SIAM J. NUMER. ANAL.
Vol. 52, No. 4, pp. 1600–1622

HIGH ORDER NUMERICAL APPROXIMATION OF THE INVARIANT MEASURE OF ERGODIC SDES*

ASSYR ABDULLE†, GILLES VILMART‡, AND KONSTANTINOS C. ZYGALAKIS§

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.

Key words. stochastic differential equations, weak convergence, modified differential equations, backward error analysis, invariant measure, ergodicity

AMS subject classifications. 65C30, 60H35, 37M25

DOI. 10.1137/130935616

1. Introduction.

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

$$dX(t) = f(X(t))dt + g(X(t))dW(t), \quad X(0) = X_0, \tag{1.1}$$

where $X(t)$ is the solution in the space E, $X_0 \in E$ is the initial condition assumed deterministic for simplicity, $f : E \mapsto E, g : E \mapsto E^m$ are smooth, and $W(t)$ is a standard d-dimensional Wiener process. The space E denotes either $E = \mathbb{R}^d$ or the torus $E = \mathbb{T}^d$ and is specified when needed. With the exception of some special cases, the solutions to (1.1) are not explicitly known, and numerical methods are needed. We first state our results on the torus and then explain extensions to $\mathbb{R}^d$. Working on the torus permits having automatically finite moments of any order for both the exact and numerical solutions of (1.1), thus avoiding technicalities. We consider a one step numerical integrator for the approximation of (1.1) at time $t = nh$ of the form

$$X_{n+1} = \Psi(X_n, h, \xi_n), \tag{1.2}$$

where h denotes the stepsize and ξ_n are independent random vectors. The choice behind the numerical method used to approximate (1.1) 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

*Received by the editors September 4, 2013; accepted for publication (in revised form) March 27, 2014; published electronically July 10, 2014. The work of the first and second authors was partially supported by the Swiss National Foundation under grants 200021_140692 and 200020_144313/1, respectively.
http://www.siam.org/journals/sinum/52-4/93561.html

†Mathematics Section, École Polytechnique Fédérale de Lausanne, 1015 Lausanne, Switzerland (Assyr.Abdulle@epfl.ch).
‡Current address: Section de mathématiques, Université de Genève, 1211 Genève 4, Switzerland. On leave from École Normale Supérieure de Rennes, INRIA Rennes, IRMAR, CNRS, UEB, av. Robert Schuman, F-35170 Bruz, France (Gilles.Vilmart@ens-rennes.fr).
§Mathematical Sciences, University of Southampton, Southampton SO17 1BJ, UK (k.zygalakis@soton.ac.uk).

of the solution, one is interested in its weak convergence properties. The numerical approximation (1.2) of (1.1), starting from the initial condition $X_0 = x \in E$, is said to have local weak order p if for all test functions $\phi \in \mathcal{C}_P^\infty(E, \mathbb{R})$ (with all derivatives at all orders of polynomial growth in the case $E = \mathbb{R}^d$),

$$|\mathbb{E}(\phi(X_1)) - \mathbb{E}(\phi(X(h)))| \leq C(x, \phi)h^{p+1} \tag{1.3}$$

for all h sufficiently small, where $C(x, \phi)$ is independent of h but depends on x, ϕ. Under appropriate conditions one can infer "a global weak order p" from the local weak error, also in the $\mathbb{R}^d$ setting [19] (see [20, Chap. 2.2]).

Strong and weak types of convergence relate to the finite time properties of (1.1) and its numerical approximations. We say that the process $X(t)$ is ergodic if it has a unique invariant measure μ satisfying for each smooth function ϕ and for any deterministic initial condition $X_0 = x$,

$$\lim_{T\to\infty} \frac{1}{T}\int_0^T \phi(X(s))ds = \int_E \phi(y)d\mu(y) \qquad \text{almost surely.} \tag{1.4}$$

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 [21, 17, 24, 25, 27, 11, 13, 12]. 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 (1.2) is called ergodic if it has a unique invariant probability law μ^h with finite moments of any order and

$$\lim_{N\to\infty} \frac{1}{N+1}\sum_{n=0}^{N} \phi(X_n) = \int_E \phi(y)d\mu^h(y) \qquad \text{almost surely} \tag{1.5}$$

for all deterministic initial condition $X_0 = x$ and all smooth test functions ϕ.

We will say that the numerical method (1.2) has order $r \geq 1$ with respect to the invariant measure if

$$|e(\phi, h)| \leq Ch^r \quad \text{with} \quad e(\phi, h) := \lim_{N\to\infty} \frac{1}{N+1}\sum_{n=0}^{N} \phi(X_n) - \int_E \phi(y)d\mu(y), \tag{1.6}$$

where C is independent of h small enough and X_0. In what follows, 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 methods 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 $e(\phi, h)$ have appeared in the literature. In [26] an error estimate for $e(\phi, h)$ was established for a variety of different numerical methods. In addition, in [28] with a use of a global weak error expansion, an expansion of (1.6) 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 [18]. Given an ergodic integrator of weak order p for an ergodic SDE (1.1), it is shown that it has order $r \geq p$ for the invariant measure (1.6). In [16] 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 [5, 4, 14]. Such phenomena where a low order integrator preserves certain invariants with higher order is classical in geometric integration of deterministic ODEs [8, 15]. For instance, a symplectic Runge–Kutta method of order p preserves the energy of Hamiltonian systems at the same order p without drift over long times, but it also conserves exactly quadratic first integrals.

In this paper, we present two results for the numerical approximation of ergodic nonlinear systems of SDEs. First, we derive new sufficient conditions for an ergodic integrator to have high order (1.6) for the invariant measure, possibly larger than its weak order of accuracy (1.3). A crucial ingredient is a new expansion of the error $e(\phi, h)$ based on the work [28] and the analysis in [6, 12, 13] of numerical invariant measures. Second, 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 [8, 15]. It was recently extended to SDEs in [29, 6] for the backward error analysis of stochastic integrators and in [1] for the construction of high weak order integrators. The integrators designed using the proposed framework involve high order derivatives of the drift and diffusion functions which can make the integrators costly and inefficient in general for large systems. We show, however, that Runge–Kutta type formulation of these schemes can also be constructed to avoid the such derivatives.

The paper is organized as follows. In section 2, we present the framework used for the analysis, based on the backward Kolmogorov and Fokker–Planck equations. 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. The extension of our results to $\mathbb{R}^d$ is discussed in section 4. In section 5, we apply our methodology and construct a range of new integrators based on the stochastic θ-method for Brownian dynamics. Finally in section 6, we present various numerical investigations that illustrate the behavior of our new integrators and corroborate the claimed orders of convergence.

2. Preliminaries.

In this section, we describe some preliminary results related to ergodicity of SDEs and their numerical approximations, using the standard framework of the backward Kolmogorov and Fokker–Planck equations. We also recall the formal expansion of the solution of the backward Kolmogorov equation, the weak Taylor expansion for a numerical integrator, and a series expansion of the numerical invariant measure based on backward error analysis.

2.1. Setting.

We start by recalling that the differential operator $\mathcal{L}$

$$\mathcal{L} := f \cdot \nabla + \frac{1}{2} g g^T : \nabla^2, \tag{2.1}$$

where $\nabla^2 \phi$ denotes the Hessian of ϕ, is called the generator of the SDE (1.1).[1] We next state our basic assumptions.

Assumption 2.1. We assume the following:

- 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 (1.1).

[1] We use the notation of the scalar product $A : B = \text{trace}(A^T B)$ on matrices.

Under these assumptions, there exists a unique the invariant measure μ. It has a density function ρ_∞ which is the unique solution of the equation

$$\mathcal{L}^* \rho_\infty = 0, \tag{2.2}$$

where $\mathcal{L}^*$ is the L^2-adjoint of the generator $\mathcal{L}$, given by

$$\mathcal{L}^* \phi = -\nabla_y \cdot (f\phi) + \frac{1}{2} g g^T : \nabla^2 \phi. \tag{2.3}$$

We also have a unique solution for $u(t,x) = \mathbb{E}\left(\phi(X(t)) | X_0 = x\right)$ satisfying the backward Kolmogorov equation

$$\frac{\partial u}{\partial t} = \mathcal{L} u, \qquad u(x,0) = \phi(x), \tag{2.4}$$

where $\phi \in \mathcal{C}^\infty(\mathbb{T}^d, \mathbb{R})$. If we denote by $\rho(y,t)$ the probability density of the random variable $X(t)$ defined by (1.1) with initial condition $X_0 = x$, we have

$$\mathbb{E}\left(\phi(X(t)) | X_0 = x\right) = \int_E \phi(y) \rho(y,t) dy, \tag{2.5}$$

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

$$\frac{\partial \rho}{\partial t} = \mathcal{L}^* \rho, \tag{2.6a}$$

$$\rho(y,0) = \delta(y-x), \tag{2.6b}$$

where δ denotes the Dirac measure in zero and $\mathcal{L}^*$ is given by (2.3). We further assume that there exists a constant λ and for all integer $k \geq 0$ constants C_k, κ_k such that for all $t \geq 0$

$$\|u(t,\cdot) - \int_{\mathbb{T}^d} \phi(y) \rho_\infty(y) dy\|_{\mathcal{C}^k} \leq C_k (1 + t^{\kappa_k}) e^{-\lambda t} \|\phi\|_{\mathcal{C}^k}, \tag{2.7}$$

where $\|v(t,\cdot)\|_{\mathcal{C}^k}$ denotes the norm of the function $v(x,t)$ and its derivatives with respect to x up to order k. Notice that setting $t \to \infty$ in (2.7) and using (2.4) yields

$$\lim_{t\to\infty} u(x,t) = \phi(x) + \int_0^\infty \mathcal{L} u(t,x) dt = \int_{\mathbb{T}^d} \phi(y) \rho_\infty(y) dy. \tag{2.8}$$

We refer to [3, 18, 6, 12, 13] for a discussion of Assumption 2.1 and (2.7).

Assumption 2.1 is naturally satisfied for Brownian and Langevin dynamics on the torus, but also in $\mathbb{R}^d$ under appropriate smoothness and growth assumptions on the potential involved (see [13, 12, 17]). The Brownian dynamics equation describes the motion of a particle in a potential subject to thermal noise [23, 7]

$$dX(t) = -\nabla V(X(t)) dt + \sigma dW(t), \tag{2.9}$$

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

$$\rho_\infty = Z e^{-2V(x)/\sigma^2}, \tag{2.10}$$

where Z is a renormalization constant such that $\int_{\mathbb{T}^d} \rho_\infty dx = 1$. We also mention the Langevin equation which has the invariant measure density with the same form (2.10), where $V(p,q) = \beta\, H(p,q)$ and $H(p,q) = \frac{1}{2}p^2 + U(q)$ denotes the Hamiltonian. It describes the motion of a particle in the potential $U(q)$ subject to linear friction and molecular diffusion [23, 7]

$$dq = pdt, \qquad dp = -(\gamma p + \nabla U(q))dt + \sqrt{2\beta^{-1}\gamma}dW(t), \tag{2.11}$$

where $q(t) \in \mathbb{T}^d, p(t) \in \mathbb{R}^d$, $U : \mathbb{T}^d \to \mathbb{R}$ is a smooth potential, $\gamma, \beta > 0$ are constants, and $W = (W_1, \dots, W_d)^T$ is a standard d-dimensional Wiener process.

2.2. Series expansion of the numerical invariant measure. A formal Taylor series expansion in terms of the generator operator $\mathcal{L}$ of the Markov process is derived in [29] for u and a rigorous finite term expansion is proposed in [6], namely,

$$u(x,h) - \phi(x) = \sum_{j=1}^{l} \frac{h^j}{j!}\mathcal{L}^j \phi(x) + h^{l+1} r_l(f,g,\phi)(x), \tag{2.12}$$

where for all positive integer l, the remainder $r_l(f,g,\phi)$ is bounded on the torus.

In terms of the numerical solution (1.2) one can define for all smooth test function ϕ,

$$U(x,h) = \mathbb{E}(\phi(X_1)|X_0 = x)) \tag{2.13}$$

for the expectation at time h. We make the following regularity and consistency assumption on the integrator, which is easily satisfied by any reasonable numerical method.

Assumption 2.2. We assume that (2.13) has a weak Taylor series expansion of the form,

$$U(x,h) = \phi(x) + hA_0(f,g)\phi(x) + h^2 A_1(f,g)\phi(x) + \cdots, \tag{2.14}$$

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 (2.1), which means that the method has (at least) local order one in the weak sense,

$$A_0(f,g) = \mathcal{L}. \tag{2.15}$$

Example 2.1. Consider the stochastic θ-method [10] for (1.1), where $g = \sigma I$ and $d = m$ (additive noise case) defined as

$$X_{n+1} = X_n + h(1-\theta)f(X_n) + \theta f(X_{n+1}) + \sigma\sqrt{h}\xi_n. \tag{2.16}$$

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

(2.14) is given by

$$\begin{aligned} A_1\phi &= \frac{1}{2}\phi''(f,f) + \frac{\sigma^2}{2}\sum_{i=1}^{d}\phi'''(e_i,e_i,f) + \frac{\sigma^4}{8}\sum_{i,j=1}^{d}\phi^{(4)}(e_i,e_i,e_j,e_j) \\ &\quad + \theta\phi'\left(f'f + \frac{\sigma^2}{2}\sum_{i=1}^{d} f''(e_i,e_i)\right) + \theta\sigma^2\sum_{i=1}^{d}\phi''(f'e_i,e_i), \end{aligned} \tag{2.17}$$

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, ..., forms, respectively. In dimension $d=1$, it reduces to $A_1\phi = \frac{1}{2}f^2\phi'' + \frac{\sigma^2}{2}f\phi''' + \frac{\sigma^4}{8}\phi^{(4)} + \theta(f'f\phi' + \frac{\sigma^2}{2}f''\phi' + \sigma^2 f'\phi'')$.

Since on the torus, all numerical moments are automatically bounded, Assumption 2.2 immediately implies that we have for all $\phi \in \mathcal{C}^\infty(\mathbb{T}^d,\mathbb{R})$ the rigorous expansion

$$U(x,h) = \phi(x) + \sum_{i=0}^{l} h^{i+1}A_i(f,g)\phi(x) + h^{l+2}R_l(f,g,\phi)(x), \tag{2.18}$$

where for all positive integers l, the remainer $R_l(f,g,\phi)$ is bounded on the torus.

We next recall the main result in [6], which permits expanding the numerical invariant measure μ^h of an ergodic method in series with respect to h. The idea, originating from backward error analysis for ODEs [8, 15], 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 (2.14) 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

$$L_n = A_n - \frac{1}{2}(\mathcal{L}L_{n-1} + L_{n-1}\mathcal{L} + \cdots) - \cdots - \frac{1}{(n+1)!}\mathcal{L}^{n+1}, \tag{2.19}$$

where $A_i,\ i=1,\dots,n$ are the differential operators defined in (2.14). Equation (2.19) has been derived in [29] in the framework of modified equations and coincides with an expression used in [6] involving the Bernoulli numbers.

LEMMA 2.1 (see [6]). *Let $E=\mathbb{T}^d$ and suppose that Assumptions* 2.1 *and* 2.2 *and* (2.7) *hold. Consider L_n the operators defined in* (2.19)*. Then there exists a sequence of functions $(\rho_n(x))_{n\geq 0}$ such that $\rho_0=\rho_\infty$ and for all $n\geq 1$, $\int_{\mathbb{T}^d}\rho_n(x)dx = 0$, and*

$$\mathcal{L}^*\rho_n = -\sum_{l=1}^{n}(L_l)^*\rho_{n-l}. \tag{2.20}$$

For any positive integer M, setting

$$\rho_M^h(x) = \rho_\infty(x) + \sum_{n=1}^{M} h^n\rho_n(x),$$

there exists a constant $C(M,\phi)$ such that for all $\phi \in \mathcal{C}^\infty(\mathbb{T}^d,\mathbb{R})$,

$$\left|\int_{\mathbb{T}^d} \phi(x)d\mu^h(x) - \int_{\mathbb{T}^d} \phi(x)\rho_M^h(x)dx\right| \leq C(M,\phi)h^{M+1}, \tag{2.21}$$

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

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 on the torus $E = \mathbb{T}^d$ with order of at least $p+k$ with $k \geq 1$. In section 3.1, we provide a characterization of numerical methods with high order invariant measure. We then introduce in section 3.2 a framework based on modified equations to construct numerical method with high order invariant measure. Extensions of our results to $\mathbb{R}^d$ are discussed in section 4.

3.1. A characterization of high order numerical invariant measure. We observe that Lemma 2.1 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. In Theorem 3.1, we prove that a sufficient condition for a numerical integrator of weak order p to have rth order of convergence for the ergodic averages is that Assumption 2.2 holds with

$$A_j^*\rho_\infty = 0 \quad \text{for} \quad j = 1,\dots r-1. \tag{3.1}$$

Remark 3.1. An interpretation of (3.1) is that the invariant measure μ is invariant through one step of the numerical integrator up to a $\mathcal{O}(h^r)$ error. Precisely,

$$\left|\mathbb{E}(\phi(X_1)|X_0 \sim \mu) - \int_{\mathbb{T}^d} \phi(x)d\mu(x)\right| \leq Ch^r,$$

where C is independent of h but depends on the test function $\phi \in \mathcal{C}^\infty(\mathbb{T}^d,\mathbb{R})$.

An obvious way to achieve (3.1) is by choosing a method of weak order r (which implies $A_j^*\rho_\infty = 0$ for all $j < r$, since $(j+1)!A_j = \mathcal{L}^{j+1}$), but as shown below, we can also achieve this by using a numerical integrator only of weak order one. For example, ρ_1 and ρ_2 in Lemma 2.1 satisfy

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

where $L_1^* = A_1^* - \frac{1}{2}(\mathcal{L}^*)^2$. Assuming $A_1^*\rho_\infty = 0$ and using (2.2) then yields $L_1^*\rho_\infty = 0$ and thus $\rho_1 = 0$. We obtain

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

and thus $\rho_2 = 0$ if in addition $A_2^*\rho_\infty = 0$. We thus see using (2.21) with $M = 2$ 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. More generally, we have the following result.

THEOREM 3.1. *Consider the SDE* (1.1) *on $\mathbb{T}^d$ satisfying Assumption* 2.1 *and* (2.7)*, and solved by an ergodic numerical method satisfying Assumption* 2.2 *and* (3.1)*. Then it has (at least) order r in* (1.6) *for the invariant measure. More precisely, the invariant measure error $e(\phi,h)$ in* (1.6) *satisfies for all $\phi \in \mathcal{C}^\infty(\mathbb{T}^d,\mathbb{R})$ and $h \to 0$,*

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

where $u(x,t)$ solves the backward Kolmogorov equation (2.4).

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

$$\lim_{N\to\infty} \frac{1}{N+1} \sum_{n=0}^{N} \phi(X_n) = \int_{\mathbb{T}^d} \phi(y) d\mu^h(y)$$

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

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

On the other hand, Lemma 2.1 allows one to expand $\rho_M^h(y)$ in powers of h and allows for an explicit characterization of each term in the expansion. Using (2.2), (2.19), and (2.20), we prove by induction on j that $\mathcal{L}^*\rho_j = A_j^*\rho_\infty = 0$ and $\rho_k = 0$ for $j = 1, \ldots, r-1$. Finally, using (2.21) with $M = r$, observing that $\rho_r^h(y) = \rho_\infty(y) + \rho_r(y)$ implies

$$\left| \int_{\mathbb{T}^d} \phi(y) d\mu^h(y) - \int_{\mathbb{T}^d} \phi(y) \rho_\infty(y) dy - h^r \int_{\mathbb{T}^d} \phi(y) \rho_r(y) dy \right| \leq C h^{r+1},$$

where C depends on r, ϕ but is independent of h. Using (2.8) and $\int_{\mathbb{T}^d} \rho_r(x) dx = 0$, we deduce

$$\int_{\mathbb{T}^d} \phi(y) \rho_r(y) dy = -\int_0^\infty \int_{\mathbb{T}^d} \mathcal{L}u(t,y) \rho_r(y) dy dt = \int_0^\infty \int_{\mathbb{T}^d} u(t,y) A_r^* \rho_\infty(y) dy dt,$$

where we used $\mathcal{L}^*\rho_r = -A_r^*\rho_\infty$ in the last identity. This concludes the proof. ∎

Remark 3.2. Under the hypotheses of Theorem 3.1, we may also deduce high accuracy results in finite time. Following [6, Thm. 2.1], there exist constants $C, \lambda, \kappa > 0$ such that for all $k \geq 0$,

$$\left| \mathbb{E}(\phi(X_k)) - \int_{\mathbb{T}^d} \phi(x) \rho_\infty(x) dx \right| \leq C((1 + |t_k|^\kappa) e^{-\lambda t_k} + h^r),$$

where $t_k = kh$, the constants C, λ, κ are independent of k, h, and h is assumed small enough (and C depends on ϕ).

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 (1.6) for the numerical invariant measure. We explain how Theorem 3.1 permits constructing high order integrators for the invariant measure by considering the framework of modified differential equations, an approach first considered in [29, 6] in the context of backward error analysis for the study of stochastic integrators, and extended in [1] for the construction of high weak order integrators.

Precisely, given an ergodic integrator (1.2) with order p for the invariant measure of an ergodic system of SDEs (1.1), we search for modified vector fields f_h and g_h of the form

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

such that the integrator (1.2) applied to the modified SDE

$$dX = f_h dt + g_h dW$$

has order $r = p + m$ in (1.6) with respect to the invariant measure. To this aim, we consider an ergodic SDE (1.1) and assume that it has an invariant measure whose Gibbs density function has the form

$$\rho_\infty(x) = Ze^{-V(x)}, \tag{3.2}$$

where $Z = (\int_{\mathbb{T}^d} e^{-V(x)}dx)^{-1}$ is a normalization constant. We assume that the potential function $V : \mathbb{T}^d \to \mathbb{R}$ is a smooth function in $C^\infty(\mathbb{T}^d, \mathbb{R})$. Notice that the above assumptions on ρ_∞ are automatically satisfied if ρ_∞ is a smooth strictly positive function on the torus $\mathbb{T}^d$. Furthermore, in the case $E = \mathbb{R}^d$, such an assumption is satisfied in the case of Brownian and Langevin dynamics (see section 5).

The following lemma shows, using integration by parts, that for any high order linear differential operator B with smooth coefficients, there exist an order one differential operator $\widetilde{B}$ such that $B^*\rho_\infty = \widetilde{B}^*\rho_\infty$. For instance, for the differential operator A_1 of order 4 given in (2.17) for the θ-method applied to Brownian dynamics (2.9), one can construct a vector field f_1 such that $A_1^*\rho_\infty = \mathrm{div}(f_1\rho_\infty)$. (See more details in Proposition 5.1 in section 5.)

Lemma 3.2. *For all $\phi, w \in C^\infty(\mathbb{T}^d, \mathbb{R})$, consider the linear differential operator*

$$B\phi := w\frac{\partial^j \phi}{\partial x_{k_1}\cdots\partial x_{k_j}}, \tag{3.3}$$

where $k_i, i = 1, \dots, j$ are indices with $1 \le k_i \le d$. Then, the following identity holds:

$$\int_{\mathbb{T}^d}(B\phi)\rho_\infty dx = \int_{\mathbb{T}^d}(\widetilde{B}\phi)\rho_\infty dx \quad \textit{for all } \phi \in C^\infty(\mathbb{T}^d, \mathbb{R}), \tag{3.4}$$

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, $1 \le i \le d$ the linear differential operator defined as

$$D_i w := w\frac{\partial V}{\partial x_i} - \frac{\partial w}{\partial x_i}, \tag{3.5}$$

where V is the potential involved in the density (3.2).

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

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

We conclude using repeatedly the identity

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

for all w and all $i = k_2, \dots, k_j$ (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 theorem on the construction of numerical integrators that approximate (1.1) with high order for the invariant measure.

Theorem 3.3. *Consider an ergodic system of SDEs* (1.1) *in $\mathbb{T}^d$ with an invariant measure of the form* (3.2) *and a numerical method* (1.2) *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*

$$dX = (f + h^p f_p + \cdots + h^{p+m-1} f_{p+m-1})dt + gdW \tag{3.6}$$

such that the numerical method applied to this modified SDE satisfies

$$A_k^*(f + h^p f_p + \cdots + h^{p+m-1} f_{p+m-1}, g)\rho_\infty = 0 \quad k = p, \ldots, p+m-1. \tag{3.7}$$

Furthermore, if the numerical method applied to this modified SDE is ergodic, then this yields a method of order (at least) $r = p + m$ *in* (1.6) *for the invariant measure of* (1.1).

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

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

and the corresponding weak expansion (2.18) involving the differential operators $A_j(f + \cdots + h^{k-1} f_{k-1}, g)$, $j = 1, 2, 3, \ldots$. It follows from Lemma 3.2 that for all differential operator of the form (3.3), we have $B^*\rho_\infty = \widetilde{B}^*\rho_\infty$, where $\widetilde{B}$ is a differential operator of order one. Since by Assumption 2.2, A_k is a sum of such differential operator,[2] we obtain that there exists a vector field f_k such that $A_k^*(f + \cdots + h^{k-1} f_{k-1}, g)\rho_\infty = \widetilde{A}_k^*\rho_\infty$, where $\widetilde{A}_k = -f_k \cdot \nabla$, or, equivalently,

$$A_k^*(f + \cdots + h^{k-1} f_{k-1}, g)\rho_\infty = \mathrm{div}(f_k \rho_\infty). \tag{3.8}$$

Using (2.15) and the definition (2.1), we have

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

which yields

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

Using (3.8), this achieves the proof of (3.7). Applying Theorem 3.1, we conclude that the scheme applied to the modified SDE (3.6) has order $p + m$ for the invariant measure. $\square$

Note that the proof of Theorem 3.3 not only shows the existence of the vector fields f_i but also provides an explicit way for calculating them. This is exemplified in section 5, where we discuss long time integrators for Brownian and Langevin dynamics.

4. Extension to $\mathbb{R}^d$.

In this section, we explain how the results of Theorems 3.1 and 3.3 derived on the torus can be generalized to $\mathbb{R}^d$.

4.1. Basic tools. We denote $\mathcal{C}_P^\infty(\mathbb{R}^d, \mathbb{R})$ the set of $\mathcal{C}^\infty$ functions whose derivatives up to any order have a polynomial growth of the form

$$|\phi(x)| \leq C(1 + |x|^s) \tag{4.1}$$

for some constants s and C independent of x. For simplicity, following [28, Lem. 2], we assume that f, g in (1.1) are C^∞ with bounded derivatives up to any order. This together with the assumption $\phi(x) \in \mathcal{C}_P^\infty(\mathbb{R}^d, \mathbb{R})$ implies that the backward Kolmogorov equation (2.4) has a unique smooth solution $u(x,t) \in \mathbb{R}^d$ whose derivatives up to

[2] See, for example, the expression for A_1 in (2.17) for the θ-method.

any order have a polynomial growth with respect to $x \in \mathbb{R}^d$. This makes the Taylor expansion (2.12) also rigorous in $\mathbb{R}^d$ with a remainder with a polynomial growth with respect to x.

In addition, we make the following assumption which guaranties that the numerical moments of all orders remain bounded along time.

Assumption 4.1. We assume that the numerical integrator (1.1) satisfies for all $x \in \mathbb{R}^d$ that

$$|\mathbb{E}(X_1 - X_0 | X_0 = x)| \leq C(1 + |x|)h, \qquad |X_1 - X_0| \leq M(1 + |X_0|)\sqrt{h}, \tag{4.2}$$

where C is independent of h small enough and M is a random variable that has bounded moments of all orders independent of h and X_0.

In addition, assuming that Assumption 2.2 holds for all $\phi \in \mathcal{C}_P^\infty(\mathbb{R}^d, \mathbb{R})$ immediately implies for all $\phi \in \mathcal{C}_P^\infty(\mathbb{R}^d, \mathbb{R})$ that the numerical solution $U(x, h) = \mathbb{E}(\phi(X_1)|X_0 = x)$ has the rigorous expansion

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

for all positive integers l with a remainder satisfying $|R_l(f, g, \phi)(x)| \leq C_l(1 + |x|^{k_l})$ for some constants C_l, k_l. We also deduce that the moments of the numerical solution are uniformly bounded, as stated in the following lemma, shown in the proof of [20, Lem. 2.2, p. 102].

PROPOSITION 4.1 (see [20]). *Assume Assumption* 4.1. *Then, for all positive integers k, there exist constants C_k, D_k independent of n, h such that*

$$\mathbb{E}\big(|X_n|^k\big) \leq C_k e^{D_k t_n} \quad \text{with} \quad t_n = nh. \tag{4.3}$$

The following theorem permits one to infer the global weak order of convergence from the local order p of convergence of a given numerical integrator. Using Assumption 2.2 in $\mathbb{R}^d$, the local weak order p of the numerical scheme can be written out as

$$\mathbb{E}(\phi(X(h))) - \mathbb{E}(\phi(X_1)) = h^{p+1}\left(\frac{\mathcal{L}^{p+1}}{(p+1)!} - A_p\right)\phi(X_0) + \mathcal{O}(h^{p+2}). \tag{4.4}$$

Theorem 4.2 combines results derived by Talay and Milstein. Precisely, the expression (4.5) has been proved in [28] for specific methods (e.g., the 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 [19] and can be found in [20, Chap. 2.2, 2.3]. The proof of Theorem 4.2 is thus omitted.

THEOREM 4.2. *Assume that f, g in* (1.1) *are C^∞ with bounded derivatives up to any order. Let X_N be a numerical solution of* (1.1) *on $[0, T]$ $(E = \mathbb{R}^d)$ satisfying Assumption* 2.2 *in $\mathbb{R}^d$, Assumption* 4.1, *and the local weak order p estimate* (1.3), *where $C(x)$ has a polynomial growth of the form* (4.1). *Then, we have the following expansion of the global error, for all $\phi \in C_P^\infty(\mathbb{R}^d, \mathbb{R})$,*

$$\mathbb{E}(\phi(X(T))) - \mathbb{E}(\phi(X_N)) = h^p \int_0^T \mathbb{E}(\psi_e(X(s), s))ds + \mathcal{O}(h^{p+1}), \tag{4.5}$$

where $\psi_e(x, t)$ satisfies

$$\psi_e(x, t) = \left(\frac{1}{(p+1)!}\mathcal{L}^{p+1} - A_p\right) v(x, t) \tag{4.6}$$

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

$$\frac{\partial v}{\partial t} + \mathcal{L}v = 0, \qquad v(x,T) = \phi(x). \tag{4.7}$$

4.2. High order for the numerical invariant measure in $\mathbb{R}^d$. In what follows, we also assume that the solution $X(t)$ of (1.1) is ergodic with an invariant measure μ with density function ρ_∞ that has bounded moments of any order, i.e., for all $n \geq 0$,

$$\int_{\mathbb{R}^d} |x|^n \rho_\infty(x) dx < \infty. \tag{4.8}$$

These assumptions hold if one supposes the following sufficient conditions (see [9]).

Assumption 4.2. We assume the following:

1. f, g are of class C^∞ with bounded derivatives of any order, and g is bounded.
2. The generator $\mathcal{L}$ in (2.1) 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 $x \in \mathbb{R}^d$, $x^T f(x) \leq -\beta x^T x + C$.

Using Theorem 4.2 one can obtain a similar expansion to (4.5) for the difference between the true and the numerical ergodic averages. In particular we have the following theorem which provides an explicit expression of the first term in the error $e(\phi, h)$ in (1.6) for the invariant measure. It will next be the key result in deriving integrators that have an order for the invariant measure strictly larger than the weak order of accuracy.

THEOREM 4.3. *Assume that the hypotheses of Theorem* 4.2 *and Assumption* 4.2 *hold. Then, if a numerical method of weak order p is ergodic, its invariant measure error in* (1.6) *satisfies for all* $\phi \in \mathcal{C}_P^\infty(\mathbb{R}^d, \mathbb{R})$ *and* $h \to 0$,

$$e(\phi, h) = -\lambda_p h^p + \mathcal{O}(h^{p+1}) \tag{4.9}$$

for any deterministic initial condition with λ_p *defined as*

$$\lambda_p = \int_0^{+\infty} \int_{\mathbb{R}^d} \left(\frac{1}{(p+1)!} \mathcal{L}^{p+1} - A_p \right) u(y,t) \rho_\infty(y) dy dt, \tag{4.10}$$

where $u(x,t)$ *is the solution of* (2.4).

Proof. The proof is similar to the one found in [28, Thm. 4], with the main difference being that now (4.5) is used as the starting point of the proof instead of the specific formula for the Euler–Maruyama method used in [28]. ∎

An immediate consequence of Theorem 4.3 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 4.4. *Assume the hypothesis of Theorem* 4.3. *If an ergodic integrator of weak order p satisfies* $A_p^* \rho_\infty = 0$ *in the weak Taylor expansion* (2.14)*, then it has ergodic order (at least)* $r = p+1$ *in* (1.6).

Proof. We consider the identity (4.10) and use the L^2-adjoint of the differential operator $\frac{1}{(p+1)!}\mathcal{L}^{p+1} - A_p$. This implies

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

Using (2.2) yields $(\mathcal{L}^*)^{p+1} \rho_\infty = 0$, which concludes the proof. ∎

THEOREM 4.5. *Consider an ergodic system of SDEs* (1.1) *in* $\mathbb{R}^d$ *with an invariant measure of the form* (3.2) *and a numerical method* (1.2) *or order p for the invariant* *measure, and satisfying Assumption* 2.2. *Then, there exists a smooth vector field f_p such that the numerical method applied to the modified SDE*

$$dX = (f + h^p f_p)dt + gdW \tag{4.11}$$

satisfies

$$A_p^*(f + h^p f_p, g)\rho_\infty = 0.$$

Furthermore, if the numerical method applied to this modified SDE is ergodic and satisfies Assumption 4.1, *then it has order (at least) $r = p+1$ in* (1.6) *for the invariant measure.*

Proof. By Assumption 2.2, the differential operator A_p in (2.14) is a sum of differential operators of the form (3.3), where w is an expression involving f and g and their derivatives. We observe that Lemma 3.2 remains valid replacing the space $\mathcal{C}^\infty(\mathbb{T}^d, \mathbb{R})$ by $\mathcal{C}_P^\infty(\mathbb{R}^d, \mathbb{R})$. It follows that there exists a smooth vector field $f_p : \mathbb{R}^d \to \mathbb{R}^d$ such that $A_p^*\rho_\infty = \widetilde{A}_p^*\rho_\infty$, where $\widetilde{A}_p = -f_p \cdot \nabla$, or, equivalently, $A_p^*\rho_\infty = \mathrm{div}(f_p\rho_\infty)$. Using (2.15) and the definition (2.1), we deduce

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

Applying Theorem 4.4, we obtain that the numerical method applied to (4.11) yields an approximation of order $p + 1$ for the invariant measure of (1.1). ∎

4.3. Brownian dynamics in $\mathbb{R}^d$. The results in [6] on the torus were recently extended to $\mathbb{R}^d$ for Brownian and Langevin dynamics in [13] and [12], respectively. The main difficulty is to fullfill Assumption 2.1 and (2.7) in this context. For Brownian dynamics (2.9), in the nonglobally Lipschtiz setting of semiconvex potentials $V = V_1 + V_2$, where $V_1, V_2 \in \mathcal{C}_P^\infty(\mathbb{R}^d, \mathbb{R})$ with V_1 convex and V_2 bounded, assuming the third condition in Assumption 4.2 and (4.8), it is proved in [13] that Assumption 2.1 and (2.7) hold in $\mathbb{R}^d$ for all $\phi \in \mathcal{C}_P^\infty(\mathbb{R}^d, \mathbb{R})$, and Assumption 2.2 and the boundedness of numerical moments hold for two specific implicit schemes.[3] The implicitness is used in [13] to guarantee the boundedness of the numerical moments in spite of the fact that ∇V is nonglobally Lipschiz. Here, we shall rather assume its global Lipschizness using the first condition in Assumption 4.2. This permits us to consider a general class of integrators satisfying Assumptions 2.2 and 4.1, which also have bounded numerical moments. We may now state the following lemma, which is a variant in the globally Lipschiz case of the main result in [13].

LEMMA 4.6. *For Brownian dynamics* (2.9) *on $\mathbb{R}^d$ satisfying the conditions in Remark* 4.2, *consider a numerical integrator fulfilling Assumptions* 2.2, 4.1. *Then there exists a sequence of functions $(\rho_n(x))_{n\geq 0}$ such that $\rho_0 = \rho_\infty$ and for all $n \geq 1$, $\int_{\mathbb{R}} \rho_n(x)dx = 0$ and setting $\rho_M^h(x) = \rho_\infty(x) + \sum_{n=1}^M h^n\rho_n(x)$,* (2.20), (2.21) *hold with $\mathbb{T}^d$ replaced by $\mathbb{R}^d$ and for $\phi \in \mathcal{C}_P^\infty(\mathbb{R}^d, \mathbb{R})$.*

Based on Lemma 4.6, we may extend Theorem 3.1 in $\mathbb{R}^d$ for Brownian dynamics.

THEOREM 4.7. *For Brownian dynamics* (2.9) *on $\mathbb{R}^d$, assume the hypotheses of Lemma* 4.6 *and assume that* (3.1) *holds for a given r. Then the integrator has order (at least) r in* (1.6) *for the invariant measure. More precisely, the invariant measure*

[3]Namely, we refer to the implicit Euler scheme $X_{n+1} = X_n - h\nabla V(X_{n+1}) + \sigma\Delta W_n$ and the implicit split-step scheme $X_{n+1}^* = X_n - h\nabla V(X_{n+1}^*)$, $X_{n+1} = X_{n+1}^* + \sigma\Delta W_n$, where $\Delta W_n \sim \mathcal{N}(0, hI)$ are independent Gaussian random vectors with dimension d.

error in (1.6) *satisfies for all* $\phi \in \mathcal{C}^\infty(\mathbb{R}^d, \mathbb{R})$ *and* $h \to 0$,

$$e(\phi, h) = h^r \int_0^\infty \int_{\mathbb{R}^d} u(x,t) A_r^* \rho_\infty(x) dx dt + \mathcal{O}(h^{r+1}),$$

where $u(x,t)$ *solves the backward Kolmogorov equation* (2.4).

Analogously, we have the following theorem, which extends to Brownian dynamics in $\mathbb{R}^d$ the statement of Theorem 3.3.

THEOREM 4.8. *For Brownian dynamics* (2.9) *on* $\mathbb{R}^d$, *assume the hypotheses of Lemma* 4.6. *Assume that the numerical method* (1.2) *has order* p *for the invariant measure. Then, for all fixed* $m \geq 1$, *there exist a modified SDE of the form*

$$dX = (f + h^p f_p + \cdots + h^{p+m-1} f_{p+m-1})dt + g dW$$

such that the numerical method applied to this modified SDE satisfies

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

Furthermore, if the numerical method applied to this modified SDE is ergodic and satisfies Assumption 4.1, *then this yields a method of order (at least)* $r = p + m$ *in* (1.6) *for the invariant measure of* (1.1).

Proof of Theorem 4.7 *and Theorem* 4.8. The proofs are identical to that of Theorems 4.7 and 3.3, respectively, with the exception that we now rely on Lemma 4.6 in $\mathbb{R}^d$ instead of Lemma 2.1 in $\mathbb{T}^d$. □

Remark 4.1. One can extend to arbitrarily high order the extrapolation results described in [28] for the Euler and Milstein methods. In particular, under the hypotheses of Theorem 4.2, a straightforward calculation shows that if one considers the Romberg extrapolation

$$Z_n^h = \frac{2^p}{2^p - 1}\phi(X_{2n}^{h/2}) - \frac{1}{2^p - 1}\phi(X_n^h), \tag{4.12}$$

where X_n^h denotes the numerical solution of weak order p 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 3.1 (for $E = \mathbb{T}^d$) and Theorem 4.7 (for $E = \mathbb{R}^d$), the Romberg extrapolation (4.12) yields an approximation of order $p+1$ for the invariant measure, i.e.,

$$\left| \lim_{N\to\infty} \frac{1}{N+1} \sum_{n=0}^N Z_n^h - \int_E \phi(y)\rho_\infty(y)dy \right| \leq Ch^{p+1}.$$

5. Examples of high order integrators. We highlight that the Brownian dynamics (2.9) and the Langevin dynamics (2.11) are two wide classes of ergodic SDEs that have an invariant measure of the form (3.2), with a wide range of applications in different branches of physics, biology, and chemistry.

In this section, we shall focus on the class of Brownian SDEs (2.9) and construct numerical integrators that have low weak order of accuracy but high order with respect to the invariant measure (1.6). We emphasize that similar constructions could be obtained in the context of the Langevin equation (2.11).

For the nonlinear system of SDEs (2.9), consider the standard θ-method defined in (2.16), where $f = -\nabla V$. For general nonlinear systems (2.16), it can be checked that the weak order and the error (1.6) 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 [16] that the method samples exactly the invariant measure for linear problems (i.e., $e(\phi, h) = 0$ in (1.6) 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 (1.6) of accuracy for the invariant measure for nonlinear systems.

5.1. 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,

$$dX = -\gamma X dt + \sigma dW. \tag{5.1}$$

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 $h \le 1/(2\gamma)$,

$$\tilde{\gamma}_h = h^{-1}(1 - \sqrt{1-2h\gamma}) = \gamma + \frac{h\gamma^2}{2} + \frac{h^2\gamma^3}{2} + \frac{5h^3\gamma^4}{8} + \frac{7h^4\gamma^5}{8} + \cdots \tag{5.2}$$

yields a method which is exact for the invariant measure ($\rho_\infty^h = \rho_\infty$), i.e., the left-hand side in (1.6) is zero, even-though the approximation has only weak order 2. Notice also that truncating (5.2) after the h^{p-1} term and applying the Euler–Maryuama yields a scheme of order p for the invariant measure.

5.2. 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.,

$$X_{n+1} = X_n + (1-\theta)(f + hf_1)(X_n) + \theta(f + hf_1)(X_{n+1}) + \sqrt{h}\sigma\xi_n. \tag{5.3}$$

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

Proposition 5.1. *Let $E = \mathbb{R}^d$ or $\mathbb{T}^d$. Consider the numerical method* (5.3) *applied to* (2.9), *where*

$$f_1 = -(1-2\theta)\left(\frac{1}{2}f'f + \frac{\sigma^2}{4}\Delta f\right). \tag{5.4}$$

Assume Assumption 2.1 *and* (2.7) *for $E = \mathbb{T}^d$, and the hypotheses of Theorem* 4.3 *for $E = \mathbb{R}^d$, respectively. If* (5.3) *is ergodic, then it has order $r = 2$ for the invariant measure in* (1.6).

Remark 5.1. In [1], a modified weak order two θ scheme was constructed for general systems of SDEs with noncommutative noise. In the context of additive noise (2.9) it has the form

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

It can be observed that both the drift and diffusion functions are modified in contrast to the scheme (5.3), where only the drift function is modified. Notice that for $\theta = 1/2$, we have $f_1 = 0$ in (5.4) which is not surprising because in this case, the θ-method has weak order two of accuracy.

Applying the recursive procedure of Theorem 3.3, we may next derive a modification of the θ method of order 3.

PROPOSITION 5.2. *Let $E = \mathbb{R}^d$ or $\mathbb{T}^d$. Consider the Euler–Maruyama method applied to the modified SDE defined by $dX = (f + hf_1 + h^2 f_2)dt + \sigma dW$, i.e.,*

$$X_{n+1} = X_n + hf(X_n) + h^2 f_1(X_n) + h^3 f_2(X_n) + \sqrt{h}\xi_n, \tag{5.5}$$

where $f = -\nabla V$, f_1 is defined in (5.4) *with $\theta = 0$ and f_2 is defined by*

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

Assume Assumption 2.1 *and* (2.7) *for $E = \mathbb{T}^d$ and the hypotheses of Lemma* 4.6 *for $E = \mathbb{R}^d$, respectively. If* (5.5) *is ergodic, then it has order $r = 3$ for the invariant measure in* (1.6).

The proof of Proposition 5.2 is postponed until the appendix.

Remark 5.2. We highlight that integrators with arbitrarily higher order for the invariant measure could be constructed analogously using Theorem 3.3. The statement of Proposition 5.2 can be generalized to the θ-method (2.16) and yield again an order 3 method for the invariant measure, but the calculation becomes rather tedious. In the linear case (5.1), the obtained scheme reduces to

$$\begin{aligned} X_{n+1} = x_n &- \left(h\gamma + (1-2\theta)h^2\frac{\gamma^2}{2} + (1-2\theta)^2 h^3 \frac{\gamma^3}{2}\right)\big((1-\theta)X_n + \theta X_{n+1}\big) \\ &+ \sigma\sqrt{h}\xi_n. \end{aligned} \tag{5.6}$$

For $\theta = 1/2$, it coincides with the standard θ-method (2.16) which is not surprising because it samples the invariant measure exactly in this linear context [16].

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

6. Numerical experiments. In this section, we illustrate numerically our main results. We consider first the linear case (5.1), where $V(x) = x^2/2$, and compare the Euler–Maruyama method and the modifications of orders 2 (Proposition 5.2, $\theta = 0$) and 3 (Proposition 5.2, $\theta = 0$). In Figure 1, we plot the error $e(\phi, h)$ defined in (1.6) for $\phi(x) = x^2$ (second moment error) and many different stepsizes h. In theory, computing one long trajectory suffices, but in practice computing several long trajectories also allows one 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,[4] $X_0 = -2$. 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

[4]Recall that the choice of the initial condition has no influence on the numerical ergodic average in (1.6).

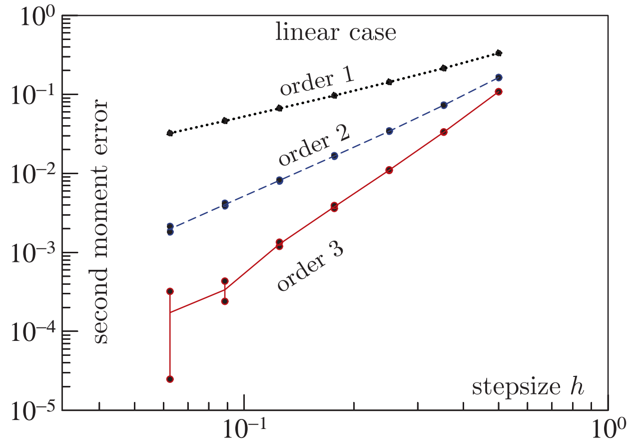

FIG. 1. *Linear case ($V(x) = x^2/2$). Euler–Maruyama method (order 1) and modifications of orders 2 and 3. Error for the second moment $\int_{\mathbb{R}} x^2\rho(x)dx$ versus time stepsize h obtained using 10 trajectories on a long time interval of length $T = 10^8$. The vertical bars indicate the standard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

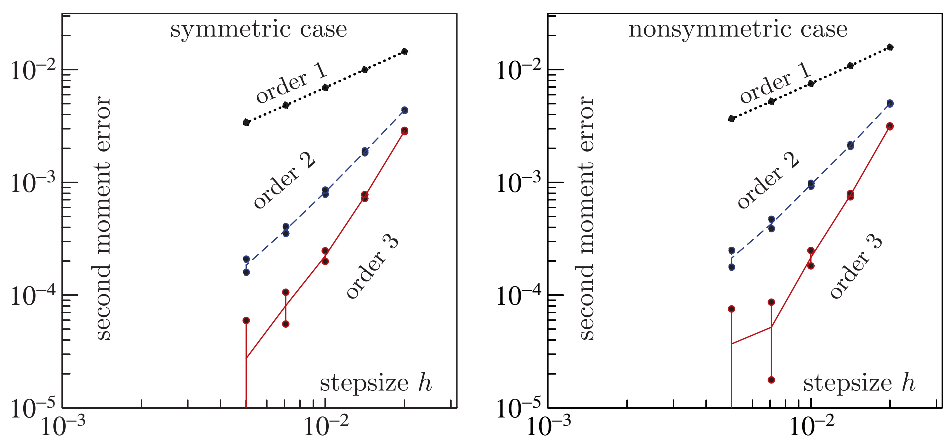

FIG. 2. *Nonlinear problem with double-well potential. Left picture: $V(x) = (1-x^2)^2$ (symmetric). Right picture: $V(x) = (1-x^2)^2 - x/2$ (nonsymmetric). Euler–Maruyama method (order 1) and modifications of orders 2 and 3. Error for the second moment $\int_{\mathbb{R}} x^2\rho(x)dx$ versus time stepsize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

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 nonsymmetric (right picture). Again, we observe the expected lines of slopes $1, 2, 3$, which corroborate Propositions 5.1 and 5.2.

We finally consider the case of Brownian dynamics (2.9) for the following two-dimensional quartic potential:

$$V(x) = (1-x_1^2)^2 + (1-x_2^2)^2 + \frac{x_1x_2}{2} + \frac{x_2}{5}. \tag{6.1}$$

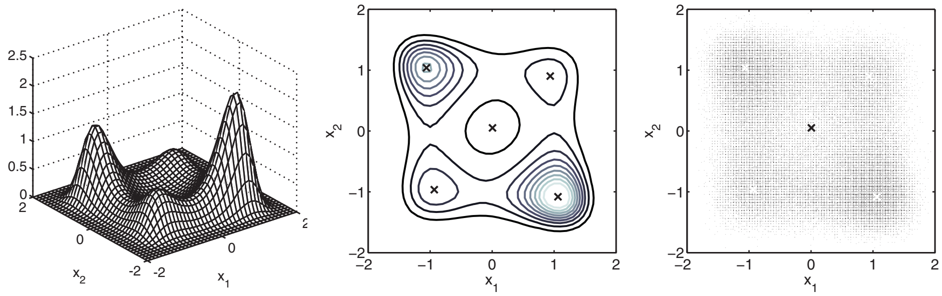

FIG. 3. *Two-dimensional problem* (2.9)–(6.1). *Left picture: Three-dimensional plot of the Gibbs density* (3.2). *Middle picture: Ten level curves of the Gibbs density are represented in solid lines (the five extrema are represented with crosses). Right picture: A numerical trajectory* $\{X_n\}$ *of the scheme* (6.3) *(with* $h = 0.02$, $T = 2 \cdot 10^3$*).*

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 (3.2) 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 (6.3) (discussed 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 (5.3)–(5.4) is not convenient in general for multidimensional systems and can be computational expensive, we introduce the following Runge–Kutta type scheme for (2.9)

$$
\begin{aligned}
Y_1 &= X_n + \sqrt{2}\sigma\sqrt{h}\xi_n, \\
Y_2 &= X_n - \frac{3}{8}hf(Y_1) + \frac{\sqrt{2}}{4}\sigma\sqrt{h}\xi_n, \\
X_{n+1} &= X_n - \frac{1}{3}hf(Y_1) + \frac{4}{3}hf(Y_2) + \sigma\sqrt{h}\xi_n,
\end{aligned}
\tag{6.2}
$$

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 (2.14) of the schemes (6.2) and (5.3)–(5.4) coincide up to order 2, i.e., they have the same operators A_0, A_1 and thus the same order 2 in (1.6) for the invariant measure, and the same weak order 1. This is detailed in the appendix (see Proposition 8.1).

Our investigations indicate that there does not exist a similar Runge–Kutta type approximation of the scheme (5.5) with only 3 evaluations of the function f per timestep. We thus propose the following Runge–Kutta type method which has order 2 in (1.6) for general nonlinear multidimensional problems (2.9), but order 3 for linear problems,

$$
\begin{aligned}
Y_1 &= X_n + \sigma\sqrt{h}\xi_n, \\
Y_2 &= X_n - \frac{h}{2}f(Y_1) + \frac{\sigma}{2}\sqrt{h}\xi_n, \\
Y_3 &= X_n + 3hf(Y_1) - 2hf(Y_2) + \sigma\sqrt{h}\xi_n, \\
X_{n+1} &= X_n - \frac{3}{2}hf(Y_1) + 2hf(Y_2) + \frac{1}{2}hf(Y_3) + \sigma\sqrt{h}\xi_n,
\end{aligned}
\tag{6.3}
$$

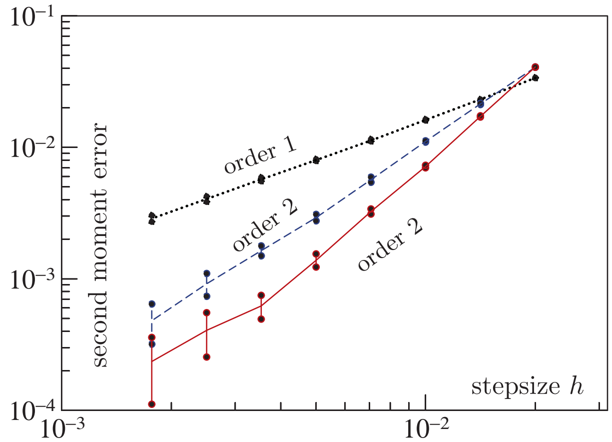

FIG. 4. *Two-dimensional problem* (2.9)–(6.1). *Errors for* $\phi(x) = x^2 + y^2$ *for the Euler–Maruyama method (order* 1*), and the modifications* (6.2) *(order* 2*) and* (6.3) *(order* 2 *but* 3 *for linear problems) with* $T = 10^7$.

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 (6.2) and (6.3). We observe the expected lines of slope $1, 2$. Notice that the error constant for the variant (6.3) is about twice smaller than the error for (6.2). The results for the scheme (5.3) are not included in this plot but are nearly identical to that of (6.2).

7. Conclusion. To achieve high order of convergence in sampling the invariant measure of ergodic nonlinear systems of SDEs, we have proved that the usual approach of using a high weak order method is not necessary. We presented a general methodology based on modified differential equations and inspired by backward error analysis which permits one to construct arbitrarily high order methods for approximating the invariant measure of ergodic SDEs, while their standard weak order remains low on short time intervals. The approach was illustrated with several high order integrators applied to Brownian dynamics. In [2], we shall analyze specifically the case of splitting methods for Langevin dynamics to investigate their order of accuracy in sampling the invariant measure, again independently of their standard weak order of convergence.

In [21], it is shown in the nonglobally Lipschitz case, relevant in most applications, that explicit SDE integrators can still be applied successfully by introducing and justifying theoretically *the concept of rejecting exploding trajectories*. This approach is applied in [22] to weak methods with high order for the efficient calculation of ergodic limits of Langevin-type equations. This is done taking advantage of the exponentially fast convergence to ergodic limits, which allows one to consider relatively short time interval trajectories rather than a single long one. This approach could be extended straightforwardly to the new class of methods proposed here, with high order for the approximation of invariant measures but a low standard weak order.

8. Appendix. We provide in this appendix the proofs of Propositions 5.1, 5.2, 8.1.

Proof of Proposition 5.1. Consider the weak Taylor expansion (2.14) for the θ method. Applying Lemma 3.2 to each differential operator of order greater than 1 in

A_1 given in (2.17) and using $f = -\nabla V$, we obtain

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

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

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

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

where we use the notation $\langle u\rangle = \int_E u(x)\rho_\infty(x)dx$, the sums are for $i,j = 1,\dots,d$, and e_i is the canonical basis of $\mathbb{R}^d$. Using the above identities, a straightforward calculation then yields that f_1 in (5.4) satisfies $\langle A_1\phi\rangle = \langle f_1\cdot\nabla\phi\rangle$, equivalently, $A_1^*\rho_\infty = \operatorname{div}(f_1\rho_\infty)$. Theorem 4.4 (for $E = \mathbb{R}^d$) and Theorem 3.3 (for $E = \mathbb{T}^d$) conclude the proof. ∎

Proof of Proposition 5.2. Consider the weak Taylor expansion (2.14) for the modified θ method (5.3) ($\theta = 0$). We have $A_0 = \mathcal{L}$ because the method has weak order 1, and by the construction of Theorem 3.3, $A_1^*\rho_\infty = 0$. 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 3.2 (see the proof of Proposition 5.1) then yields

$$\begin{aligned}
&\left\langle \sigma^2\phi''(f'e_i,f'e_i)\right\rangle\\
&\quad = \left\langle \phi'(-\sigma^2 f''(e_i,f'e_i) - f'\nabla(\sigma^2\operatorname{div} f + \|f\|^2))\right\rangle\\
&\langle \phi''(f,f'f)\rangle\\
&\quad = \left\langle -\phi'\left(f'f'f + f''(f,f) + (\operatorname{div} f)f'f + \frac{2}{\sigma^2}\|f\|^2 f'f\right)\right\rangle\\
&\langle \phi''(f,f''(e_i,e_i))\rangle\\
&\quad = \left\langle -\phi'\left(f'''(f,e_i,e_i) + (\operatorname{div} f)f''(e_i,e_i) + \frac{2}{\sigma^2}\|f\|^2 f''(e_i,e_i)\right)\right\rangle\\
&\left\langle \sigma^4\phi^{(3)}(f''(e_i,e_i),e_j,e_j)\right\rangle\\
&\quad = \big\langle -\phi'(\sigma^4 f^{(4)}(e_i,e_i,e_j,e_j) + 4\sigma^2 f'''(f,e_i,e_i)\\
&\qquad + 2(\operatorname{div} f)f''(e_i,e_i) + 4\|f\|^2 f''(e_i,e_i))\big\rangle\\
&\left\langle \sigma^2\phi^{(3)}(f'f,e_i,e_i)\right\rangle
\end{aligned}$$

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

$$\left\langle \sigma^2 \phi^{(3)}(f, f'e_i, e_i) \right\rangle = \left\langle -\sigma^2 \phi''(f, f''(e_i, e_i)) - \sigma^2 \phi''(f'e_i, f'e_i) - 2\phi''(f'f, f) \right\rangle$$

$$\left\langle \sigma^4 \phi^{(4)}(e_i, e_i, f'e_j, e_j) \right\rangle = \left\langle -\sigma^4 \phi^{(3)}(f''(e_i, e_i), e_j, e_j) - 2\sigma^2 \phi^{(3)}(f'f, e_i, e_i) \right\rangle$$

$$\left\langle \sigma^2 \phi^{(4)}(f, f, e_i, e_i) \right\rangle = \left\langle -2\phi^{(3)}(f, f, f) - 2\sigma^2 \phi^{(3)}(f, f'e_i, e_i) \right\rangle$$

$$\left\langle \sigma^4 \phi^{(5)}(f, e_i, e_i, e_j, e_j) \right\rangle = \left\langle -2\sigma^2 \phi^{(4)}(f, f, e_i, e_i) - \sigma^4 \phi^{(4)}(e_i, e_i, f'e_j, e_j) \right\rangle$$

$$\left\langle \sigma^6 \phi^{(6)}(e_i, e_i, e_j, e_j, e_k, e_k) \right\rangle = \left\langle -2\sigma^4 \phi^{(5)}(f, e_i, e_i, e_j, e_j) \right\rangle,$$

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 $\nabla \operatorname{div} f = \Delta f$ and $\nabla(\|f\|^2) = 2f'f$ in the first equality and we obtain $A_2^* \rho_\infty = \operatorname{div}(f_2 \rho_\infty)$. Theorem 4.7 (for $E = \mathbb{R}^d$) and Theorem 3.3 (for $E = \mathbb{T}^d$) conclude the proof. ∎

PROPOSITION 8.1. *Consider the method* (6.2) *for* (2.9) *on the space* $E = \mathbb{T}^d$ *(assuming assumptions* (2.1) *and* (2.7)*) or* $E = \mathbb{R}^d$ *(assuming the hypotheses of Theorem* 4.3*), and assume that it is ergodic. Then,* (6.2) *has order* $r = 2$ *in* (1.6) *for the invariant measure.*

Proof. We justify the construction of the derivative free implementation (6.2) of the scheme (5.3) ($\theta = 0$). 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,

$$X_1 = X_0 + h \left(\sum_{i=1}^s b_i \right) f + \sqrt{h} \sigma \xi_n + h^{3/2} \sigma \left(\sum_{i=1}^s b_i \overline{c}_i \right) f' \xi_n + h^2 \left(\sum_{i=1}^s b_i c_i \right) f'f + \frac{h^2 \sigma^2}{2} \left(\sum_{i=1}^s b_i \overline{c}_i^2 \right) f''(\xi_n, \xi_n) + \cdots$$

and we deduce the differential operators in the weak Taylor expansion (2.14),

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

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 $A_1\phi = -(\frac{1}{2}f'f + \frac{\sigma^2}{4}\Delta f)\cdot\nabla\phi$ as for the scheme (5.3) ($\theta = 0$) and thus the same order two for the invariant measure. ∎

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 (2012), pp. 1800–1823.

[2] A. Abdulle, G. Vilmart, and K. C. Zygalakis, *Long Time Accuracy of Lie-Trotter Splitting Methods for Second Order Stochastic Dynamics*, MATHICSE Technical report 18 2014, École Polytechnique Fédérale de Lausanne, 2014.

[3] V. Bally and D. Talay, *The law of the Euler scheme for stochastic differential equations.* I. *Convergence rate of the distribution function*, Probab. Theory Related Fields, 104 (1996), pp. 43–60.

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

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

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

[7] C. W. Gardiner, *Handbook of Stochastic Methods*, 2nd ed., Springer-Verlag, Berlin, 1985.

[8] E. Hairer, C. Lubich, and G. Wanner, *Geometric Numerical Integration. Structure-Preserving Algorithms for Ordinary Differential Equations*, 2nd ed., Springer Ser. Comput. Math. 31, Springer-Verlag, Berlin, 2006.

[9] R. Hasminskii, *Stochastic Stability of Differential Equations*, Sijthoff and Noordhoff, The Netherlands, 1980.

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

[11] 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 (2011), pp. 1563–1576.

[12] M. Kopec, *Weak Backward Error Analysis for Langevin Process*, preprint, 2013.

[13] M. Kopec, *Weak Backward Error Analysis for Overdamped Langevin Processes*, IMA J. Numer. Anal., doi:10.1093/imanum/dru016, 2014.

[14] B. Leimkuhler and C. Matthews, *Rational construction of stochastic numerical methods for molecular sampling*, Appl. Math. Res. Express AMRX, 2013 (2013), pp. 34–56.

[15] B. Leimkuhler and S. Reich, *Simulating Hamiltonian Dynamics*, Cambridge Monogr. Appl. Comput. Math. 14, Cambridge University Press, Cambridge, 2004.

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

[17] J. Mattingly, A. Stuart, and D. Higham, *Ergodicity for SDES and approximations: Locally Lipschitz vector fields and degenerate noise*, Stochastic Process. Appl., 101 (2002), pp. 185–232.

[18] 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 (2010), pp. 552–577.

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

[20] G. Milstein and M. Tretyakov, *Stochastic Numerics for Mathematical Physics*, Sci. Comput., Springer-Verlag, Berlin, 2004.

[21] G. N. Milstein and M. V. Tretyakov, *Numerical integration of stochastic differential equations with nonglobally Lipschitz coefficients*, SIAM J. Numer. Anal., 43 (2005), pp. 1139–1154.

[22] G. N. Milstein and M. V. Tretyakov, *Computing ergodic limits for Langevin equations*, Phys. D, 229 (2007), pp. 81–95.

[23] H. Risken, *The Fokker–Planck Equation*, Springer Ser. Synergetics 18, Springer-Verlag, Berlin, 1989.

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

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

[26] 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.

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

[28] D. Talay and L. Tubaro, *Expansion of the global error for numerical schemes solving stochastic differential equations*, Stoch. Anal. Appl., 8 (1990), pp. 483–509.

[29] K. C. Zygalakis, *On the existence and the applications of modified equations for stochastic differential equations*, SIAM J. Sci. Comput., 33 (2011), pp. 102–130.