SIAM J. NUMER. ANAL.
Vol. 53, No. 1, pp. 1–16

LONG TIME ACCURACY OF LIE–TROTTER SPLITTING METHODS FOR LANGEVIN DYNAMICS*

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

Abstract. A new characterization of sufficient conditions for the Lie–Trotter splitting to capture the numerical invariant measure of nonlinear ergodic Langevin dynamics up to an arbitrary order is discussed. Our characterization relies on backward error analysis and needs weaker assumptions than assumed so far in the literature. In particular, neither high weak order of the splitting scheme nor symplecticity are necessary to achieve high order approximation of the invariant measure of the Langevin dynamics. Numerical experiments confirm our theoretical findings.

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

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

DOI. 10.1137/140962644

1. Introduction.

In many applications (see [25, 16, 5, 19] and references therein), one is interested in simulating the invariant measure of a stochastic differential equation (SDE) by running a numerical scheme that approximates its time dynamics. In particular, one uses the numerical trajectories to construct an empirical measure either by averaging one single long trajectory or by averaging over many realizations to obtain a finite ensemble average (see, for example, [25]). One immediately is faced with the question regarding the quality of such an approximation. Another issue is the ergodicity of the numerical approximation itself. This second question is still an area of active research [20, 33, 28, 31] but is not addressed here.

In this paper, we focus on the case of second order stochastic dynamics generated by the Langevin equation, of the form

$$dq(t) = M^{-1}p(t)dt, \tag{1.1a}$$

$$dp(t) = (-\nabla V(q(t)) - \gamma p(t))dt + \sqrt{2\beta^{-1}\gamma}M^{1/2}dW(t), \tag{1.1b}$$

where the scalar function $V : \mathbb{R}^d \to \mathbb{R}$ denotes the potential, the mass matrix M is symmetric positive definite, $\gamma > 0$ (the friction), $\beta > 0$ (related to the temperature) are fixed constants, $p(t), q(t) \in \mathbb{R}^d$, and $W(t)$ denotes a standard d-dimensional Wiener process. For simplicity, we assume that the mass matrix is the identity matrix $M = I$, but we emphasize that our analysis would apply straightforwardly to general

*Received by the editors March 28, 2014; accepted for publication (in revised form) September 10, 2014; published electronically January 8, 2015.
http://www.siam.org/journals/sinum/53-1/96264.html

†Mathematics Section, École Polytechnique Fédérale de Lausanne, 1015 Lausanne, Switzerland (Assyr.Abdulle@epfl.ch). The research of this author was partially supported by Fonds National Suisse project 200021_140692.

‡Université de Genève, Section de mathématiques, CP 64, 1211 Genève 4, Switzerland, and École Normale Supérieure de Rennes, INRIA Rennes, IRMAR, CNRS, F-35170 Bruz, France (Gilles.Vilmart@ens-rennes.fr). The research of this author was partially supported by Fonds National Suisse Project 200020_144313/1.

§Mathematical Sciences, University of Southampton, Southampton SO17 1BJ, UK (k.zygalakis@soton.ac.uk).

positive definite matrices M. Under appropriate smoothness and growth assumptions on the Hamiltonian energy (see, e.g., [21, 14]),

$$H(p,q) = \frac{1}{2}p^T p + V(q), \tag{1.2}$$

one can show that the dynamics generated by (1.1) are ergodic, i.e., for arbitrary initial conditions $q(0), p(0)$ (assumed deterministic for simplicity), and for all smooth test function $\phi \in \mathcal{C}_P^\infty(\mathbb{R}^{2d}, \mathbb{R})$ with derivatives of all orders with polynomial growth, the time average of the trajectories of (1.1) satisfy with probability 1

$$\lim_{T\to\infty} \frac{1}{T}\int_0^T \phi(p(t),q(t))dt = \int_{\mathbb{R}^d\times\mathbb{R}^d} \phi(p,q)d\mu(p,q).$$

The invariant measure μ is characterized by the Gibbs density function

$$\rho_\infty(p,q) = Ze^{-\beta H(p,q)}, \tag{1.3}$$

where Z is the normalization constant, which ensures that $\int_{\mathbb{R}^d\times\mathbb{R}^d} \rho_\infty(p,q)dpdq = 1$. Similarly to [4], we consider the following class of schemes based on the Lie–Trotter splitting:

$$X_{n+1} = \Phi_h \circ \Theta_{h,n}(X_n), \tag{1.4}$$

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

$$dq(t) = p(t)dt, \qquad dp(t) = -\nabla V(q(t))dt, \tag{1.5}$$

while $\Theta_{h,n}$ is a numerical integrator for the stochastic part

$$dq(t) = 0, \tag{1.6a}$$

$$dp(t) = -\gamma p(t)dt + \sqrt{2\beta^{-1}\gamma}dW(t). \tag{1.6b}$$

The exact solution of (1.6) is given by the variation of constants formula,

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

The trajectories of the above stochastic integral itself need not be simulated. Indeed, the properties of the stochastic integral [26] imply that the solution $p(t)$ of (1.6) given $p(0)$ is Gaussian and thus characterized by its mean and covariance. Using this, we obtain the following flow $(q_{n+1}, p_{n+1}) = \Theta_{h,n}(q_n, p_n)$ for (1.6) which is exact in law,[1]

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

where $\xi_n \sim \mathcal{N}(0, I)$ are independent d-dimensional Gaussian random variables.

The aim of this paper is to analyze splitting schemes such as (1.4) for a variety of different deterministic integrators Φ_h used for solving (1.5), for the numerical approximation of the invariant measure, i.e., we quantify the error

$$e(\phi,h) := \lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^N \phi(p_n,q_n) - \int_{\mathbb{R}^d\times\mathbb{R}^d}\phi(p,q)\rho_\infty(p,q)dpdq \tag{1.7}$$

[1] This means that the exact solution $(p(h), q(h))$ and the numerical solution $\Theta_{h,n}(p(0), q(0))$ have the same law of probability for a given state $(p(0), q(0))$.

for all smooth test function $\phi \in \mathcal{C}_P^\infty(\mathbb{R}^{2d}, \mathbb{R})$. We say that the method has order r for sampling the invariant measure if $|e(\phi, h)| \leq Ch^r$ for all h small enough, where C is independent of h, but depends on the smooth test function ϕ. In particular, using recent advances in the theory of modified integrators for SDEs and ODEs [2, 8] and backward error analysis for SDEs [35, 9, 14, 15, 3], we provide a new characterization of sufficient conditions for the Lie–Trotter splitting to capture the numerical invariant measure of nonlinear ergodic Langevin dynamics up to an arbitrary order.

It is well known that controlling the weak error $|\mathbb{E}(\phi(X_n)) - \mathbb{E}(\phi(X(t_n))|$ up to a given order allows to control the accuracy error $|e(\phi, h)|$ for the invariant measure up to the same order [34, 22]. However, the converse is not true, and new sufficient conditions independent of the weak order of accuracy for a numerical method to approximate with a high order of accuracy the invariant measure of ergodic SDEs have recently been derived in [3]. The accuracy of the numerical computation of the invariant measure for Langevin dynamics has recently been studied in a number of papers [4, 16, 17]. In [4], the splitting method (1.4) is proposed and analyzed, while several other splitting methods with high order for the invariant measure are considered and studied in [16, 17]. It is shown in [4] that if Φ_h is a variational integrator (equivalently a symplectic scheme) which preserves the energy up to order r, then the Lie–Trotter splitting (1.4) has order r of accuracy for the invariant measure, although the standard weak error is only of order one. This is another example illustrating the fact that a high weak order of accuracy is not necessary to achieve high order for the invariant measure.

In this paper, we provide weaker assumptions to achieve order r in the approximation of the numerical invariant measure using the splitting (1.4). Our results show that to achieve order r for the invariant measure by the splitting (1.4), neither an order r of accuracy in the weak sense nor an order r of accuracy in the deterministic sense for Φ_h in (1.4) is needed. This is also related to the results in [6, 7] which provide a detailed analysis of the accuracy of the numerical invariant measure in the case of a linear drift, i.e., a quadratic potential $V(q)$. It is shown there that among a certain class of stochastic Runge–Kutta methods applied to Langevin dynamics with a linear drift (which does not include the Lie–Trotter splitting (1.4)), the *(stochastic) implicit midpoint rule is the only (stochastic) Runge–Kutta method with a nonsingular tableau matrix that reproduces the exact stationary distribution for all values of damping.* See also related results in [19]. Our conclusion for Lie–Trotter splitting methods (1.4) is different: we prove in particular that it samples exactly the invariant measure for a linear drift if the deterministic integrator Φ_h is a consistent symmetric Runge–Kutta method.

Our new characterization is expressed in terms of functions arising in the modified differential equation for the Hamiltonian system. We thus heavily rely on backward error analysis for our finding. The first error term in the error estimate for the numerical invariant measure can also be expressed in term of functions arising in the modified differential equation. This is also a new expression for error estimates à la Talay and Tubaro [34].

To simplify the presentation of our results, we focus on the globally Lipschitz case and we assume that the potential V is a $\mathcal{C}^\infty$ function where ∇V has bounded derivatives of any order (which implies the global Lipschitzness of ∇V) and satisfies the following standard growth condition, which guarantees that (1.1) is ergodic:

$$q^T \nabla V(q) \geq C_1 q^T q - C_2 \qquad \text{for all } q \in \mathbb{R}^d, \tag{1.8}$$

for some constants $C_1, C_2 > 0$. We further assume that the numerical flow Φ_h is globally Lipschitz in $\mathbb{R}^{2d}$,

$$\|\Phi_h(x_1) - \Phi_h(x_2)\| \leq L\|x_1 - x_2\| \tag{1.9}$$

for all $x_1, x_2 \in \mathbb{R}^{2d}$ and all stepsize h small enough. In a more general nonglobally Lipschitz setting, important in most applications, two implicit schemes are studied in [14], and the implicitness is used to guarantee the boundedness of the numerical moments along time. Using implicit schemes for the Hamiltonian part is one way to extend our results in the nonglobally Lipschitz setting. Alternatively, notice that the class of integrators considered in this paper (which are not assumed implicit) could still be applied rigorously in the nonglobally Lipschitz setting by following the methodology of *rejecting exploding trajectories* proposed in [24] and applied to Langevin-type ergodic systems in [25].

We close this introduction by mentioning that we focus here only on the discrepancy error (1.7) and do not discuss the Monte Carlo error that can't be avoided in practice. One could use the so-called multilevel Monte Carlo method (MLMC) [10], which is a popular technique for reducing the variance. Indeed, it applies not only to first weak order integrators for SDEs but also to higher orders weak schemes, as shown in [11], where the antithetic MLMC has been introduced. Also in the context of stiff SDEs, it was shown in [1] that applying a weak second order method at the finer level of the MLMC permits us to significantly improve the error constant. Finally we mention the schemes based on Markov chain Monte Carlo methods (see, e.g., the survey [29]), for which the bias error is reduced to zero, thanks to an appropriate acceptance/rejection criteria at each step.

The paper is organized as follows. In section 2, we present the basic ingredients of backward error analysis and the framework in [3] that allows us to characterize the long time behavior of numerical integrators of ergodic SDEs. In section 3, we prove the main result of this paper, namely, sufficient conditions for a Lie–Trotter splitting scheme to capture the numerical invariant measure of nonlinear ergodic Langevin dynamics up to an arbitrary order. We also relate this characterization to geometric properties of the deterministic integrator and present in section 4 various numerical investigations, for both linear and nonlinear problems that corroborate such theoretical findings.

2. Preliminaries. In section 2.1, we discuss some important concepts related to backward error analysis for ODEs, while in section 2.2 we discuss the standard framework of backward Kolmogorov equation and weak Taylor expansions. Finally, in section 2.3, we recall the order conditions framework from [3] for the approximation of the invariant measure.

2.1. Deterministic backward error analysis. Backward error analysis is a powerful tool for the analysis of numerical integrators for differential equations [30, 18, 13]. In particular, it is the main ingredient for the proof of the good energy conservation (without drift) of symplectic Runge–Kutta methods when applied to deterministic Hamiltonian systems over exponentially long time intervals. In our context it is useful to characterize the "generator" of the deterministic method Φ_h of (1.5). Given a consistent integrator $y_{n+1} = \Phi_h(y_n)$ for a system of ODEs

$$\frac{dy(t)}{dt} = f(y(t)), \tag{2.1}$$

the idea of backward error analysis is to search for a modified differential equation written as a formal series in powers of the stepsize h,

$$\frac{d\widetilde{y}}{dt} = f(\widetilde{y}) + hf_1(\widetilde{y}) + h^2 f_2(\widetilde{y}) + \cdots, \quad \widetilde{y}(0) = y_0, \tag{2.2}$$

such that (formally) $y_n = \widetilde{y}(t_n)$, where $t_n = nh$ (in the above differential equation, we omit the time variable for brevity). This means that the numerical solution can be interpreted as the exact solution of a modified ODE. The vector fields f_j in the series can be constructed for all reasonable integrators by induction on j [18, 13], setting $f_0 = f$. Notice, however, that this series diverges in general for nonlinear systems and needs to be truncated. Considering the truncated modified ODE at order s

$$\frac{d\widetilde{y}}{dt} = f(\widetilde{y}) + hf_1(\widetilde{y}) + h^2 f_2(\widetilde{y}) + \cdots + h^s f_s(\widetilde{y}), \quad \widetilde{y}(0) = y_0, \tag{2.3}$$

we have $y_n = \widetilde{y}(t_n) + \mathcal{O}(h^{s+1})$ for $h \to 0$ for bounded times $t_n = nh \le T$. We note that the flow $\widetilde{\Phi}_h(y)$ of the modified differential equation (2.3) can be expressed as

$$\widetilde{\Phi}_h = \left(\sum_{k=0}^{M} \frac{h^k \widetilde{\mathcal{L}}_D^k}{k!} \right) I + \mathcal{O}(h^{M+1}), \qquad \widetilde{\mathcal{L}}_D = F_0 + hF_1 + h^2 F_2 + \cdots + h^s F_s, \tag{2.4}$$

for all $M \ge 0$, where I is the identity map, $F_j\phi = f_j \cdot \nabla\phi$, $j = 0, \ldots, s$, and $f_0 = f$. This is a particular case ($\phi(y) = Iy$) of the general result [13, sect. III.5.1]

$$\phi \circ \widetilde{\Phi}_h = \left(\sum_{k=0}^{M} \frac{h^k \widetilde{\mathcal{L}}_D^k}{k!} \right) \phi + \mathcal{O}(h^{M+1}) \tag{2.5}$$

for all smooth test function ϕ. The constants symbolized by $\mathcal{O}$ in (2.4), (2.5) are independent of $h \to 0$ but depend on M, s, and ϕ.[2]

2.2. Order conditions for the numerical invariant measure. A numerical method applied to (1.1) is called ergodic if it has a unique invariant law μ^h with finite moments of any order and

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \phi(p_n, q_n) = \int_{\mathbb{R}^d \times \mathbb{R}^d} \phi(p,q) d\mu^h(p,q).$$

The aim of this subsection is to briefly describe the conditions that the numerical method applied to (1.1) should satisfy in order to approximate the invariant measure μ with an error (1.7) of size $\mathcal{O}(h^r)$. These conditions relate directly to the expansion of one-step numerical expectations in powers of h. In particular, let $X_1 = (p_1, q_1)$, $x = (p, q)$, and

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

We assume that it is possible to expand U in powers of h to obtain the expansion

$$U(x,h) = \phi(x) + hA_0\phi(x) + h^2 A_1\phi(x) + h^3 A_2\phi(x) + \cdots, \tag{2.6}$$

where $A_i, i = 0, 1, \ldots$, are linear differential operators with coefficients depending on the choice of the numerical integrator and depending smoothly on the Hamiltonian function $H(p,q)$ and its derivatives. For a consistent method (i.e., of weak order at least one), we have $A_0 = \mathcal{L}$, where $\mathcal{L}$ denotes the generator of (1.1) given by

$$\mathcal{L} := p \cdot \nabla_q - (\nabla V(q) + \gamma p) \cdot \nabla_p + \beta^{-1}\gamma\Delta_p, \tag{2.7}$$

[2]For all h small enough, the sum in (2.4) (and (2.5)) can be shown to converge for $M \to \infty$ in the case of analytic vector fields f_j (and analytic test functions ϕ), which permits us to remove the $\mathcal{O}$ remainder.

where ∇_p, ∇_q denote the gradient differential operators with respect to p and q, and Δ_p is the Laplacian with respect to the p variable. Truncating the expansion (2.6) yields rigorous estimates (see, for instance, [32]).

Example 2.1. Consider the Euler–Maruyama method for (1.1) defined as

$$\begin{aligned} q_{n+1} &= q_n + hp_n, \\ p_{n+1} &= p_n - h(\nabla V(q_n) + \gamma p_n) + \sqrt{2\beta^{-1}\gamma h}\xi_n, \end{aligned}$$

where $\xi_n \sim \mathcal{N}(0, I)$ are independent d-dimensional Gaussian random variables. A straightforward calculation yields that A_1 is a differential operator of order 4,

$$A_1\phi = \frac{1}{2}\phi''(f,f) + \beta^{-1}\gamma(\Delta_p\phi)'f + \frac{\beta^{-2}\gamma^2}{2}\Delta_p^2\phi \quad \text{with} \quad f = \begin{pmatrix} -\nabla V(q) - \gamma p \\ p \end{pmatrix},$$

where $\phi''(\cdot,\cdot)$ denotes the second differential of ϕ (bilinear form), and ϕ' is the first derivative (linear form). We refer to [2, 35] for other examples of such weak Taylor series calculations.

We can now state the conditions which ensure that μ^h approximates μ to order r. We have the following theorem, which involves the adjoint operators A_j^* of the differential operators A_j involved in (2.6), defined as

$$\int_{\mathbb{R}^{2d}} A_j^*\phi_1(x)\phi_2(x)dx = \int_{\mathbb{R}^{2d}} \phi_1(x)A_j\phi_2(x)dx$$

for all smooth test functions ϕ_1, ϕ_2 with compact support on $\mathbb{R}^{2d}$. The following result is proved in [3].

THEOREM 2.1. *For the ergodic system* (1.1)*, consider a consistent ergodic numerical method satisfying* (2.6) *and*

$$A_j^*\rho_\infty = 0 \quad \textit{for} \quad j = 1, \dots, r-1. \tag{2.8}$$

Then one obtains that the error in (1.7) *satisfies for all test functions* $\phi \in \mathcal{C}_P^\infty(\mathbb{R}^{2d}, \mathbb{R})$

$$e(\phi, h) = h^r \int_0^\infty \int_{\mathbb{R}^d \times \mathbb{R}^d} A_r u(p,q,t)\rho_\infty(p,q)dpdqdt + \mathcal{O}(h^{r+1}), \tag{2.9}$$

where ρ_∞ *is the invariant measure density given in* (1.3)*,* h *is assumed small enough, and* $u(x,t)$ *solves the backward Kolmogorov equation*

$$\frac{\partial u}{\partial t}(p,q,t) = \mathcal{L}u(p,q,t), \qquad t > 0, p, q \in \mathbb{R}^d, \tag{2.10a}$$

$$u(p,q,0) = \phi(p,q), \tag{2.10b}$$

where $\mathcal{L}$ *is the generator of* (1.1) *defined in* (2.7)*.*

Theorem 2.1 generalizes a standard result by Talay and Tubaro [34] and Milstein [22] (see, e.g., [23, Chap. 2.2, 2.3] for a proof) for methods which have weak order r. Here, the order r assumption is replaced by the weaker condition (2.8). An interpretation of (2.8) is the following: the invariant measure μ with density ρ_∞ of (1.1) is preserved up to an $\mathcal{O}(h^r)$ error by a single time step $X_0 \mapsto X_1$ of the integrator. Precisely, if X_0 is a random variable with law μ, then for all test functions $\phi \in \mathcal{C}_P^\infty(\mathbb{R}^{2d}, \mathbb{R})$ and $h \to 0$,

$$\mathbb{E}(\phi(X_1)) = \int_{\mathbb{R}^{2d}} \phi(x)\rho_\infty(x)dx + \mathcal{O}(h^r),$$

where the constant symbolized by $\mathcal{O}$ is independent of h assumed small enough.

Remark 2.1. Assuming that ∇V is smooth with bounded derivatives at all orders, (2.10) has a unique solution for all test function $\phi \in \mathcal{C}_P^\infty(\mathbb{R}^{2d}, \mathbb{R})$. It admits a rigorous Taylor expansion up to arbitrary order p for all h small enough

$$u(x,h) = \phi(x) + h\mathcal{L}\phi(x) + \frac{h^2}{2}\mathcal{L}^2\phi(x) + \cdots + \frac{h^p}{p!}\mathcal{L}^p\phi(x) + \mathcal{O}(h^{p+1}),$$

where the function symbolized in $\mathcal{O}$ is smooth with respect to x with all its derivatives having a polynomial growth, independently of the smallness of the stepsize h. Notice that applying this result to the deterministic problem (2.3), we recover the statement (2.5).

3. Accuracy of Lie–Trotter splitting methods in sampling the invariant measure. In this section, we derive our main results. We first present in section 3.1 sufficient order conditions for the Lie–Trotter splitting to achieve order r for the invariant measure. We discuss the relation of our new characterization with geometric properties such as energy and volume conservation and show in section 3.2 that high order for the invariant measure can be achieved without high order conservation of energy and volume.

3.1. Order conditions to capture the invariant measure to a given order. Using arguments from backward error analysis, we prove the following result for the accuracy of the sampling the invariant measure (1.3) of Langevin dynamics of the form (1.1).

THEOREM 3.1. *Consider the ergodic system* (1.1). *Assume that* Φ_h *is a consistent method for* (1.5) *satisfying* (1.9) *with modified equation from backward error analysis* (2.2). *If the vector fields in* (2.2) *satisfy*

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

then, assuming ergodicity, the Lie–Trotter splitting (1.4) *has order* r *of accuracy for the invariant measure. Precisely, we have for all* $\phi \in \mathcal{C}_P^\infty(\mathbb{R}^d, \mathbb{R})$ *and* $h \to 0$,

$$e(\phi,h) = -h^r \int_0^\infty \int_{\mathbb{R}^d \times \mathbb{R}^d} u(p,q,t) \operatorname{div}(f_r(p,q)\rho_\infty(p,q)) dp dq dt + \mathcal{O}(h^{r+1}), \tag{3.1}$$

where $e(\phi,h)$ *is defined in* (1.7) *and* $u(x,t)$ *is defined in* (2.10). *The constant symbolized in* $\mathcal{O}$ *is independent of the stepsize* h *assumed small enough.*

Proof. Consider the method (1.4) and its corresponding one-step expansion in powers of h

$$\mathbb{E}(\phi(X_1)|X_0 = x) = \phi(x) + h\mathcal{L}\phi(x) + \cdots + h^r A_{r-1}\phi(x) + h^{r+1} A_r \phi(x) + \mathcal{O}(h^{r+1}).$$

In order to prove our theorem, it is enough to show that

$$A_j^* \rho_\infty = 0 \quad \text{for} \quad j = 1, \dots r-1, \qquad A_r^* \rho_\infty = \operatorname{div}(f_r \rho_\infty). \tag{3.2}$$

The result then follows immediately from Theorem 2.1 using the identity

$$\int_{\mathbb{R}^{2d}} A_r u(x,t) \rho_\infty(x) dx = -\int_{\mathbb{R}^{2d}} u(x,t) \operatorname{div}(f_r(x)\rho_\infty(x)) dx. \tag{3.3}$$

We now start with the calculation of A_j. In particular, given a smooth test function ϕ and $x \in \mathbb{R}^{2d}$, using the semigroup property of the Markov process we have

$$\mathbb{E}(\phi(X_1)|X_0 = x) = \mathbb{E}(\phi(\Phi_h \circ \Theta_{h,n})(X_0)|X_0 = x) = e^{h\mathcal{L}_S}(\phi \circ \Phi_h)(x), \tag{3.4}$$

where $e^{h\mathcal{L}_S}\phi$ denotes the exact flow of the Kolmogorov backward equation corresponding to (1.6), with generator $\mathcal{L}_S$ given by

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

We next recall the generator (2.4) of the truncated modified equation (2.3) of the integrator Φ_h

$$\widetilde{\mathcal{L}}_D = F_0 + hF_1 + \cdots + h^r F_r,$$

where we define the differential operators $F_j\phi = f_j \cdot \nabla\phi$ (with $f_0 = f$). We have, using (3.4), applying Remark 2.1 for $\mathcal{L}_S$, and using (2.5) with $M = s = r$,

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

where we denote

$$A_k = \sum_{j=0}^{k+1} \mathcal{L}_S^{k+1-j}\left(\sum_{\substack{1 \le i \le j \\ n_1+n_2+\cdots+n_i = j-i}} \frac{1}{i!(k+1-j)!} F_{n_1}\cdots F_{n_i}\right),$$

where the second sum above is over integers $n_1, \ldots, n_i \ge 0$ and is equal to the identity I when $j = 0$. We obtain for all $k \ge 1$,

$$A_k^*\rho_\infty = \sum_{j=0}^{k+1}\left(\sum_{\substack{1 \le i \le j \\ n_1+n_2+\cdots+n_i = j-i}} \frac{1}{i!(k+1-j)!} F_{n_i}^*\cdots F_{n_1}^*\right)(\mathcal{L}_S^*)^{k+1-j}\rho_\infty.$$

Using $\mathcal{L}_S^*\rho_\infty = 0$ and $F_i^*\rho_\infty = 0$, $i = 1, \ldots r-1$, we see that for $k \le r$, the only possibly nonzero term in the above sum is obtained for $j = r+1, k = r, i = 1$, i.e., $F_r^*\rho_\infty = \operatorname{div}(f_r\rho_\infty)$. Hence, we deduce (3.2), which permits us to conclude the proof. ∎

Remark 3.1. Theorem 3.1 remains valid for the analogous scheme $X_{n+1} = \Theta_{h,n} \circ \Phi_h(X_n)$ with a similar analysis.

An immediate consequence of Theorem 3.1 is the following corollary using the identity

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

COROLLARY 3.2. *Under the assumptions of Theorem* 3.1, *if*

$$\operatorname{div}(f_j) - \beta f_j \cdot \nabla H = 0, \qquad j = 1, \dots, r-1, \tag{3.5}$$

then, assuming ergodicity, the Lie–Trotter splitting (1.4) *has order* r *of accuracy for the invariant measure.*

In addition to the estimates on the bias (1.7), as discussed in [3] for Theorem 2.1, we also obtain the following exponential convergence estimate (3.6) which is classical for weak methods [23, Chap. 2] (see also [33, 20, 25] in the context of nonglobally Lipchitz vector fields).

Remark 3.2. Under the assumptions of Theorem 2.1 or Theorem 3.1 or Corollary 3.2, there exists constants $\lambda, C, K(x) > 0$ such that for all $n \geq 0$,

$$\left| \mathbb{E}(\phi(X_n)) - \int_{\mathbb{R}^d \times \mathbb{R}^d} \phi(p,q)\rho(p,q)dpdq \right| \leq K(x)e^{-\lambda t_n} + Ch^p, \tag{3.6}$$

where $t_n = nh$, the constants $C, K(x), \lambda$ are independent of n, h, and h is assumed small enough ($C, K(x)$ depend on ϕ and $K(x)$ depends on $X_0 = x$).

The simplest way for condition (3.5) to be satisfied is to consider a method Φ_h of order r, since in this case $f_j = 0, j = 1, \dots r-1$. Indeed, in this case, standard results from backward error analysis show that the right-hand side of the modified equation (2.3) reads $f(\widetilde{y}) + h^r f_r(y) + \cdots$ and $f_j = 0,\ j = 1 \dots, r-1$.

Relation with results of [4]. In the paper [4] the authors show that if a variational (equivalently symplectic) integrator Φ_h preserving the energy with order r is used for solving the deterministic dynamics (1.5), then the Lie–Trotter splitting (1.4) has order r of accuracy for the invariant measure. Now a symplectic method applied to a Hamiltonian system is volume preserving hence $\operatorname{div}(f_j) = 0$ for all $j \geq 0$. Furthermore, if the method Φ_h has order r, then it also preserves energy up to order r, i.e., for all vectors y,

$$H(\Phi_h(y)) = H(y) + \mathcal{O}(h^{r+1}) \tag{3.7}$$

this condition is equivalent to $f_j \cdot \nabla H = 0$ for all $1 \leq j \leq r-1$. If both terms in (3.5) are required to vanish independently, we see that we obtain order r of accuracy for the invariant measure under weaker assumptions than in [4]. It is not necessary for the method to be symplectic; only volume preservation up to to order r is required i.e.,

$$\det\left(\frac{\partial \Phi_h(y)}{\partial y}\right) = 1 + \mathcal{O}(h^{r+1}), \tag{3.8}$$

which is equivalent to $\operatorname{div}(f_j) = 0$ for all $1 \leq j \leq r-1$. Note that any deterministic method Φ_h of order r automatically satisfies (3.7) and (3.8) and thus has the desired properties. We will further see in section 3.2 that energy and volume conservation up to order r of the deterministic integrator Φ_h in (1.4) is not necessary to get order r of accuracy for the invariant measure, as one can construct deterministic methods that satisfy (3.5) with $j = r+1$, while conserving energy and volume only up to order r.

Remark 3.3. In general, it is difficult to construct a numerical scheme which is simultaneously energy and volume preserving, so a choice has to be made. Indeed, already for problems with only one degree of freedom ($d = 1$), energy preserving

means that the numerical method has an exact trajectory $\{y \in \mathbb{R}^{2d}; H(y) = H(y_0)\}$ and thus is a time transformation of the exact solution. Such a scheme can also be volume preserving if and only if this time transformation has divergence zero, which is not satisfied by known energy preserving methods such as the AVF method [27] or energy preserving variants of collocation methods [12].

Order of accuracy in the linear case. The following corollary of Theorem 3.1 states that for (nonpartitioned) Runge–Kutta methods and quadratic potentials, the order of accuracy for the invariant measure of the Lie–Trotter splitting (1.4) is always odd, i.e., if the Runge–Kutta method has order r, the order of accuracy for the invariant measure of the Lie–Trotter splitting is $r+1$ if r is even, and r if r is odd. Moreover, if the Runge–Kutta method is time-symmetric, i.e., $\Phi_h^{-1} = \Phi_{-h}$ then the Lie–Trotter splitting (1.4) samples exactly the invariant measure. We recall [13] that the method Φ_h is symmetric if and only if (2.2) has an expansion in even powers of h, i.e.,

$$f_{2j+1} = 0 \quad \text{for all } j. \tag{3.9}$$

COROLLARY 3.3. *In the linear case (i.e., $V(q)$ is quadratic), assume that Φ_h is a consistent (nonpartitioned) Runge–Kutta method for* (1.5)*. Then the order of accuracy r for sampling the invariant measure of the Lie–Trotter splitting* (1.4) *is odd. Moreover, if Φ_h is symmetric, then the error* (1.7) *of* (1.4) *is zero for all h small enough.*

Proof. Since the problem is linear, the Hamiltonian vector field in (1.5) has the form $f(y) = Ay = J^{-1}Sy$, where S is a constant symmetric matrix and

$$J = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} \tag{3.10}$$

is the $2d \times 2d$ symplectic matrix. For a Runge–Kutta method applied to a linear problem, then we have the identity $\Phi(x) = R(hA)x$, where $R(z)$ is a rational function. The modified differential equation for backward error analysis of Φ_h then has the form (see [13, sect. IX])

$$\frac{d\widetilde{y}}{dt} = h^{-1}\log(R(hA))\widetilde{y} = (A + \alpha_1 hA^2 + \alpha_2 h^2 A^3 + \alpha_3 h^3 A^4 + \cdots)\widetilde{y}, \tag{3.11}$$

with constants $\alpha_1, \alpha_2, \alpha_3, \ldots$. Using (3.9), we have $\alpha_j = 0$ for all odd j in the case of a symmetric method. Since $f_j(y) = \alpha_j A^{j+1}$, to conclude the proof using Corollary 3.2, it is sufficient to show that $\operatorname{div}(f_j) = f_j \cdot \nabla H = 0$ for all even integer j. Indeed, for Φ_h of order r, we already have $f_j = 0$ for all $1 \le j < r$, while for symmetric Φ_h, we already have $f_j = 0$ for all odd j. Since j is even, we have that JA^{j+1} is a symmetric matrix and thus f_j is a Hamiltonian vector field ($f_j = J^{-1}\nabla H_j$ with $H_j(y) = \frac{1}{2}y^T JA^{j+1}y$), which yields $\operatorname{div}(f_j) = 0$. In addition, using the skew-symmetry of SA^{j+1}, we obtain $f_j(y) \cdot \nabla H(y) = \alpha_j y^T SA^{j+1}y = 0$. This concludes the proof. ∎

3.2. Energy and volume conservation are sufficient but not necessary conditions. Corollary 3.2 shows that energy and volume conservation up to order r of the deterministic integrator Φ_h in (1.4) are sufficient conditions for the splitting method (1.4) to have order r for the invariant measure. These are, however, not necessary, as shown in the following theorem, which proves that applying an appropriate time transformation to a symplectic integrator of order r permits us to increase to order $r+1$ the accuracy for the invariant measure, while the order of accuracy for the conservation of both the energy (3.7) and volume (3.8) remains r. We emphasize

that this time transformed numerical method is used only to illustrate a theoretical finding and we do not advocate its use in practice. Here, we use the fact that for a symplectic (possibly partitioned) Runge–Kutta method Φ_h applied to a (deterministic) Hamiltonian system $\dot y = f(y) = J^{-1}\nabla H(y)$, the modified equation for backward error analysis (2.2) also possesses an Hamiltonian structure of the form [13, sect. IX.3]

$$\frac{d\widetilde y}{dt} = J^{-1}\nabla H_h(\widetilde y), \qquad H_h = H + h^r H_r + h^{r+1}H_{r+1} + \cdots, \tag{3.12}$$

where J in defined in (3.10), and the formal series H_h is the modified Hamiltonian.

THEOREM 3.4. *Assume that Φ_h is a consistent symplectic integrator of order r for (1.5) with modified Hamiltonian (3.12). Consider the modified integrator $\overline{\Phi}_h$ defined via a time transformation by*

$$\overline{\Phi}_h(y) := \Phi_{\alpha_h(y)h}(y), \qquad \alpha_h(y) := 1 - h^r \beta H_r(y).$$

Then, assuming ergodicity, the splitting method

$$X_{n+1} = \overline{\Phi}_h \circ \Theta_{h,n}(X_n) \tag{3.13}$$

has order $r+1$ for the invariant measure of (1.1).

Proof. We denote $f_h = f + h^r f_r$ ($f_r = J^{-1}\nabla H_r$) and $\overline{f}_h = f + h^r \overline{f}_r$ the vector fields (2.3) of backward error analysis of Φ_h and $\overline{\Phi}_h$, repectively, truncated at order r. Using $\alpha_h = 1 + \mathcal{O}(h^r)$, we have $\overline{f}_h = \alpha_h f_h + \mathcal{O}(h^{r+1})$. We deduce

$$\begin{aligned}
\operatorname{div}(\overline{f}_h) - \beta \overline{f}_h \cdot \nabla H &= \alpha_h \operatorname{div}(f_h) + \nabla \alpha_h \cdot f_h - \beta \alpha_h f_h \cdot \nabla H + \mathcal{O}(h^{r+1}) \\
&= \alpha_h \operatorname{div}(f_h) - \beta \alpha_h (\nabla H + h^r \nabla H_r) \cdot f_h + \mathcal{O}(h^{r+1}) \\
&= \mathcal{O}(h^{r+1}),
\end{aligned}$$

where we use $\operatorname{div}(f_h) = 0$ and an easy computation shows that $(\nabla H + h^r \nabla H_r)\cdot f_h = 0$. This yields $\operatorname{div}(\overline{f}_r) - \beta \overline{f}_r \cdot \nabla H = 0$. This shows that $\overline{\Phi}_h$ fulfills the assumptions of Corollary 3.2 with r replaced by $r+1$ in (3.5), which concludes the proof. ∎

It is known for deterministic geometric integrators [13, sect. VIII.2] that a nonconstant time transformation destroys in general the geometric properties of symplecticity and volume conservation. Thus, it is not surprising that the scheme (3.13) is energy and volume preserving up to order r but not up to order $r+1$. Indeed, we have that $\operatorname{div}(\overline{f}_h) = \operatorname{div}(\alpha_h f_h) + \mathcal{O}(h^{p+1}) = \nabla \alpha_h \cdot f_h + \mathcal{O}(h^{r+1}) = -h^r \beta \alpha_h \nabla H_r \cdot f_h + \mathcal{O}(h^{r+1})$ is $\mathcal{O}(h^r)$ in general because $\nabla H_r \cdot f_h \neq 0$ (otherwise, Φ_h would be energy preserving up to order $r+1$). This implies that the idea of Theorem 3.4 cannot be applied repeatedly to achieve arbitrarily high order.

Remark 3.4. The time transformation $\alpha_h(y)$ in Theorem 3.4 is nonglobally Lipschitz in general (see, e.g., the formula (4.4) for the symplectic Euler method). This may lead to a nonglobally Lipschitz numerical flow. One possibility to make $\alpha_h(y)$ Lipschitz and bounded is to consider a modification, e.g., $\alpha_h = 1 - (1 + h|H_r|^2)^{-1}\beta h^r H_r)$, which does not affect $f, \bar f_r$ in the modified equation of $\overline{\Phi}_h$.

4. Numerical experiments. In this section we illustrate the various theoretical results of this paper. We start in section 4.1 with an example with a quadratic potential that illustrates the results of Corollary 3.3. We then illustrate in section 4.2 the finding of Theorem 3.1 with an example with quartic potential with various deterministic integrators Φ_h. This section also corroborates that symplecticity is not a necessary geometric property to capture the invariant measure of Langevin dynamics with high order.

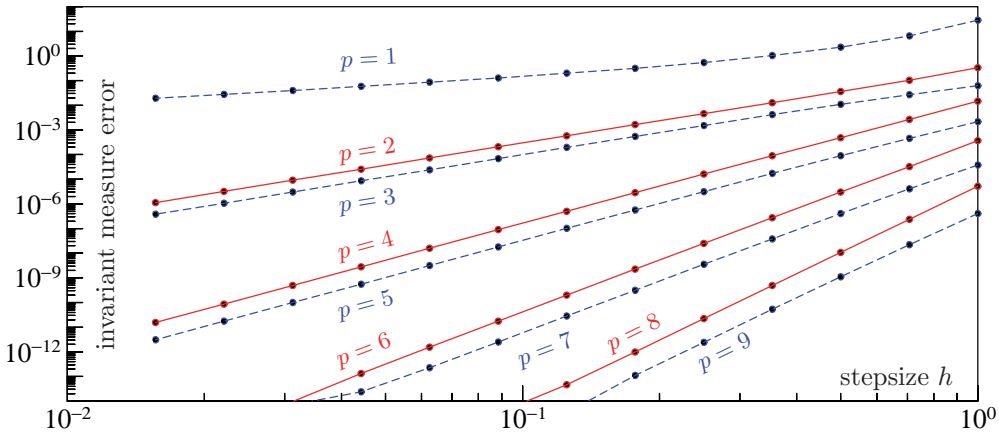

FIG. 1. *Accuracy of the numerical invariant measure (covariance matrix error) of the Lie–Trotter splitting for Langevin dynamics* (1.1) *in the linear case* ($V(q) = \frac{1}{2}q^2$). *The lines correspond to explicit deterministic integrators of orders* $p = 1, 2, 3, \dots, 9$ *(from top to bottom), respectively. The orders of accuracy for the invariant measure are always odd.*

4.1. Illustration for quadratic potential. In Figure 1, we consider the Langevin dynamics (1.1) in $\mathbb{R}^2$ with one degree of freedom ($d = 1$) with quadratic potential $V(q) = \frac{1}{2}q^2$, which reduces (1.5) to the harmonic oscillator. We observe that (1.4) has the form $X_{n+1} = U_h X_n + V_h \xi_n$, where U_h, V_h are constant matrices. This yields that X_n is a Gaussian variable for all n, and a calculation shows that the covariance matrix Σ_h of the corresponding numerical invariant measure is given by the Lyapunov equation

$$U_h \Sigma_h U_h^T - \Sigma_h = V_h V_h^T.$$

We compare in Figure 1 the covariance matrix Σ_h with the covariance matrix $\Sigma = \beta^{-1} I$ of the exact Gaussian invariant measure and plot the 2-norm of the difference. As predicted by Corollary 3.3, we observe the expected lines of with odd slopes $1, 3, 3, 5, 5, 7, 7, 9, 9$ for the explicit deterministic integrators $\Phi_h(y) = (I + A + \frac{A^2}{2} + \cdots + \frac{A^p}{p!})y$ (with $A = J^{-1}$) of order p, where $p = 1, 2, 3, \dots, 9$, respectively.

4.2. Illustration for quartic potential. Although our analysis applies only to globally Lipschitz vector fields, in Figure 2 we consider the Langevin dynamics (1.1) with one degree of freedom with quartic potential $V(q) = (1 - q^2)^2 - \frac{1}{2}q$. We compare the invariant measure error (1.7) for $\phi(p, q) = p^2 + q^2$ of the Lie–Trotter splitting method (1.4) versus the stepsize h for various choices of the deterministic integrator Φ_h:

- the explicit Euler method,

$$p_{n+1} = p_n - h\nabla V(q_n), \qquad q_{n+1} = q_n + hp_n; \tag{4.1}$$

- the Heun method,

$$p_{n+1} = p_n - h\nabla V\left(q_n + \frac{h}{2}p_n\right), \qquad q_{n+1} = q_n + h\left(p_n - \frac{h}{2}\nabla V(q_n)\right); \tag{4.2}$$

- the standard symplectic Euler method,

$$p_{n+1} = p_n - h\nabla V(q_n), \qquad q_{n+1} = q_n + hp_{n+1}; \tag{4.3}$$

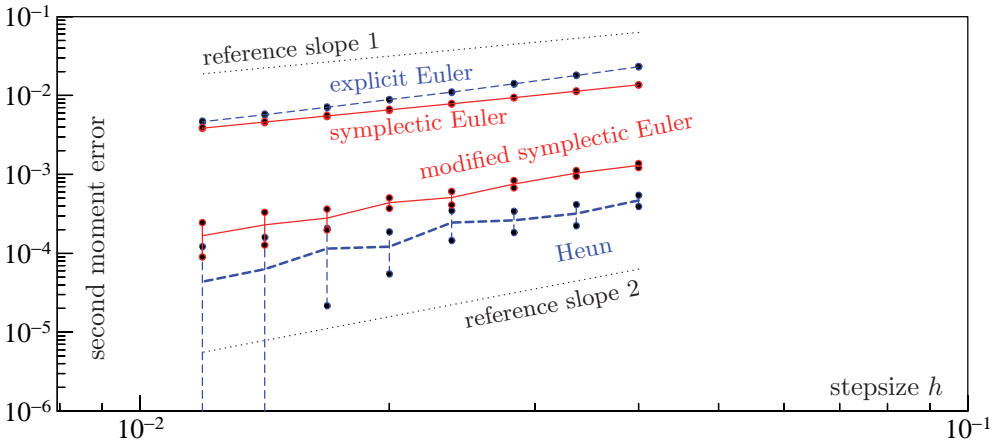

FIG. 2. *Invariant measure error* (1.7) *for* $\phi(p,q) = p^2+q^2$ *of the Lie–Trotter splitting method,* (1.4) *versus the stepsize* h *for Langevin dynamics* (1.1) *with quartic potential* $V(q) = (1-q^2)^2 - \frac{1}{2}q$ $(d = 1)$*, with* $\gamma = 4, \beta = 2$ *and initial condition* $p(0) = q(0) = -1.5$*. We compare various choices of the deterministic integrator* Φ_h*: explicit Euler* (4.1)*, symplectic Euler* (4.4)*, Heun explicit RK2* (4.2)*, and the modified symplectic Euler method* (4.4)*. The curves are obtained as the averages over* 10 *independent long trajectories of time length* $T = 10^8$*. The vertical bars measure the standard deviation intervals due to the Monte Carlo error among these* 10 *trajectories.*

- a time transformed version of the symplectic Euler method, (4.4)

$$p_{n+1} = p_n - \alpha_n h \nabla V(q_n), \quad q_{n+1} = q_n + \alpha_n h p_{n+1}, \quad \alpha_n = 1 + \frac{h}{2}\beta p_n \cdot \nabla V(q_n).$$

The explicit Euler method, the standard symplectic Euler method, and its time transformed version have deterministic order 1 for (1.5), while the Heun method has deterministic order 2. In Figure 2, we plot the error (1.7) for $\phi(p,q) = p^2 + q^2$ for the methods considered above. We consider for each method the average over 10 independent long time trajectories with length $T = 10^8$. The vertical bars represent the standard deviation intervals due to the Monte Carlo error among these 10 trajectories. As predicted by Corollary 3.2, the expected convergence slopes 1 and 2 can be observed in Figure 2. Precisely, the method (4.4) has deterministic order 1 and preserves the energy and the volume only up to order $r = 1$ in (3.7), (3.8), but it has order 2 of accuracy for the invariant measure, as predicted by Theorem 3.4. We observe in Figure 2 that the order 2 schemes have an accuracy improved by a factor of about 100 compared to the order 1 methods. Again, as mentioned in section 3.2, the time transformed symplectic Euler method (1.1) is presented only to illustrate Theorem 3.4 and we do not claim that it is the best method at hand to solve problem (1.1).

Geometric properties of symplectic and time transformed symplectic methods. In this experiment we take again the quartic potential $V(q) = (1-q^2)^2 - \frac{1}{2}q$, temperature $\beta = 2$, and initial condition $p(0) = q(0) = -1.5$. In Figure 3, we plot the evolution of the Hamiltonian energy (1.2) along time for the standard symplectic Euler method (4.3) and the time transformation version (4.4) as described in Theorem 3.4 for two values of γ.

In the top picture of Figure 3, we consider the deterministic case ($\gamma = 0$). In contrast to the standard symplectic Euler method for which no drift in the energy

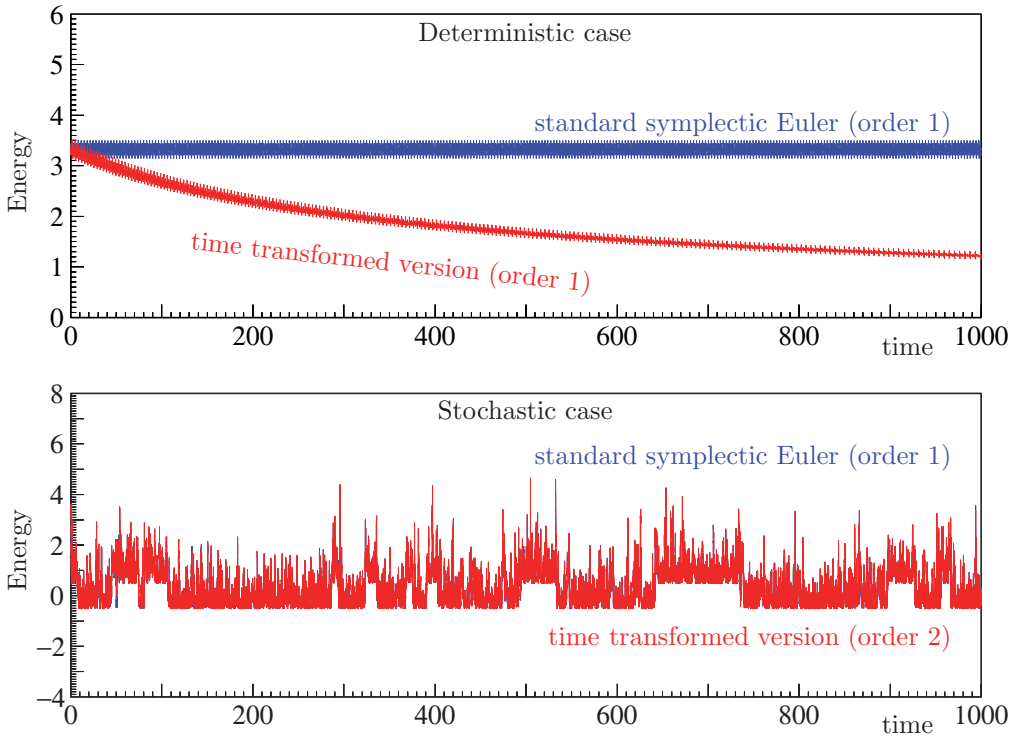

FIG. 3. *Evolution of the Hamiltonian energy* (1.2) *along time for the standard symplectic Euler method* (4.3) *and the time transformed version* (4.4) *(nonsymplectic) applied to* (1.1) *with* $V(q) = (1-q^2)^2 - \frac{1}{2}q$ *(quartic potential). Top picture: deterministic case* ($\gamma = 0$). *Bottom picture: stochastic case* ($\gamma = 4$). *Stepsize is* $h = 0.02$ *and* $\beta = 2$.

can be observed, we obtain a drift in the energy for the time transformed version (4.4) which is nonsymplectic. In the bottom picture, we consider the stochastic case ($\gamma = 4$) and plot sample trajectories for both methods using the same sets of generated random variables. We observe no linear drift in the energy for both methods in this stochastic case. It can be observed in Figure 2 that the modified symplectic Euler method yields order 2 of accuracy compared to the standard version of order 1, as predicted by Theorem 3.4. This last numerical example shows that among the standard symplectic Euler method (4.3) and the time transformed version (4.4) (applied in the Lie–Trotter splitting (1.4)), the most accurate scheme is not the same in the deterministic and stochastic contexts, as predicted by Theorem 3.4.

5. Conclusion. We have presented an analysis supported by numerical experiments of Lie–Trotter splitting methods applied to nonlinear Langevin dynamics. In particular, we obtained a new characterization of sufficient conditions for the Lie–Trotter splitting to capture the numerical invariant measure of nonlinear ergodic Langevin dynamics up to arbitrary order. While in previous results in the literature symplecticity and energy conservation (up to a given order) were required to prove high order accuracy for the numerical invariant measure, we showed that although sufficient, these conditions are not necessary. More precisely, we showed that the order of convergence of the deterministic integrator provides a lower bound for the order of convergence for the numerical invariant measure of the Lie–Trotter splitting method. We illustrated this finding by constructing a first order nonsymplectic deterministic

integrator that, despite its poor geometric properties, captures the invariant measure of Langevin dynamics to second order.

REFERENCES

[1] A. Abdulle and A. Blumenthal, *Improved stabilized Multilevel Monte Carlo method for stiff stochastic differential equations*, in Numerical Mathematics and Advanced Applications ENUMATH 2013, Lect. Notes Comput. Sci. Eng. 103, Springer, Heidelberg, 2015, pp. 537–545.

[2] 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.

[3] A. Abdulle, G. Vilmart, and K. C. Zygalakis, *High order numerical approximation of the invariant measure of ergodic SDEs*, SIAM J. Numer. Anal., 52 (2014), pp. 1600–1622.

[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] N. Bou-Rabee and E. Vanden-Eijnden, *Pathwise accuracy and ergodicity of metropolized integrators for SDEs*, Comm. Pure Appl. Math., 63 (2010), pp. 655–696.

[6] K. Burrage, I. Lenane, and G. Lythe, *Numerical methods for second-order stochastic differential equations*, SIAM J. Sci. Comput., 29 (2007), pp. 245–264.

[7] 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.

[8] P. Chartier, E. Hairer, and G. Vilmart, *Numerical integrators based on modified differential equations*, Math. Comp., 76 (2007), pp. 1941–1953.

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

[10] M. B. Giles, *Multilevel Monte Carlo path simulation*, Oper. Res., 56 (2008), pp. 607–617.

[11] M. B. Giles and L. Szpruch, *Antithetic multilevel Monte Carlo estimation for multi-dimensional SDEs without Lévy area simulation*, Ann. Appl. Probab., 24 (2014), pp. 1585–1620.

[12] E. Hairer, *Energy-preserving variant of collocation methods*, JNAIAM J. Numer. Anal. Ind. Appl. Math., 5 (2010), pp. 73–84.

[13] 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.

[14] M. Kopec, *Weak backward error analysis for Langevin process*, preprint, 2013.

[15] M. Kopec, *Weak backward error analysis for overdamped Langevin processes*, IMA J. Numer. Anal. (2014).

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

[17] B. Leimkuhler, C. Matthews, and G. Stoltz, *The computation of averages from equilibrium and nonequilibrium Langevin molecular dynamics*, IMA J. Numer. Anal. (2014).

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

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

[20] 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 (2002), pp. 185–232.

[21] 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.

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

[23] G. N. Milstein and M. V. Tretyakov, *Stochastic numerics for mathematical physics*, in Scientific Computing, Springer-Verlag, Berlin, 2004.

[24] 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.

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

[26] B. Oksendal, *Stochastic Differential Equations*, 6th ed., Springer-Verlag, Berlin, 2005.

[27] G. R. W Quispel and D. I. McLaren, *A new class of energy-preserving numerical integration methods*, J. Phys. A, 41 (2008), 045206.
[28] G. O. Roberts and R. L. Tweedie, *Exponential convergence of Langevin distributions and their discrete approximations*, Bernoulli, 2 (1996), pp. 341–363.
[29] J. M. Sanz-Serna, *Markov Chain Monte Carlo and Numerical Differential Equation*, Lectures Notes in Math. 2082, Springer, Berlin, 2014.
[30] J. M. Sanz-Serna and M. P. Calvo, *Numerical Hamiltonian Problems*, Appl. Math. Math. Comput. 7, Chapman & Hall, London, 1994.
[31] 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.
[32] D. Talay, *Discrétisation d'une équation différentielle stochastique et calcul approché d'espérances de fonctionnelles de la solution*, RAIRO Modél. Math. Anal. Numér., 20 (1986), pp. 141–179.
[33] 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.
[34] D. Talay and L. Tubaro, *Expansion of the global error for numerical schemes solving stochastic differential equations*, Stochastic Anal. Appl., 8 (1990), pp. 483–509.
[35] 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.