Long time accuracy of Lie-Trotter splitting methods for Langevin dynamics

Assyr Abdulle, Gilles Vilmart, Konstantinos C. Zygalakis
Long time accuracy of Lie-Trotter splitting methods for Langevin dynamics

Assyr Abdulle\textsuperscript{1}, Gilles Vilmart\textsuperscript{2}, and Konstantinos C. Zygalakis\textsuperscript{3}

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, modified differential equations, backward error analysis, invariant measure.

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

1 Introduction

In many applications (see [20, 13, 4, 16] and references therein), one is interested in simulating the invariant measure of a stochastic differential equation (SDE) by running a numerical scheme that approximates the time dynamics of the SDE. In particular, one uses the numerical trajectories to construct an empirical measure either by averaging one single long trajectory or by averaging over many realisations to obtain a finite ensemble average (see for example [20]). One immediately faces 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 [17, 27, 23, 25], 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

\begin{align}
 dq(t) &= M^{-1}p(t)dt, \\
 dp(t) &= (-\nabla V(q(t)) - \gamma p(t))dt + \sqrt{2\beta^{-1}\gamma}M^{1/2}dW(t),
\end{align}

where the scalar function $V : \mathbb{R}^d \rightarrow \mathbb{R}$ 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 positive definite matrices \(M\). Under appropriate smoothness and growth assumptions on the Hamiltonian energy (see e.g. [18, 11]),

\[
H(p, q) = \frac{1}{2} p^T p + V(q),
\]

one can show that the dynamics generated by (1) are ergodic, i.e. for arbitrary initial conditions \(q(0), p(0)\) (assumed deterministic for simplicity), and for all smooth test function \(\phi \in C_0^\infty(\mathbb{R}^{2d}, \mathbb{R})\) with derivatives of all orders with polynomial growth, the time average of the trajectories of (1) satisfy with probability 1,

\[
\lim_{T \to \infty} \frac{1}{T} \int_0^T \phi(p(t), q(t)) = \int_{\mathbb{R}^{2d}} \phi(x) d\mu(x).
\]

The invariant measure \(\mu\) is characterized by the Gibbs density function

\[
\rho_\infty(p, q) = Z e^{-\beta H(p, q)},
\]

where \(Z\) is the renormalization constant, which ensures that \(\int_{\mathbb{R}^d \times \mathbb{R}^d} \rho_\infty(p, q) dp dq = 1\). Similarly to [3], we consider the following class of schemes based on the Lie-Trotter splitting,

\[
X_{n+1} = \Phi_h \circ \Theta_{h,n}(X_n),
\]

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

\[
dq(t) = p(t) dt, \quad dp(t) = -\nabla V(q(t)) dt,
\]

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

\[
dq(t) = 0, \quad dp(t) = -\gamma p(t) dt + \sqrt{2\beta^{-1}\gamma} dW(t).
\]

The exact solution of (6) is given by the variation of constants formula,

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

The trajectories of the above stochastic integral itself need not to be simulated. Indeed, the properties of the stochastic integral [21] imply that the solution \(p(t)\) of (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 (6) which is exact in law,\(^1\)

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

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

\(^1\)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 \((p(0), q(0))\).
The aim of this paper is to analyze splitting schemes such as (4) for a variety of different deterministic integrators Φₕ used for solving (5), for the numerical approximation of the invariant measure, i.e., we quantify the error
\[ e(φ, h) := \lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} φ(p_n, q_n) - \int_{\mathbb{R}^d \times \mathbb{R}^d} φ(p, q) \rho_∞(p, q) dp dq, \] (7) for all smooth test function φ ∈ C_∞(\mathbb{R}^{2d}, \mathbb{R}). We say that the method has order r for sampling the invariant measure if |e(φ, h)| ≤ 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 [1, 7] and backward error analysis for SDEs [29, 8, 11, 12, 2], 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.

Note that the problem of quantifying e(φ, h) has been considered before in a number of papers [3, 13, 14]. In [3], a splitting method called the Geometric Langevin Algorithm is proposed and it is shown that if Φₕ is a variational integrator (equivalently a symplectic scheme) which preserves the energy up to order r, then the Lie-Trotter splitting (4) has order r of accuracy for the invariant measure. In this paper, we provide weaker assumptions to achieve order r in the approximation of the numerical invariant measure. In [13, 14] the ergodic measure accuracy of several splitting methods for (1) is analyzed. In particular, schemes that conserve the invariant measure up to order two (based on Strang splitting) are discussed. These schemes have weak second order of accuracy (symmetric Strang splitting methods), i.e. the weak error \( \mathbb{E}(φ(X_n)) - \mathbb{E}(φ(X(t_n))) \) has size \( \mathcal{O}(h^2) \) for all \( t_n = nh \leq T \). Our results show that such a weak second order is not necessary to achieve high order for the invariant measure, because the Lie-Trotter splitting (4) which has weak order of convergence only one, is sufficient to achieve high order for the invariant measure. This is also related to the results in [5, 6] 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 the class of stochastic Runge-Kutta methods applied to Langevin dynamics with a linear drift (which does not include the Lie-Trotter splitting (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 [16]. Our conclusion for Lie-Trotter splitting methods (4) is different: we prove in particular that it samples exactly the invariant measure for a linear drift if the deterministic integrator Φₕ 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 terms of functions arising in the modified differential equation. This is also a new expression for error estimates à la Talay-Tubaro [28].

To simplify the presentation of our results, we focus on the globally Lipschitz case and we assume that the potential \( V \) is a \( C^∞ \) function where \( \nabla V \) has bounded derivatives of any order (which implies the global Lipschitzness of \( \nabla V \)) and satisfies the following standard growth condition, which guarantees that (1) is ergodic,
\[ q^T \nabla V(q) \geq C_1 q^T q - C_2, \quad \text{for all } q \in \mathbb{R}^d, \] (8)
for some constants $C_1, C_2 > 0$. We further assume that the numerical flow $\Phi_h$ is globally Lipschitz in $\mathbb{R}^{2d}$,
\[\|\Phi_h(x_1) - \Phi_h(x_2)\| \leq L\|x_1 - x_2\|,\] (9)
for all $x_1, x_2 \in \mathbb{R}^{2d}$ and all stepsize $h$ small enough. In a more general non-globally Lipschitz setting, important in most applications, two implicit schemes are studied in [11], and the implicitness is used to guaranty the boundedness of the numerical moments along time. Using implicit schemes for the Hamiltonian part is one way to extend our results in the non-globally 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 non-globally Lipschitz setting by following the methodology of rejecting exploding trajectories proposed in [19] and applied to Langevin-type ergodic systems in [20].

The paper is organized as follows. In section 2, we present the basic ingredients of backward error analysis and the framework in [2] that allows us to characterise the long time behaviour 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, both for 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 [2] 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 [24, 15, 10]. 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 $\Phi_h$ of (5). Given a consistent integrator $y_{n+1} = \Phi_h(y_n)$ for a system of ODEs
\[
\frac{dy(t)}{dt} = f(y(t)),
\] (10)
the idea of backward error analysis is to search for a modified differential equation written as a formal series in powers of the stepsize $h$,
\[
\frac{d\tilde{y}}{dt} = f(\tilde{y}) + hf_1(\tilde{y}) + h^2f_2(\tilde{y}) + \ldots, \quad \tilde{y}(0) = y_0
\] (11)
such that (formally) $y_n = \tilde{y}(t_n)$, 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 $j$ [15, 10], 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 \( s \)

\[
\frac{d\tilde{y}}{dt} = f(\tilde{y}) + h f_1(\tilde{y}) + h^2 f_2(\tilde{y}) + \ldots + h^s f_s(y), \quad \tilde{y}(0) = y_0
\]  

(12)
we have \( y_n = \tilde{y}(t_n) + \mathcal{O}(h^{s+1}) \) for \( h \to 0 \) for bounded times \( t_n = nh \leq T \). We note that the flow \( \tilde{\Phi}_h(y) \) of the modified differential equation (12) can be expressed as

\[
\tilde{\Phi}_h = \left( \sum_{k=0}^{M} \frac{h^k \tilde{L}_D^{(k)}}{k!} \right) I + \mathcal{O}(h^{M+1}), \quad \tilde{L}_D = F_0 + h F_1 + h^2 F_2 + \ldots + h^s F_s,
\]  

(13)
for all \( M \geq 0 \), 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 [10, Sect. III.5.1]

\[
\phi \circ \tilde{\Phi}_h = \left( \sum_{k=0}^{M} \frac{h^k \tilde{L}_D^{(k)}}{k!} \right) \phi + \mathcal{O}(h^{M+1}),
\]  

(14)
for all smooth test function \( \phi \). The constants symbolized by \( \mathcal{O} \) in (13), (14), are independent of \( h \to 0 \) but depend on \( M, s \) and \( \phi \).2

2For all \( h \) small enough, the sum in (13) and (14) can be shown to converge for \( M \to \infty \) in the case of analytic vector fields \( f_j \) (and analytic test functions \( \phi \)), which permits to remove the \( \mathcal{O} \) remainder.

2.2 Order conditions for the numerical invariant measure

A numerical method applied to (1) is called ergodic if it has a unique invariant law \( \mu^h \) 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) should satisfy in order to approximate the invariant measure \( \mu \) with an error \( (7) \) 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 following expansion

\[
U(x, h) = \phi(x) + h A_0 \phi(x) + h^2 A_1 \phi(x) + h^3 A_2 \phi(x) + \ldots
\]  

(15)
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) given by

\[
\mathcal{L} := p \cdot \nabla_q - (\nabla V(q) + \gamma p) \cdot \nabla p + \beta^{-1} \gamma \Delta p.
\]  

(16)
Remark 2.1. Assuming that $\nabla V$ is smooth with bounded derivatives at all orders, (19) has a unique solution for all test function $\phi \in C_0^\infty(\mathbb{R}^{2d}, \mathbb{R})$. It admits a rigorous Taylor expansion up to arbitrary order $p$ for all $h$ small enough

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

where the function symbolized in $O$ is smooth with respect to $x$ with its all derivatives having a polynomial growth, independently of the smallness of the stepsize $h$. Notice that applying this result to the deterministic problem (12), we recover the statement (14).

Example 2.2. Consider the Euler-Maruyama method for (1) defined as

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

where $\xi_n \sim \mathcal{N}(0, 1)$ are independent $d$-dimensional Gaussian random variables. A straightforward calculation yields that $A_1$ is a differential operator of order 4,

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

where $\phi''(\cdot, \cdot)$ denotes the second differential of $\phi$ (bilinear form), and $\phi'$ is the first derivative (linear form). We refer to [1, 29] for other examples of such weak Taylor series calculations.

We can now state the conditions which ensure that $\mu^h$ approximates $\mu$ to order $r$. We have the following theorem which involves the adjoint operators $A_j^*$ of the differential operators $A_j$ involved in (15), 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 $\phi_1, \phi_2$ with compact support on $\mathbb{R}^{2d}$. The following result is proved in [2].

Theorem 2.3. For the ergodic system (1), consider a consistent ergodic numerical method satisfying (15) and

$$A_j^*\rho_\infty = 0, \text{ for } j = 1, \ldots, r - 1. \quad (17)$$

Then one obtains that the error in (7) satisfies for all test functions $\phi \in C_0^\infty(\mathbb{R}^{2d}, \mathbb{R})$,

$$e(\phi, h) = h^r \int_0^\infty \int_{\mathbb{R}^{2d} \times \mathbb{R}^{2d}} A_r u(p, q, t)\rho_\infty(p, q)dpdqdt + O(h^{r+1}) \quad (18)$$

where $\rho_\infty$ is the invariant measure density given in (3) and $u(x, t)$ solves the backward Kolmogorov equation

$$\frac{\partial u}{\partial t}(p, q, t) = L u(p, q, t), \quad t > 0, p, q \in \mathbb{R}^d, \quad (19a)$$
$$u(p, q, 0) = \phi(p, q), \quad (19b)$$

and $L$ is the generator of (1) defined in (16).
A geometric interpretation of the condition (17) is the following: the invariant measure \( \mu \) with density \( \rho_{\infty} \) of (1) is preserved up to a \( O(h^r) \) error by a single time step \( X_0 \mapsto X_1 \) of the integrator. Precisely, if \( X_0 \) is a random variable with law of probability \( \mu \), then for all test functions \( \phi \in C^\infty_\mathbb{P}(\mathbb{R}^{2d}, \mathbb{R}) \) and \( h \to 0 \),

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

where the constant symbolized by \( O \) is independent of \( h \) assumed small enough.

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 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 form backward error analysis, we prove the following result for the accuracy of the sampling the invariant measure (3) of Langevin dynamics of the form (1).

**Theorem 3.1.** Consider the ergodic system (1). Assume that \( \Phi_h \) is a consistent method for (5) satisfying (9) with modified equation from backward error analysis (11). If the vector fields in (11) satisfy

\[
\text{div}(f_j \rho_{\infty}) = 0, \quad j = 1, \ldots, r - 1,
\]

then, assuming ergodicity, the Lie-Trotter splitting (4) has order \( r \) of accuracy for the invariant measure. Precisely, we have for all \( \phi \in C^\infty_\mathbb{P}(\mathbb{R}^{2d}, \mathbb{R}) \) and \( h \to 0 \),

\[
e(\phi, h) = -h^r \int_{0}^{\infty} \int_{\mathbb{R}^{2d} \times \mathbb{R}^{2d}} u(p, q, t) \text{div}(f_r(p, q)\rho_{\infty}(p, q)) dpdqdt + O(h^{r+1}),
\]

where \( e(\phi, h) \) is defined in (7) and \( u(x, t) \) is defined in (19). The constant symbolized in \( O \) is independent of the stepsize \( h \) assumed small enough.

**Proof.** Consider the method (4) and its corresponding one step expansion in powers of \( h \)

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

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

\[
A_j^* \rho_{\infty} = 0 \quad \text{for} \quad j = 1, \ldots, r - 1, \quad A_r^* \rho_{\infty} = \text{div}(f_r \rho_{\infty}).
\]

The result then follows immediately from Theorem 2.3 using the identity

\[
\int_{\mathbb{R}^{2d}} A_r u(x, t)\rho_{\infty}(x)dx = -\int_{\mathbb{R}^{2d}} u(x, t) \text{div}(f_r(x)\rho_{\infty}(x))dx.
\]
We now start with the calculation of \( A_j \). In particular, given a smooth test function \( \phi \) and \( x \in \mathbb{R}^{2d} \), using the semi-group property of the Markov process we have
\[
\mathbb{E}(\phi(X_1)|X_0 = x) = \mathbb{E}(\phi(\Phi_h \circ \Theta_{h,n})(X_0)|X_0 = x) = e^{h\mathcal{L}_S(\phi \circ \Phi_h)(x)},
\]
where \( e^{h\mathcal{L}_S}\phi \) denotes the exact flow of the Kolmogorov backward equation corresponding to (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 (13) of the truncated modified equation (12) of the integrator \( \Phi_h \)
\[
\tilde{\mathcal{L}}_D = F_0 + h F_1 + \ldots + h^r F_r,
\]
where we define the differential operators \( F_j \phi = f_j \cdot \nabla \phi \) (with \( f_0 = f \)). We have, using (23), applying Remark 2.1 for \( \mathcal{L}_S \), and using (14) with \( M = s = r \),
\[
\mathbb{E}(\phi(X_1)|X_0 = x) = \left( \sum_{k=0}^{r} \frac{h^k \mathcal{L}_{S_k}^*}{k!} \right) \left( \sum_{k=0}^{r} \frac{h^k \tilde{\mathcal{L}}_{D_k}^*}{k!} \right) \phi(x) + O(h^{r+1})
\]
\[
= \phi(x) + h \mathcal{L} \phi(x) + \sum_{k=1}^{r} h^{k+1} A_k \phi(x) + O(h^{r+1})
\]
where we denote
\[
A_k = \sum_{j=0}^{k+1} \mathcal{L}_{S_j}^{k+1-j} \left( \sum_{\substack{1 \leq i_1 \leq j \leq \ldots \leq i_n \leq j \leq 1 \leq i_1 \leq j \leq \ldots \leq i_n \leq j - \sum_{j=1}^{n}}} \frac{1}{i!(k+1-j)!} F_{n_1} \cdots F_{n_i} \right),
\]
where the second sum above is over integers \( n_1, \ldots, n_i \geq 0 \) and is equal to the identity \( I \) when \( j = 0 \). We obtain for all \( k \geq 1 \),
\[
A_k \rho_{\infty} = \sum_{j=0}^{k+1} \left( \sum_{\substack{1 \leq i_1 \leq j \leq \ldots \leq i_n \leq j \leq 1 \leq i_1 \leq j \leq \ldots \leq i_n \leq j - \sum_{j=1}^{n}}} \frac{1}{i!(k+1-j)!} F_{n_1}^{*} \cdots F_{n_i}^{*} \right) (\mathcal{L}_{S_j}^{k+1-j} \rho_{\infty}).
\]
Using \( \mathcal{L}_{S_j}^{k} \rho_{\infty} = 0 \) and \( F_{i}^{*} \rho_{\infty} = 0, i = 1, \ldots, r-1 \) we see that for \( k \leq r \), the only possibly non-zero term in the above sum is obtained for \( j = r+1, k = r, i = 1, \ldots, r-1 \). Hence, we deduce (21) which permits to conclude the proof. \( \square \)

**Remark 3.2.** 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
\[
\text{div}(f_j \rho_{\infty}) = (\text{div}(f_j) - \beta f_j \cdot \nabla H) \rho_{\infty}.
\]

**Corollary 3.3.** Under the assumptions of Theorem 3.1, if
\[
\text{div}(f_j) - \beta f_j \cdot \nabla H = 0, \quad j = 1, \ldots, r - 1,
\]
then, assuming ergodicity, the Lie-Trotter splitting (4) has order \( r \) of accuracy for the invariant measure.

The simplest way for condition (24) to be satisfied is to consider a method \( \Phi_h \) of order \( r \), since in this case \( f_j = 0, j = 1, \ldots, r-1 \). Indeed, in this case, standard results from backward error analysis show that the right-hand side of the modified equation (12) reads \( f(\bar{y}) + h^r f_r(y) + \ldots \) and the \( f_j = 0, j = 1, \ldots, r-1 \).
Relation with results of [3]. In the paper [3] the authors show that if a variational (equivalently symplectic) integrator $\Phi_h$ preserving the energy with order $r$ is used for solving the deterministic dynamics (5), then the Lie-Trotter splitting (4) has order $r$ of accuracy for the invariant measure. Now a symplectic method applied to a Hamiltonian system is volume preserving hence $\text{div}(f_j) = 0$ for all $j \geq 0$. Furthermore, if the method $\Phi_h$ has order $r$, then it also preserves energy up to order $r$, i.e. for all vectors $y$,

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

(25)

this condition is equivalent to $f_j \cdot \nabla H = 0$ for all $1 \leq j \leq r - 1$. If both term in (24) are required to vanish independently, we see that we obtain order $r$ of accuracy for the invariant measure under weaker assumptions than in [3]. It is not necessary for the method to be symplectic, only volume preservation up to order $r$ is required i.e.,

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

(26)

which is equivalent to $\text{div}(f_j) = 0$ for all $1 \leq j \leq r-1$. Note that any deterministic method $\Phi_h$ of order $r$ automatically satisfies (25) and (26), 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 $\Phi_h$ in (4) are not necessary to get order $r$ of accuracy for the invariant measure, as one can construct deterministic methods that satisfy (24) with $j = r + 1$, while conserving energy and volume only up to order $r$.

**Remark 3.4.** 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 $\{y \in \mathbb{R}^2; H(y) = H(y_0)\}$, and thus it a time transformation of the exact solution. Such scheme can be also volume preserving if and only if this time transformation has divergence zero, which is not satisfied by known energy preserving methods such has the AVF method [22] or energy preserving variants of collocation methods [9].

**Order of accuracy in the linear case.** The following corollary of Theorem 3.1 states that for (non-partitioned) Runge-Kutta methods and quadratic potentials, the order of accuracy for the invariant measure of the Lie-Trotter splitting (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 (4) samples exactly the invariant measure. We recall [10] that the method $\Phi_h$ is symmetric if and only if (11) has an expansion in even powers of $h$, i.e.

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

(27)

**Corollary 3.5.** In the linear case (i.e. $V(q)$ is quadratic), assume that $\Phi_h$ is a consistent (non-partitioned) Kutta method for (5). Then the order of accuracy for the invariant measure of the Lie-Trotter splitting (4) is odd. Moreover, if $\Phi_h$ is symmetric, then (4) is exact for all $h$ small enough.
Proof. Since the problem is linear, the Hamiltonian vector field in (5) has the form \( f(y) = Ay = J^{-1}Sy \) where \( S \) is a constant symmetric matrix and
\[
J = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}
\] (28)
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 \( \Phi_h \) then has the form (see [10, Sect. IX])
\[
\frac{d\tilde{y}}{dt} = \frac{1}{h} \log(R(hA))\tilde{y} = (A + \alpha_1 hA^2 + \alpha_2 h^2 A^3 + \alpha_3 h^3 A^4 + \ldots)\tilde{y},
\] (29)
with constants \( \alpha_1, \alpha_2, \alpha_3, \ldots \). Using (27), 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 3.3, it is sufficient to show that \( \text{div}(f_j) = f_j \cdot \nabla H = 0 \) for all even integer \( j \). Indeed, for \( \Phi_h \) of order \( r \), we already have \( f_j = 0 \) for all \( 1 \leq j < r \), while for symmetric \( \Phi_h \), 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 \( (f_j = J^{-1}\nabla H_j \text{ with } H_j(y) = \frac{1}{2}y^T J A^{j+1} y) \), which yields \( \text{div}(f_j) = 0 \). 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. \( \square \)

3.2 Energy and volume conservation are sufficient, but not necessary conditions

Corollary 3.3 shows that energy and volume conservation up to order \( r \) of the deterministic integrator \( \Phi_h \) in (4) are sufficient conditions for the splitting method (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 to increase to order \( r + 1 \) the accuracy for the invariant measure, while the order of accuracy for the conservation of both the energy (25) and volume (26) remains \( r \). Here, we use the fact that for a symplectic (possibly partitioned) Runge-Kutta method \( \Phi_h \) applied to a (deterministic) Hamiltonian system \( \dot{y} = f(y) = J^{-1}\nabla H(y) \), the modified equation for backward error analysis (11) also possesses an Hamiltonian structure of the form [10, Sect. IX.3]
\[
\frac{d\tilde{y}}{dt} = J^{-1}\nabla H_h(\tilde{y}), \quad H_h = H + h^r H_r + h^{r+1} H_{r+1} + \ldots,
\] (30)
where \( J \) in defined in (28), and the formal series \( H_h \) is the modified Hamiltonian.

Theorem 3.6. Assume that \( \Phi_h \) is a consistent symplectic integrator of order \( r \) for (5) with modified Hamiltonian (30). Consider the modified integrator \( \overline{\Phi}_h \) defined via a time transformation by
\[
\overline{\Phi}_h(y) := \Phi_{\alpha_h(y)}(y), \quad \alpha_h(y) := 1 - h^r \beta H_r(y).
\]
Then, assuming ergodicity, the splitting method
\[
X_{n+1} = \overline{\Phi}_h \circ \Theta_{h,n}(X_n)
\] (31)
has order \( r + 1 \) for the invariant measure of (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 \((12)\) of backward error analysis of \( \Phi_h \) and \( \overline{\Phi}_h \), respectively, truncated at order \( r \). Using \( \alpha_h = 1 + O(h^r) \), we have \( \overline{f}_h = \alpha_h f_h + O(h^{r+1}) \). We deduce

\[
\text{div}(\overline{f}_h) - \beta \overline{f}_h \cdot \nabla H = \alpha_h \text{div}(f_h) + \nabla \alpha_h \cdot f_h - \beta \alpha_h f_h \cdot \nabla H + O(h^{r+1})
\]

\[
= \alpha_h \text{div}(f_h) - \beta \alpha_h (\nabla H + h^r \nabla H_r) \cdot f_h + O(h^{r+1})
\]

\[
= O(h^{r+1})
\]

where we use \( \text{div}(f_h) = 0 \) and an easy computation shows that \( (\nabla H + h^r \nabla H_r) \cdot f_h = 0 \). This yields \( \text{div}(\overline{f}_h) - \beta \overline{f}_h \cdot \nabla H = 0 \). This shows that \( \overline{\Phi}_h \) fulfills the assumptions of Corollary 3.3 with \( r = 2 \), which concludes the proof. \( \square \)

It is known for deterministic geometric integrators [10, Sect. VIII.2] that a non constant time transformation destroys in general the geometric properties of symplecticity and volume conservation. Thus, it is not surprising that the scheme \((31)\) is energy and volume preserving up to order \( r \) but not up to order \( r+1 \). Indeed, we have that \( \text{div}(\overline{f}_h) = \text{div}(\alpha_h f_h) + O(h^{r+1}) = \nabla \alpha_h \cdot f_h + O(h^{r+1}) = -h^r \beta \alpha_h \nabla H_r \cdot f_h + O(h^{r+1}) \) is \( O(h^r) \) in general because \( \nabla H_r \cdot f_h \neq 0 \) (otherwise, \( \Phi_h \) would be energy preserving up to order \( r+1 \)). This implies that the idea of Theorem 3.6 cannot be applied repeatedly to achieve arbitrarily high order.

Remark 3.7. The time transformation \( \alpha_h(y) \) in Theorem 3.6 is non-globally Lipschitz in general (see e.g. the formula \((35)\) for the symplectic Euler method). This may lead to a non-globally 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 \( f, f_r \) 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 by an example with a quadratic potential, that illustrates the results of Corollary 3.5. We then illustrate in Section 4.2 the finding of Theorem 3.1 by an example with quartic potential with various deterministic integrators \( \Phi_h \). This section also corroborates that symplecticity is not a necessary geometric property to capture the invariant measure of Langevin dynamics with high order.

4.1 Illustration for quadratic potential

In Figure 1, we consider the Langevin dynamics \((1)\) in \( \mathbb{R}^2 \) with one degree of freedom \((d = 1)\) with quadratic potential \( V(q) = \frac{1}{2} q^2 \), which reduces \((5)\) to the harmonic oscillator. We observe that \((4)\) 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 \( \Sigma_h \) 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 \( \Sigma_h \) 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 3.5, we observe the expected lines of with odd slopes 1, 3, 5, 7, 9, for the explicit deterministic integrators \( \Phi_h(y) = (I + A + \frac{A^2}{2} + \ldots + \frac{A^p}{p!})y \) (with \( A = J^{-1} \)) of order \( p \), where \( p = 1, 2, 3, \ldots, 9 \), respectively.

**Figure 1:** Accuracy of the numerical invariant measure (covariance matrix error) of the Lie-Trotter splitting for Langevin dynamics (1) in the linear case \( (V(q) = \frac{1}{4}q^2) \). The lines corresponds to explicit deterministic integrators of orders \( p = 1, 2, 3, \ldots, 9 \) (from top to bottom), respectively. The orders of accuracy for the invariant measure are always odd.

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) with one degree of freedom with quartic potential \( V(q) = (1 - q^2)^2 - \frac{1}{2}q \). We compare the invariant measure error (7) for \( \phi(p, q) = p^2 + q^2 \) of the Lie-Trotter splitting method (4) versus the stepsizes \( h \) for various choices of the deterministic integrator \( \Phi_h \):

- the explicit Euler method,
  \[
  p_{n+1} = p_n - h\nabla V(q_n), \quad q_{n+1} = q_n + hp_n, \tag{32}
  \]

- the Heun method,
  \[
  p_{n+1} = p_n - h\nabla V\left(q_n + \frac{h}{2}p_n\right), \quad q_{n+1} = q_n + h\left(p_n - \frac{h}{2}\nabla V(q_n)\right), \tag{33}
  \]

- the standard symplectic Euler method,
  \[
  p_{n+1} = p_n - h\nabla V(q_n), \quad q_{n+1} = q_n + hp_{n+1}, \tag{34}
  \]

- A time transformed version of the symplectic Euler method,
  \[
  p_{n+1} = p_n - \alpha_n h\nabla V(q_n), \quad q_{n+1} = q_n + \alpha_n hp_{n+1}, \quad \alpha_n = 1 + \frac{h}{2}\beta p_n \cdot \nabla V(q_n). \tag{35}
  \]
Figure 2: Invariant measure error (7) for $\phi(p,q) = p^2 + q^2$ of the Lie-Trotter splitting method (4) versus the stepsize $h$ for Langevin dynamics (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 $\Phi_h$: explicit Euler (32), symplectic Euler (35), Heun explicit RK2 (33), and the modified symplectic Euler method (35). The curves are obtained as the averages over 10 long trajectories of time length $T = 10^7$.

The explicit Euler method and the standard symplectic Euler method have deterministic order 1 for (5), while the Heun method has deterministic order 2. We thus observe the the expected lines of slope 1 and 2 in Figure 2, as predicted by Corollary 3.3. The method (35) has deterministic order 1 and preserves the energy and the volume only up to order $r = 1$ in (25), (26), but it has order 2 of accuracy for the invariant measure, as predicted by Theorem 3.6. 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. Notice that the convergence slopes 2 are difficult to identify. This is due to the Monte-Carlo error which is still observed for the high order schemes, in spite of the very long length of the trajectories ($T = 10^7$).

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 (2) along time for the standard symplectic Euler method (34) and the time transformation version (35) as described in Theorem 3.6 for two values of $\gamma$.

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 can be observed, we obtain a drift in the energy for the time transformed version (35) which is non symplectic. 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 3.6. This last numerical example shows that among the standard symplectic Euler method (34) and the time transformed version (35) (applied in the Lie-Trotter splitting (4)), the most accurate scheme is not the same in the deterministic and stochastic contexts, as predicted by Theorem 3.6.
Figure 3: Evolution of the Hamiltonian energy (2) along time for the standard symplectic Euler method (34) and the time transformed version (35) (non symplectic) applied to (1) with $V(q) = (1 - q^2)^2 - \frac{1}{2}q$ (quartic potential). Top picture: deterministic case ($\gamma = 0$). Bottom picture: stochastic case ($\gamma = 4$). Step size is $h = 0.02$ and $\beta = 2$.

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 non symplectic deterministic integrator that, despite its poor geometric properties, captures the invariant measure of Langevin dynamics to second order. Although perhaps counter-intuitive, this last result shows that the most accurate scheme for sampling the invariant measure is not always the one with the best geometric features when applied to deterministic Hamiltonian dynamics.

Acknowledgements. The work of AA and GV was partially supported by the Fonds National Suisse, project No. 200021_140692. and No. 200020_144313/1, respectively.
References


Recent publications:
MATHEMATICS INSTITUTE OF COMPUTATIONAL SCIENCE AND ENGINEERING
Section of Mathematics
Ecole Polytechnique Fédérale
CH-1015 Lausanne

05.2014 BJÖRN ADLERBORN, BO KAGSTRÖM, DANIEL KRÉSSNER:
A parallel QZ algorithm for distributed memory HPC systems

06.2014 MICHELE BENZI, SIMONE DEPARIS, GWENOL GRANDPERRIN, ALFIO QUARTERONI:
Parameter estimates for the relaxed dimensional factorization preconditioner and application to hemodynamics

07.2014 ASSYR ABDULLE, YUN BAI:
Reduced order modelling numerical homogenization

08.2014 ANDREA MANZONI, FEDERICO NEGRE:
Rigorous and heuristic strategies for the approximation of stability factors in nonlinear parametrized PDEs

09.2014 PENG CHEN, ALFIO QUARTERONI:
A new algorithm for high-dimensional uncertainty quantification problems based on dimension-adaptive and reduced basis methods

10.2014 NATHAN COLLIER, ABDUL-LATEEF HAJI-ALI, FABIO NOBILE, ERIK VON SCHWERIN, RAÚL TEMPONE:
A continuation multilevel Monte Carlo algorithm

11.2014 LUKA GRUBISIC, DANIEL KRÉSSNER:
On the eigenvalue decay of solutions to operator Lyapunov equations

12.2014 FABIO NOBILE, LORENZO TAMELLINI, RAÚL TEMPONE:
Convergence of quasi-optimal sparse grid approximation of Hilbert-valued functions: application to random elliptic PDEs

13.2014 REINHOLD SCHNEIDER, ANDRÉ USCHMAJEW:
Convergence results for projected line-search methods on varieties of low-rank matrices via Lojasiewicz inequality

14.2014 DANIEL KRÉSSNER, PETAR SIRKOVIC:
Greedy low-rank methods for solving general linear matrix equations

15.2014 WISSAM HASSAN, MARCO PICASSO:
An anisotropic adaptive finite element algorithm for transonic viscous flows around a wing

16.2014 ABDUL-LATEEF HAJI-ALI, FABIO NOBILE, ERIK VON SCHWERIN, RAÚL TEMPONE:
Optimization of mesh hierarchies in multilevel Monte Carlo samplers

17.2014 DANIEL KRÉSSNER, FRANCISCO MACEDO:
Low-rank tensor methods for communicating Markov processes

18.2014 ASSYR ABDULLE, GILLES VILMART, CONSTANTINOS C. ZYGALAKIS:
Long time accuracy of Lie-Trotter splitting methods for Langevin dynamics