (1-2\theta)^2h^3\frac{\gamma^3}{2})\big((1-\theta)X_n + \theta X_{n+1}\big) \\ &+ \sigma\sqrt{h}\xi_n. \end{aligned} \tag{54}$$

For $\theta = 1/2$*, it coincides with the standard* θ*-method* (20) *which is not surprising because it samples the invariant measure exactly in this linear context [12].*

We shall discuss in the next Section 5 derivative free implementations of the new derived schemes.

5 Numerical experiments

In this section, we illustrate numerically our main results. We consider first the linear case (48) where $V(x) = x^2/2$, and compare the Euler-Maruyama method and the modifications of orders 2 (Proposition 4.3, $\theta = 0$). and 3 (Proposition 4.3, $\theta = 0$). In Figure 1, we plot the error $e(\phi, h)$ defined in (6) for $\phi(x) = x^2$ (second moment error) and many different stepsizes h. In theory computing one long trajectory suffices, however in practice computing several long trajectories allows also to draw some statistics such as the variance of the error. We therefore approximate the error using the average over 10 long trajectories on a time interval of length $T = 10^8$ and the deterministic initial condition[3] $X_0 = -2$. We observe the expected lines of slopes $1, 2, 3$ for the Euler-Maruyama method and the modifications of order $2, 3$.

We next consider examples of nonlinear problems in $E = \mathbb{R}^d$ which have non-globally Lipschitz coefficients. We emphasize that our results do not apply in this situation. However, numerical experiments still exhibit the high order convergence of the numerical invariant measure predicted in the Lipschitz case.

In Figure 2, we perform the same convergence experiment in the nonlinear with a quartic potential, either symmetric (left picture) or non-symmetric (right picture). Again, we observe the expected lines of slopes $1, 2, 3$ which corroborates Propositions 4.1 and 4.3.

We finally consider a multi-dimensional case ($d = 2$) of Brownian dynamics (47) with the nonlinear potential

$$V(x) = (1 - x_1^2)^2 + (1 - x_2^2)^2 + \frac{x_1 x_2}{2} + \frac{x_2}{5}. \tag{55}$$

[3]Recall that the choice of the initial condition has no influence on the results.

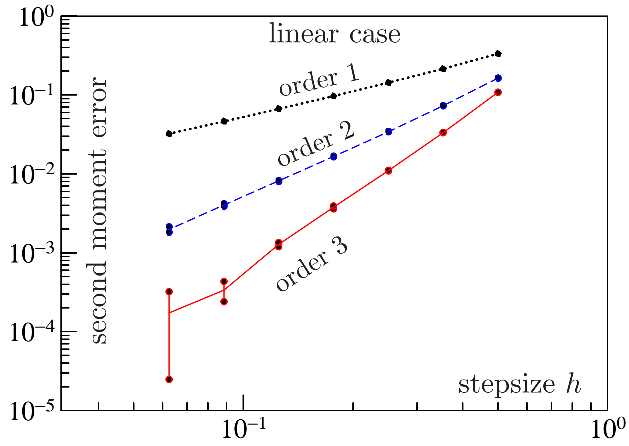

Figure 1: Linear case ($V(x) = x^2/2$). Euler-Maruyama method (order 1) and modifications of orders 2 and 3. Error for the second moment $\int_{\mathbb{R}} x^2\rho(x)dx$ versus time stepsize h obtained using 10 trajectories on a long time interval of length $T = 10^8$. The vertical bars indicate the standard deviation intervals.

This potential has one local maximum close to the origin and four local minima represented by white crosses in Figure 3 where we plot the Gibbs density function (38) together with 10 level curves (left and middle picture). The 10^5 gray dotes in the right picture indicate one numerical trajectory of the scheme (57) (discusses below) with stepsize $h = 0.02$ and time interval of size $T = 2 \cdot 10^3$ (the initial condition is $X_0 = (-2, -2)$).

Since calculating the derivative $f'f$ and Δf in (50)-(51) is not convenient in general for multi-dimensional systems and can be computational expensive, we introduce the following Runge-Kutta type scheme for (47)

$$\begin{aligned} Y_1 &= X_n + \sqrt{2}\sigma\sqrt{h}\xi_n \\ Y_2 &= X_n - \frac{3}{8}hf(Y_1) + \frac{\sqrt{2}}{4}\sigma\sqrt{h}\xi_n \\ X_{n+1} &= X_n - \frac{1}{3}hf(Y_1) + \frac{4}{3}hf(Y_2) + \sigma\sqrt{h}\xi_n \end{aligned} \tag{56}$$

where $f = -\nabla V$, $\xi_{n,i} \sim \mathcal{N}(0, 1)$ (or alternatively $\mathbb{P}(\xi_{n,i} = \pm\sqrt{3}) = 1/6$, $\mathbb{P}(\xi_{n,i} = 0) = 2/3$), are independent random variables. It can be checked straightforwardly that the weak Taylor expansions (18) of the schemes (56) and (50)-(51) coincide up to order 2, i.e. they have the same operators A_0, A_1 and thus the same order 2 in (6) for the invariant measure, and the same weak order 1. This is detailed in the Appendix (see Proposition 6.1).

Our investigations indicate that there does not exist a similar Runge-Kutta type approximation of the scheme (52) with only 3 evaluations of the function f per timestep. We thus propose the following Runge-Kutta type method which has order 2 in (6) for general

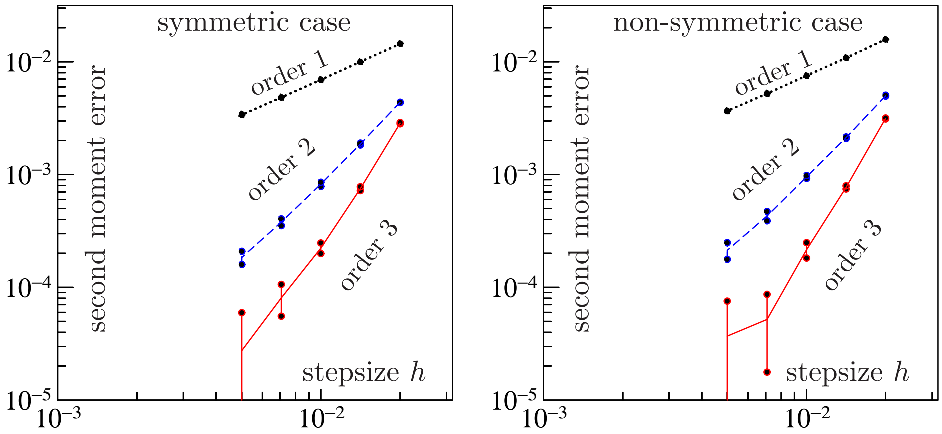

Figure 2: Nonlinear problem with double-well potential. Left picture: $V(x) = (1 - x^2)^2$ (symmetric). Right picture: $V(x) = (1 - x^2)^2 - x/2$ (non-symmetric). Euler-Maruyama method (order 1) and modifications of orders 2 and 3. Error for the second moment $\int_{\mathbb{R}} x^2 \rho(x) dx$ versus time stepsize h obtained using 10 trajectories on a long time interval of length $T = 10^8$. The vertical bars indicate the standard deviation intervals.

nonlinear multi-dimensional problems (47), but order 3 for linear problems,

$$\begin{aligned} Y_1 &= X_n + \sigma\sqrt{h}\xi_n \\ Y_2 &= X_n - \frac{h}{2}f(Y_1) + \frac{\sigma}{2}\sqrt{h}\xi_n \\ Y_3 &= X_n + 3hf(Y_1) - 2hf(Y_2) + \sigma\sqrt{h}\xi_n \\ X_{n+1} &= X_n - \frac{3}{2}hf(Y_1) + 2hf(Y_2) + \frac{1}{2}hf(Y_3) + \sigma\sqrt{h}\xi_n \end{aligned} \tag{57}$$

where $f = -\nabla V$ and ξ_n is a vector of independent random variables with $\xi_{n,j} \sim \mathcal{N}(0, 1)$. We plot in Figure 4 the errors $e(\phi, h)$ for $\phi(x) = x^2 + y^2$ for the Euler-Maruyama method, and the modifications (56) and (57). We observe the expected lines of slope $1, 2$. Notice that the error constant for the variant (57) is about twice as smaller than the error for

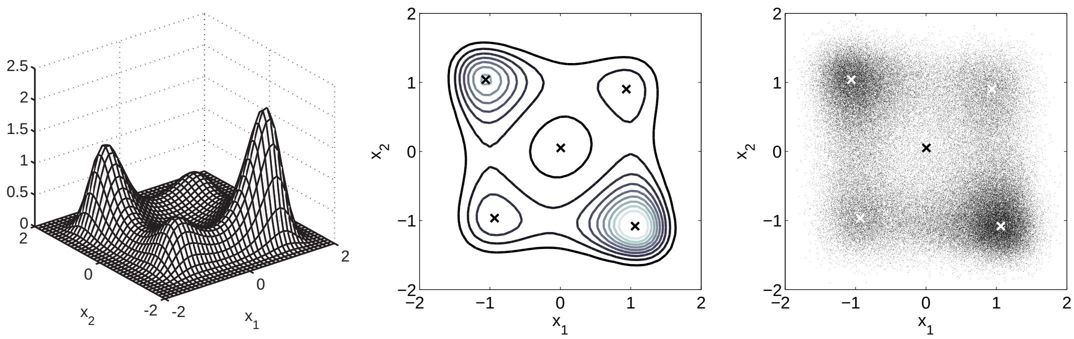

Figure 3: 2D problem (47)-(55). Left picture: 3D plot of the Gibbs density (38). Middle picture: ten level curves of the Gibbs density are represented in solid lines (the five extrema are represented with crosses). Right picture: a numerical trajectory $\{X_n\}$ of the scheme (57) (with $h = 0.02$, $T = 2 \cdot 10^3$).

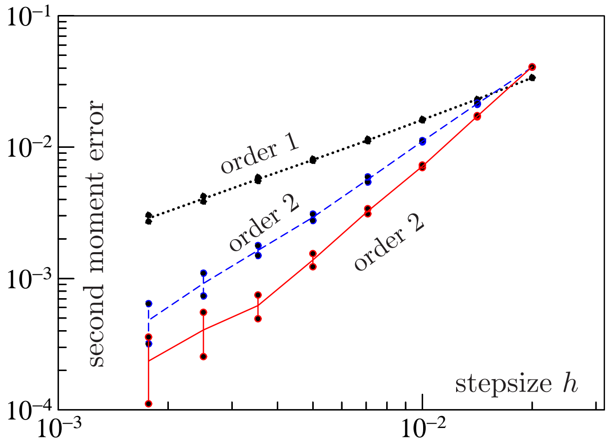

Figure 4: 2D problem (47)-(55). Errors for $\phi(x) = x^2+y^2$ for the Euler-Maruyama method (order 1), and the modifications (56) (order 2) and (57) (order 2 but 3 for linear problems) with $T = 10^7$.

(56). The results for the scheme (50) are not included in this plot, but are nearly identical to that of (56).

6 Appendix

We provide in this Appendix a proof of Theorem 2.6 for the derivation of a global error expansion for a weak numerical method of arbitrary order p. We next give the proofs of Propositions 4.1, 4.3, 6.1.

Proof of Theorem 2.6. Consider u the solution of (13), then $v(x,t) = u(x, T-t)$ is the solution of the Kolomogorov equation (28) with $v(x,t) = \mathbb{E}(\phi(X(T))|X(t) = x)$. Setting $t_i = ih$, the weak error (25) can be expressed as

$$\begin{aligned} E(\phi, h, T) &= \mathbb{E}(\phi(X(T))) - \mathbb{E}(\phi(X_N)) \\ &= \mathbb{E}(v(X_0, 0)) - \mathbb{E}(v(X_N, T)) \\ &= \sum_{i=1}^{N} \left(\mathbb{E}\left(v(X_{i-1}, t_{i-1})\right) - \mathbb{E}\left(v(X_i, t_i)\right)\right) \end{aligned} \tag{58}$$

where $Nh = T$.Now using properties of conditional expectations give

$$E(\phi, h, T) = \sum_{i=1}^{N} \Big(\mathbb{E}\big(v(X(t_i), t_i)|X(t_{i-1}) = X_{i-1}\big) - \mathbb{E}\big(v(X_i, t_i)\big)\Big).$$

If we denote by $g_i(x) = e^{(T-t_i)\mathcal{L}}\phi(x)$ the solution of (28) at time $t = ih$, we obtain

$$E(\phi, h, T) \quad = \quad \sum_{i=1}^{N} \Big(\mathbb{E}\big(g_i(X(t_i))|(X(t_{i-1}) = X_{i-1})\big) - \mathbb{E}\big(g_i(X_i)\big)\Big).$$

Using (14),(22) and the weak order of convergence p of the method, we obtain

$$E(\phi, h, T) \quad = \quad \sum_{i=1}^{N} \mathbb{E}\left(h^{p+1} D_p g_i(X_{i-1}) + h^{p+2} d_i\right)$$

$$= \sum_{i=1}^{N}\left(h^{p+1}\mathbb{E}\left(D_p e^{-h\mathcal{L}} g_{i-1}(X_{i-1})\right) + h^{p+2}\mathbb{E}(d_i)\right)$$
$$= \sum_{i=1}^{N}\left(h^{p+1}\mathbb{E}\left(D_p e^{-h\mathcal{L}} v(X_{i-1}, t_{i-1})\right) + h^{p+2}\mathbb{E}(d_i)\right)$$

where we have used the fact that $g_i = e^{-h\mathcal{L}} g_{i-1}$, and for notational brevity we have dropped the dependence of d_i on f, g, ϕ, h and $X(t), X_{i-1}$. Next using (23), we see that $D_p = \frac{1}{(p+1)!}\mathcal{L}^{p+1} - \mathcal{A}_p$. Using the regularity of the solution of (28) we obtain

$$E(\phi, T, h) = \sum_{i=1}^{N}\left(h^{p+1}\mathbb{E}(\psi_e((X_{i-1}, t_{i-1}))) + h^{p+2}\mathbb{E}(d_i)\right), \tag{59}$$
$$= h^p \int_0^T \mathbb{E}(\psi_e(X(s), s))ds + h^{p+2}\sum_{i=1}^{N}\mathbb{E}(d_i) \tag{60}$$
$$+ \; h^p\left(h\sum_{i=1}^{N}\mathbb{E}(\psi_e(X_{i-1}, t_{i-1})) - \int_0^T \mathbb{E}(\psi_e(X(s), s))ds\right). \tag{61}$$

Using Remark 2.1 and Proposition 2.4 we see that $\mathbb{E}(d_i) \leq C(T)$. To conclude the proof, it remains to show

$$h\sum_{i=1}^{N}\mathbb{E}(\psi_e(X_{i-1}, t_{i-1})) - \int_0^T \mathbb{E}(\psi_e(X(s), s))ds = \mathcal{O}(h). \tag{62}$$

Indeed, this term can be bounded by

$$h\sum_{i=1}^{N}|\mathbb{E}(\psi_e(X_{i-1}, t_{i-1})) - \mathbb{E}(\psi_e(X(t_{i-1}), t_{i-1}))| \tag{63}$$
$$+ \; \left|h\sum_{i=1}^{N}\mathbb{E}(\psi_e(X(t_{i-1}), t_{i-1})) - \int_0^T \mathbb{E}(\psi_e(X(s), s))ds\right|. \tag{64}$$

The first term is bounded by $\mathcal{O}(h)$ using $E(\psi_e(\cdot, t_i), T, h) = \mathcal{O}(h)$ uniformly in $t_i \leq T$. The second term is bounded by $\mathcal{O}(h)$ using that $s \mapsto \mathbb{E}(\psi_e(X(s), s))$ has a continuous derivative. □

Proof of Proposition 4.1. Consider the weak Taylor expansion (18) for the θ method. Applying Lemma 3.5 to each differential operator of order greater than 1 in $\mathcal{A}_1$ given in (21) and using $f = -\nabla V$, we obtain

$$\langle \phi''(f,f)\rangle = \left\langle -\phi'(f'f + (\operatorname{div} f)f + \frac{2}{\sigma^2}\|f\|^2 f)\right\rangle,$$
$$\left\langle \sigma^2\sum_i \phi'''(f, e_i, e_i)\right\rangle = \left\langle \phi'(\sigma^2\sum_i f''(e_i, e_i) + 4f'f + 2(\operatorname{div} f)f + \frac{4}{\sigma^2}\|f\|^2 f)\right\rangle,$$
$$\left\langle \sigma^2\sum_{ij}\phi^{(4)}(e_i, e_i, e_j, e_j)\right\rangle = \left\langle -\sum_i 2\phi'''(f, e_i, e_i)\right\rangle,$$

$$\left\langle \frac{\sigma^2}{2}\sum_i \phi''(f'e_i,e_i)\right\rangle = \left\langle -\phi'(\sigma^2\sum_i f''(e_i,e_i)+2f'f)\right\rangle,$$

where we use the notation $\langle u\rangle = \int_E u(x)\rho_\infty(x)dx$ and the sums are for $i,j=1,\ldots,d$ and e_i is the canonical basis of $\mathbb{R}^d$. Using the above identities, a straightforward calculation then yields that f_1 in (51) satisfies $\langle \mathcal{A}_1\phi\rangle = \langle f_1\cdot\nabla\phi\rangle$, equivalently $\mathcal{A}_1^*\rho_\infty = \mathrm{div}(f_1\rho_\infty)$. Theorem 3.1 (for $E=\mathbb{R}^d$) and Theorem 3.7 (for $E=\mathbb{T}^d$) conclude the proof. □

Proof of Proposition 4.3. Consider the weak Taylor expansion (18) for the modified θ method (50) ($\theta=0$). We have $A_0=\mathcal{L}$ because the method has weak order 1, and by the construction of Theorem 3.7, $A_1^*\rho_\infty=0$. A calculation of A_2 yields

$$\begin{aligned}
A_2\phi = &-\frac{1}{2}\phi''(f,f'f) - \sum_i \frac{\sigma^2}{4}\phi''(f,f''(e_i,e_i))\\
&-\sum_{ij}\frac{\sigma^4}{8}\phi^{(3)}(f''(e_i,e_i),e_j,e_j) - \sum_i \sigma^2\frac{1}{4}\phi^{(3)}(f'f,e_i,e_i)\\
&+\frac{1}{6}\phi^{(3)}(f,f,f) + \sum_i\frac{\sigma^2}{4}\phi^{(4)}(f,f,e_i,e_i) + \sum_{ij}\frac{\sigma^4}{8}\phi^{(5)}(f,e_i,e_i,e_j,e_j)\\
&+\sum_{ijk}\frac{\sigma^6}{48}\phi^{(6)}(e_i,e_i,e_j,e_j,e_k,e_k).
\end{aligned}$$

Applying repeatedly integration by parts as in Lemma 3.5 (see the proof of Proposition 4.1) then yields

$$\begin{aligned}
\left\langle\sigma^2\phi''(f'e_i,f'e_i)\right\rangle &= \left\langle\phi'(-\sigma^2 f''(e_i,f'e_i) - f'\nabla(\sigma^2\mathrm{div}\, f + \|f\|^2))\right\rangle\\
\left\langle\phi''(f,f'f)\right\rangle &= \left\langle -\phi'(f'f'f + f''(f,f) + (\mathrm{div}\, f)f'f + \frac{2}{\sigma^2}\|f\|^2 f'f)\right\rangle\\
\left\langle\phi''(f,f''(e_i,e_i))\right\rangle &= \left\langle -\phi'(f'''(f,e_i,e_i) + (\mathrm{div}\, f)f''(e_i,e_i) + \frac{2}{\sigma^2}\|f\|^2 f''(e_i,e_i))\right\rangle\\
\left\langle\sigma^4\phi^{(3)}(f''(e_i,e_i),e_j,e_j)\right\rangle &= \left\langle -\phi'(\sigma^4 f^{(4)}(e_i,e_i,e_j,e_j) + 4\sigma^2 f'''(f,e_i,e_i)\right.\\
&+\left. 2(\mathrm{div}\, f)f''(e_i,e_i) + 4\|f\|^2 f''(e_i,e_i))\right\rangle\\
\left\langle\sigma^2\phi^{(3)}(f'f,e_i,e_i)\right\rangle &= \left\langle\phi'(\sigma^2 f'''(f,e_i,e_i) + 2\sigma^2 f''(f'e_i,e_i) + \sigma^2 f'f''(e_i,e_i)\right.\\
&+\left. 4(f'f'f + f''(f,f)) + 2(\mathrm{div}\, f)f'f + \frac{4}{\sigma^2}\|f\|^2 f'f)\right\rangle\\
\left\langle\sigma^2\phi^{(3)}(f,f'e_i,e_i)\right\rangle &= \left\langle -\sigma^2\phi''(f,f''(e_i,e_i)) - \sigma^2\phi''(f'e_i,f'e_i) - 2\phi''(f'f,f)\right\rangle\\
\left\langle\sigma^4\phi^{(4)}(e_i,e_i,f'e_j,e_j)\right\rangle &= \left\langle -\sigma^4\phi^{(3)}(f''(e_i,e_i),e_j,e_j) - 2\sigma^2\phi^{(3)}(f'f,e_i,e_i)\right\rangle\\
\left\langle\sigma^2\phi^{(4)}(f,f,e_i,e_i)\right\rangle &= \left\langle -2\phi^{(3)}(f,f,f) - 2\sigma^2\phi^{(3)}(f,f'e_i,e_i)\right\rangle\\
\left\langle\sigma^4\phi^{(5)}(f,e_i,e_i,e_j,e_j)\right\rangle &= \left\langle -2\sigma^2\phi^{(4)}(f,f,e_i,e_i) - \sigma^4\phi^{(4)}(e_i,e_i,f'e_j,e_j)\right\rangle\\
\left\langle\sigma^6\phi^{(6)}(e_i,e_i,e_j,e_j,e_k,e_k)\right\rangle &= \left\langle -2\sigma^4\phi^{(5)}(f,e_i,e_i,e_j,e_j)\right\rangle
\end{aligned}$$

where sums should be taken over all indices $i, j, k = 1, \ldots, d$ in the above formulas (omitted for brevity of the notation). Using the symmetry of $f' = -\nabla^2 V$, we have $\nabla \mathrm{div}\, f = \Delta f$ and $\nabla(\|f\|^2) = 2f'f$ in the first equality and we obtain $A_2^* \rho_\infty = \mathrm{div}(f_2 \rho_\infty)$. We conclude the proof using Theorem 3.7. $\square$

Proposition 6.1. *Consider the method* (56) *for* (47) *on the space* $E = \mathbb{R}^d$ *or* $\mathbb{T}^d$ *and assume that it is ergodic. Then, it has order order* $r = 2$ *in* (6) *for the invariant measure.*

Proof. We justify the construction of the derivative free implementation (56) of the scheme (50) ($\theta = 0$). Consider a Runge-Kutta type scheme of the form

$$Y_i = X_n + h \sum_{j=1}^{s} a_{ij} f(Y_j) + \overline{c}_i \sqrt{h} \xi_n, \quad X_{n+1} = X_n + h \sum_{i=1}^{s} b_i f(Y_i) + \sigma \sqrt{h} \xi_n,$$

with coefficients $a_{ij}, b_j, \overline{c}_i$, with $i, j = 1, \ldots, s$. Setting $c_i = \sum_{j=1}^{s} a_{ij}$, we expand in Taylor series the numerical solution,

$$\begin{aligned} X_1 &= X_0 + h(\sum_{i=1}^{s} b_i) f + \sqrt{h} \sigma \xi_n + h^{3/2} \sigma (\sum_{i=1}^{s} b_i \overline{c}_i) f' \xi_n \\ &+ h^2 (\sum_{i=1}^{s} b_i c_i) f' f + \frac{h^2 \sigma^2}{2} (\sum_{i=1}^{s} b_i \overline{c}_i^2) f''(\xi_n, \xi_n) + \ldots \end{aligned}$$

and we deduce the differential operators in the weak Taylor expansion (18),

$$\begin{aligned} A_0 \phi &= (\sum_{i=1}^{s} b_i) f \cdot \nabla \phi + \frac{1}{2} \sigma^2 \Delta \phi, \\ A_1 \phi &= \big((\sum_{i=1}^{s} b_i c_i) f' f + (\sum_{i=1}^{s} b_i \overline{c}_i) \sigma \mathrm{div} f + \frac{\sigma^2}{2} (\sum_{i=1}^{s} b_i \overline{c}_i^2) \Delta f \big) \cdot \nabla \phi. \end{aligned}$$

Then, imposing the order conditions

$$\sum_{i=1}^{s} b_i = 1, \quad \sum_{i=1}^{s} b_i c_i = -\frac{1}{2}, \quad \sum_{i=1}^{s} b_i \overline{c}_i = 0, \quad \sum_{i=1}^{s} b_i \overline{c}_i^2 = -\frac{1}{2},$$

yields the same operators $A_0 = \mathcal{L}$ and $A_1 \phi = -\big(\frac{1}{2} f' f + \frac{\sigma^2}{4} \Delta f\big) \cdot \nabla \phi$ as for the scheme (50) ($\theta = 0$) and thus the same order two for the invariant measure. $\square$

Acknowledgements. The research of A. A. is partially supported by Swiss National Foundation Grant 200021_140692.

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] 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.

[3] K. Burrage and G. Lythe. Accurate stationary densities with partitioned numerical methods for stochastic differential equations. *SIAM J. Numer. Anal.*, 47(3):1601–1618, 2009.

[4] A. Debussche and E. Faou. Weak backward error analysis for SDEs. *SIAM J. Numer. Anal.*, 50(3):1735–1752, 2012.

[5] C. W. Gardiner. *Handbook of stochastic methods*. Springer-Verlag, Berlin, second edition, 1985. For physics, chemistry and the natural sciences.

[6] 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.

[7] R. Hasminskii. *Stochastic stability of differential equations*. Sijthoff and Noordhoff, The Netherlands, 1980.

[8] D. J. Higham. Mean-square and asymptotic stability of the stochastic theta method. *SIAM J. Numer. Anal.*, 38(3):753–769, 2000.

[9] M. Hutzenthaler, A. Jentzen, and P. E. Kloeden. Strong and weak divergence in finite time of Euler's method for stochastic differential equations with non-globally Lipschitz continuous coefficients. *Proc. Roy. Soc. Edinburgh, Sect. A*, 467(2130):1563–1576, 2011.

[10] B. Leimkuhler and C. Matthews. Rational construction of stochastic numerical methods for molecular sampling. *Applied Mathematics Research eXpress*, 2013(1):34–56, 2013.

[11] B. Leimkuhler and S. Reich. *Simulating Hamiltonian Dynamics*. Cambridge Monographs on Applied and Computational Mathematics 14. Cambridge University Press, Cambridge, 2004.

[12] T. Li, A. Abdulle, and W. E. Effectiveness of implicit methods for stiff stochastic differential equations. *Commun. Comput. Phys.*, 3(2):295–307, 2008.

[13] J. C. Mattingly, A. M. Stuart, and D. J. Higham. Ergodicity for SDEs and approximations: locally Lipschitz vector fields and degenerate noise. *Stochastic Process. Appl.*, 101(2):185–232, 2002.

[14] J. C. Mattingly, A. M. Stuart, and M. V. Tretyakov. Convergence of numerical time-averaging and stationary measures via Poisson equations. *SIAM J. Numer. Anal.*, 48(2):552–577, 2010.

[15] G. Milstein. Weak approximation of solutions of systems of stochastic differential equations. *Theory Probab. Appl.*, 30(4):750–766, 1986.

[16] G. Milstein and M. Tretyakov. *Stochastic numerics for mathematical physics.* Scientific Computing. Springer-Verlag, Berlin and New York, 2004.

[17] 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, 2005.

[18] H. Risken. *The Fokker-Planck equation*, volume 18 of *Springer Series in Synergetics.* Springer-Verlag, Berlin, 1989.

[19] G. O. Roberts and R. L. Tweedie. Exponential convergence of langevin distributions and their discrete approximations. *Bernoulli*, 2(4):pp. 341–363, 1996.

[20] T. Shardlow and A. M. Stuart. A perturbation theory for ergodic markov chains and application to numerical approximations. *SIAM J. Numer. Anal.*, 37(4):1120–1137, 2000.

[21] D. Talay. *Second Order Discretization Schemes of Stochastic Differential Systems for the Computation of the Invariant Law.* Rapports de recherche. Institut National de Recherche en Informatique et en Automatique, 1987.

[22] D. Talay. Stochastic Hamiltonian systems: exponential convergence to the invariant measure, and discretization by the implicit Euler scheme. *Markov Process. Related Fields*, 8(2):163–198, 2002. Inhomogeneous random systems (Cergy-Pontoise, 2001).

[23] D. Talay and L. Tubaro. Expansion of the global error for numerical schemes solving stochastic differential equations. *Stochastic Anal. Appl.*, 8(4):483–509 (1991), 1990.

[24] 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.