Probabilistic geometric integration of Hamiltonian systems based on random time steps

Assyr Abdulle[*]   Giacomo Garegnani[†]

Abstract

The long-time energy conservation of the random time stepping Runge-Kutta method (RTS-RK) introduced in [A. Abdulle and G. Garegnani, *Random time step probabilistic methods for uncertainty quantification in chaotic and geometric numerical integration*, arXiv:1801.01340, 2018] is studied. Departing from classical backward error analysis tools, we show that the accumulation of errors due to random perturbations causes an energy drift which grows as the square root of time. Nonetheless, we are able to prove that the approximation in the mean sense of the Hamiltonian given by our probabilistic integrator has the same quality as an equivalent deterministic symplectic method over time spans of polynomial length. Numerical examples confirm our theoretical findings and show the effectiveness of the probabilistic approach in the frame of Bayesian inverse problems.

1   Introduction

Hamiltonian dynamics are used to model a variety of physical phenomena, such as celestial orbits or the dynamics of mechanical systems. All Hamiltonian system are endowed with the property of conserving an energy function which is specific to the considered dynamic. Mathematically, Hamiltonian systems can be modelled by an ordinary differential equations (ODE) which for a smooth energy function $Q\colon \mathbb{R}^{2d} \to \mathbb{R}$ and for an initial condition $y_0 \in \mathbb{R}^{2d}$ reads

$$y' = J^{-1}\nabla Q(y), \quad y(0) = y_0 \in \mathbb{R}^{2d}, \tag{1}$$

where the matrix J is defined as

$$J = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}, \tag{2}$$

and where I is the d-dimensional identity matrix. The solution of the ODE above can be expressed in terms of the numerical flow $\varphi_t\colon \mathbb{R}^{2d} \to \mathbb{R}^{2d}$ which maps the initial condition into the solution at a time $t > 0$, i.e., $y(t) = \varphi_t(y_0)$. It is well known that the flow of Hamiltonian systems has the geometric property of conserving certain projected areas in the state space, or, in other words, it is symplectic.

Symplectic Runge-Kutta methods are a class of numerical integrators which are particularly effective for Hamiltonian systems. In fact, the numerical flow of these methods preserves the geometric properties of the exact flow, which guarantees that the solution is accurate over long time spans. In particular, the energy of the system is nearly conserved along the trajectories of the numerical solution, where the accuracy depends on the chosen time step and on the method's order of convergence. Long-time conservation of energy and accuracy are relevant for many applications, e.g., in astronomy or molecular dynamics.

While the properties of energy conservation and long-time accuracy are fundamental for the integration of Hamiltonian systems, classical deterministic methods are associated only with a priori

[*]Mathematics Section, École Polytechnique Fédérale de Lausanne (assyr.abdulle@epfl.ch)

[†]Mathematics Section, École Polytechnique Fédérale de Lausanne (giacomo.garegnani@epfl.ch)

upper bounds for the numerical error, which can be insufficient for describing the method's accuracy in certain situations. In recent years, probabilistic numerical methods for integrating ODEs have been proposed to provide a richer description of the approximation's quality [1,3,5,9,13]. The main idea underlying all these methods is to associate a probability measure to the punctual numerical solution given by classical methods, thus performing a consistent uncertainty quantification of the numerical error. This additional characterisation is particularly effective in the frame of problems of statistical inference, where applying classical schemes can lead to biased posterior concentrations and hence to inadequate solutions [1,4,5].

In our recent work [1] we introduced a probabilistic integrator of ODEs based on a random selection of time steps and explored the geometric properties of probabilistic methods. In particular, we focused on the numerical conservation of first integrals, showing theoretically and by means of numerical experiments that if a deterministic method conserves first integrals then randomisation of the time steps does not affect this property. Moreover, we showed that if the deterministic component has a symplectic flow, then the RTS-RK method has a symplectic flow. This does not imply directly that the energy conservation properties of symplectic integrators with fixed time steps are automatically inherited by randomised integrators as we cannot apply classical backward error analysis for constant time step symplectic methods. In fact, numerical experiments in [1] show that a symplectic RTS-RK method does not preserve the Hamiltonian. We recall that for deterministic symplectic integrators of order q the Hamiltonian is preserved with accuracy $\mathcal{O}(h^q)$ up to exponentially long time intervals.

Nevertheless, using backward error analysis we show in this paper that the Hamiltonian is preserved in the mean for a RTS-RK method built on a symplectic integrator. In particular, we are able to prove exploiting Brouwer's argument [2] that the mean deviation in energy is bounded and proportional to the square root of time. In particular, denoting by q the order of the deterministic method, if the strenght of the noise introduced by the probabilistic method is well-balanced with the numerical error, we show that the mean energy drift is of order $\mathcal{O}(h^q + t^{1/2}h^{2q})$. This result guarantees that for time intervals of size $t = \mathcal{O}(h^{-2q})$ the probabilistic approximation preserves the mean energy with accuracy $\mathcal{O}(h^q)$.

The paper is organised as follows. In section 2 we introduce the technique of backward error analysis for the quantification of errors, focusing on the classical results on energy conservation for Hamiltonian systems and symplectic methods. Then, in section 3 we show the main theoretical result of this work, namely an energy conservation theorem for random time step integration. Finally, in section 4 we present numerical experiments that illustrate our findings and we show an application of our method to Bayesian inverse problems.

2 Background on backward error analysis

In this section we give a quick review on backward error analysis. For more details we refer to [6,10,12]. Let us consider a smooth function $f\colon \mathbb{R}^{2d} \to \mathbb{R}^{2d}$ and the ODE

$$y' = f(y), \quad y(0) = y_0 \in \mathbb{R}^{2d}. \tag{3}$$

We then consider a Runge-Kutta method identified by its numerical flow $\Psi_h\colon \mathbb{R}^{2d} \to \mathbb{R}^{2d}$, such that the numerical solution $y_n = \Psi_h^n(y_0)$ obtained applying n times the flow Ψ_h is an approximation of the solution $y(nh)$. The error introduced by time integration can be analysed via techniques of backward error analysis, whose key ingredient consists in finding a function $\tilde{f}\colon \mathbb{R}^{2d} \to \mathbb{R}^{2d}$ of the form $\tilde{f}(y) = f(y) + hf_2(y) + h^2 f_3(y) + \ldots$, such that the numerical solution is exact when applied to the system

$$y' = \tilde{f}(y), \quad y(0) = y_0. \tag{4}$$

The functions $f_i, i = 2, 3, \ldots$ are built by matching the coefficients of equal powers of h in the expression of $\tilde{f}$ and in the Taylor expansion of Ψ_h, which reads

$$\Psi_h(y) = y + hf(y) + h^2 d_2(y) + h^3 d_3(y) + \ldots, \tag{5}$$

where the functions $d_i, i = 2, 3, \ldots$ depend on the derivatives of f and on the coefficients of the Runge-Kutta method. Once an expression for $\tilde{f}$ is given, the numerical error is then analysed in terms of the difference between the modified and the original equation. A first result states that if the Runge-Kutta integrator has order q, then the lowest power of h in the modified equation is again q, i.e.,

$$\tilde{f}(y) = f(y) + h^q f_{q+1}(y) + h^{q+1} f_{q+2}(y) + \ldots, \tag{6}$$

where the equality above is only formally true, as the series is not guaranteed to converge. In order to perform a rigorous backward error analysis it is hence necessary to choose a value $N > q$ and truncate the sum above after N terms, thus getting

$$\tilde{f}(y) = f(y) + h^q f_{q+1}(y) + h^{q+1} f_{q+2}(y) + \ldots + h^{N-1} f_N(y). \tag{7}$$

We employ a slight abuse of notation, and in the following we always denote by $\tilde{f}$ the truncated modified function.

We next focus on symplectic methods applied to Hamiltonian systems. Let us consider an analytic function $Q \colon \mathbb{R}^{2d} \to \mathbb{R}$ and the associated Hamiltonian system

$$y' = J^{-1} \nabla Q(y), \quad y(0) = y_0 \in \mathbb{R}^{2d}. \tag{8}$$

Let us first recall the definition of a symplectic differentiable map.

Definition 1 (Definition VI.2.2 in [6]). A differentiable map $g \colon U \to \mathbb{R}^{2d}$ (where $U \subset \mathbb{R}^{2d}$ is an open set) is called symplectic if the Jacobian matrix g' is everywhere symplectic, i.e., if

$$(g')^\top J g' = J. \tag{9}$$

The exact flow map φ_t of an Hamiltonian system is symplectic. Moreover, we call a Runge-Kutta integrator symplectic if its flow map Ψ_h is symplectic whenever it is applied to a smooth Hamiltonian system [6]. If a symplectic method is employed to integrate (8), it is known that the modified equation is still Hamiltonian, i.e., there exists a modified energy $\tilde{Q} \colon \mathbb{R}^{2d} \to \mathbb{R}$ defined as

$$\tilde{Q}(y) = Q(y) + h Q_2(y) + h^2 Q_3(y) + \ldots, \tag{10}$$

such that $\tilde{f} = J^{-1} \nabla \tilde{Q}$. For the same reasons as above, it is more appropriate to consider the truncated Hamiltonian, which for an integrator of order q reads

$$\tilde{Q}(y) = Q(y) + h^q Q_{q+1}(y) + \ldots + h^{N-1} Q_N(y). \tag{11}$$

As for the function $\tilde{f}$, let us remark that the notation $\tilde{Q}$ is used in the following to denote only the truncated Hamiltonian. Moreover, we denote by $\tilde{\varphi}_{N,h}$ the exact flow of the dynamical system with Hamiltonian $\tilde{Q}$, highlighting the dependence on the truncation index N and the time step h. In order to show results of energy conservation, it is necessary to introduce an assumption on the regularity of the dynamical system.

Assumption 1. The function f is analytic in a neighbourhood of the initial condition y_0 and there exist constants $M, R > 0$ such that $\|f(y)\| \le M$ for $\|y - y_0\| \le 2R$.

Under this assumption, it is possible to prove that the modified Hamiltonian $\tilde{Q}$ is Lipschitz continuous with a global constant independent of h (see e.g. [6], Section IX.7). Furthermore, a result of local error valid for any dynamical system and any Runge-Kutta method holds.

Lemma 1 (see e.g. [6], Section IX.7). *Under assumption 1 and for h sufficiently small, there exists $N = N(h)$ such that the numerical flow Ψ_h and the exact flow of the truncated modified equation $\tilde{\varphi}_{N,h}$ satisfy*

$$\|\Psi_h(y_0) - \tilde{\varphi}_{N,h}(y_0)\| \le C h e^{-\kappa/h}, \tag{12}$$

where C and κ are positive constants depending only on the method's coefficients and on the regularity of the function f.

In case of Hamiltonian systems and symplectic methods, the existence of a modified Hamiltonian guarantees that the error is small over long time spans. In particular, the following result holds.

Theorem 1 (see e.g. [6], Section IX.8). *Under assumption 1 and for h sufficiently small, the numerical solution y_n given by a symplectic method of order q applied to an Hamiltonian system satisfies*

$$\tilde{Q}(y_n) = \tilde{Q}(y_0) + \mathcal{O}(e^{-\kappa/2h}), \tag{13}$$

$$Q(y_n) = Q(y_0) + \mathcal{O}(h^q), \tag{14}$$

where the two results are valid for exponentially long times.

3   Energy conservation with random time steps

We now consider the random time stepping Runge-Kutta method (RTS-RK) introduced in [1]. Let us recall that this method is defined by the following recurrence relation

$$Y_{n+1} = \Psi_{H_n}(Y_n), \quad Y_0 = y_0, \tag{15}$$

where $\{H_i\}_{i\geq 0}$ is a sequence of independent and identically distributed (i.i.d.) positive random variables. The convergence and some geometric properties of this probabilistic numerical scheme have been studied in [1]. In particular, we report here the main assumption on the time steps which is needed for these properties to hold.

Assumption 2. The i.i.d. random variables H_k satisfy for all $k = 0, 1, \dots$

   (i) $H_k > 0$ a.s.,

   (ii) there exists $h > 0$ such that $\mathbb{E}\, H_k = h$,

   (iii) there exists $p \geq 1/2$ such that the scaled random variables $Z_k := H_k - h$ satisfy

$$\mathbb{E}\, Z_k^2 = Ch^{2p+1}, \tag{16}$$

In this work, we add an assumption on the higher moments of the time steps.

Assumption 3. The i.i.d. random variables H_k satisfy

$$|H_k^p - h^p| \leq M|H - h| \quad a.s., \tag{17}$$

for $p \geq 2$ and a constant $M \geq 0$.

We quickly summarise some results of [1]. Under Assumption 2 and for a deterministic integrator of order q, we have that the mean square convergence of the RTS-RK method is given by

$$\left(\mathbb{E}\|Y_k - y(t_k)\|^2 \right)^{1/2} \leq Ch^{\min\{q,p\}}, \quad k = 0, 1, \dots, \quad t_k = kh, \tag{18}$$

while for any smooth function $\Phi \in \mathbb{C}^\infty(\mathbb{R}^d, \mathbb{R})$ the weak convergence can be expressed as

$$|\mathbb{E}\, \Phi(Y_k) - \Phi(y(t_k))| \leq Ch^{\min\{q,2p\}}, \quad k = 0, 1, \dots, \quad t_k = kh. \tag{19}$$

Let us finally recall the following result on the symplecticity of the flow of the RTS-RK method.

Lemma 2. *If the flow Ψ_h of the deterministic integrator is symplectic, then the flow of the random time-stepping probabilistic method* (15) *is symplectic.*

This result implies that locally the random choice of the time steps does not destroy the geometric property of the underlying deterministic method. Nonetheless, this result is not sufficient to guarantee that energy is well conserved along the trajectories of the probabilistic method.

When integrating (8) with the RTS-RK method, the perturbation introduced by the random choice of the time steps propagates in time following the system's dynamics. The accumulation of errors in the numerical solution of Hamiltonian systems has been treated in several works such as [7, 17], where the main focus has been studying the impact of round-off errors on the conservation of energy when high-order symplectic method are employed. We next prove our main result on the conservation of energy in Hamiltonian system for the RTS-RK method.

Theorem 2. *Under Assumptions 1 and 2 and if the numerical solution Y_n is close enough to the initial condition y_0 almost surely, there exist constants $C_1, C_2 > 0$ independent of h such that the solution given by the RTS-RK method built on a symplectic deterministic method applied to a Hamiltonian system with Hamiltonian Q satisfies*

$$\mathbb{E}|Q(Y_n) - Q(y_0)| \leq C_1 h^q + C_2 \sqrt{t} h^{p+q}. \tag{20}$$

Proof. We exploit the conservation of $\tilde{Q}$ along the trajectories of its corresponding dynamical system, i.e., $\tilde{Q}(\tilde{\varphi}_{N,z}(y)) = \tilde{Q}(y)$ for $y \in \mathbb{R}^d$ and $z > 0$ and employ a telescopic sum to obtain

$$
\begin{aligned}
\mathbb{E}\left(\tilde{Q}(Y_n) - \tilde{Q}(y_0)\right) &= \sum_{j=1}^n \mathbb{E}\left(\tilde{Q}(Y_j) - \tilde{Q}(Y_{j-1})\right) \\
&= \sum_{j=1}^n \mathbb{E}\left(\tilde{Q}(Y_j) - \tilde{Q}(\tilde{\varphi}_{N,h}(Y_{j-1}))\right).
\end{aligned} \tag{21}
$$

We then add and subtract the modified Hamiltonian computed on the numerical flow with fixed time step h to obtain

$$
\begin{aligned}
\mathbb{E}|\tilde{Q}(Y_n) - \tilde{Q}(y_0)| &= \sum_{j=0}^{n-1} \mathbb{E}|\tilde{Q}(Y_{j+1}) - \tilde{Q}(\Psi_h(Y_j)) + \tilde{Q}(\Psi_h(Y_j)) - \tilde{Q}(\tilde{\varphi}_{N,h}(Y_j))| \\
&\leq \sum_{j=0}^{n-1} \mathbb{E}|\tilde{Q}(\Psi_{H_j}(Y_j)) - \tilde{Q}(\Psi_h(Y_j))| + \sum_{j=0}^{n-1} \mathbb{E}|\tilde{Q}(\Psi_h(Y_j)) - \tilde{Q}(\tilde{\varphi}_{N,h}(Y_j))| \\
&\leq \sum_{j=0}^{n-1} \mathbb{E}|I_{1,j}| + \sum_{j=0}^{n-1} \mathbb{E}|I_{2,j}|,
\end{aligned} \tag{22}
$$

where

$$
\begin{aligned}
I_{1,j} &= \tilde{Q}(\Psi_{H_j}(Y_j)) - \tilde{Q}(\Psi_h(Y_j)), \\
I_{2,j} &= \tilde{Q}(\Psi_h(Y_j)) - \tilde{Q}(\tilde{\varphi}_{N,h}(Y_j)).
\end{aligned} \tag{23}
$$

Let us first consider the second term. Thanks to Lemma 1 and since $\tilde{Q}$ is a Lipschitz-continuous function, we have

$$\sum_{j=0}^{n-1} \mathbb{E}|I_{2,j}| \leq nCe^{-\kappa/h}. \tag{24}$$

Hence, this term is bounded by $e^{-\kappa/2h}$ for exponentially long time spans, i.e., for $n \leq e^{\kappa/2h}$. For the first term, we consider the Taylor expansion of the numerical solution, which can be written for any $y \in \mathbb{R}^{2d}$ and $h > 0$ as

$$\Psi_h(y) = y + hJ^{-1}\nabla Q(y) + h^2 d_2(y) + \dots \tag{25}$$

where the d_i functions are analytic and bounded thanks to Assumption 1 (see [6, Theorem IX.7.2]). Moreover, thanks to Assumption 1 we can compute a Taylor expansion of $\tilde{Q}$ and obtain for each $j = 0, \dots, n-1$

$$I_{1,j} = \nabla \tilde{Q}(Y_j)^\top J^{-1} \nabla Q(Y_j)(H_j - h) + \mathcal{O}((H_j^2 - h^2)). \tag{26}$$

Let us denote by $S: \mathbb{R}^{2d} \to \mathbb{R}$ the function $S(y) = Q_{q+1}(y) + hQ_{q+2}(y) + \dots + h^{N-1}Q_N(y)$, i.e., $\tilde{Q}(y) = Q(y) + h^q S(y)$, and remark that

$$
\begin{aligned}
\nabla \tilde{Q}(Y_j)^\top J^{-1} \nabla Q(Y_j) &= \nabla Q(Y_j)^\top J^{-1} \nabla Q(Y_j) + h^q \nabla S(Y_j)^\top J^{-1} \nabla Q(Y_j) \\
&= h^q \nabla S(Y_j)^\top J^{-1} \nabla Q(Y_j),
\end{aligned} \tag{27}
$$

5

since Q is a first integral of motion for (8). Thanks to Assumption 2, 3 and 1, we can write $I_{1,j} = h^q \eta_j$ for a sequence of independent random variables $\{\eta_j\}_{j=0}^{n-1}$ such that $\mathbb{E}\,\eta_j = 0$ and $\mathrm{Var}\,\eta_j \leq Ch^{2p}$ for a constant $C > 0$ independent of h. Hence, by Jensen's inequality we get

$$
\sum_{j=0}^{n-1} \mathbb{E}|I_{1,j}| \leq h^q \Big(\mathbb{E} \big(\sum_{j=0}^{n-1} |\eta_j| \big)^2 \Big)^{1/2} = h^q \Big(\sum_{j=0}^{n-1} \mathbb{E}\,\eta_j^2 \Big)^{1/2}
$$
$$
\leq Cn^{1/2} h^{p+q+1/2} = Ct^{1/2} h^{p+q}. \tag{28}
$$

The desired result is given by Assumption 1 and as by the triangular inequality we have

$$
\mathbb{E}|Q(Y_n) - Q(y_0)| \leq h^q \big(\mathbb{E}|S(Y_n)| + |S(y_0)| \big) + \mathbb{E}|\tilde{Q}(Y_n) - \tilde{Q}(y_0)|
$$
$$
\leq C_1 h^q + C_2 t^{1/2} h^{p+q}. \tag{29}
$$

$\square$

Remark 1. The result of Theorem 2 is consistent with the theory of deterministic symplectic integrators. In fact, in the limit $p \to \infty$, we have

$$
\mathbb{E}|Q(Y_n) - Q(y_0)| = \mathcal{O}(h^q), \tag{30}
$$

and the expectation $\mathbb{E}\,Q(Y_n) \to Q(y_n)$, where y_n is the numerical solution given by the deterministic method.

Remark 2. If we choose the natural scale for the noise $p = q + 1/2$, we have

$$
\mathbb{E}|Q(Y_n) - Q(y_0)| \leq C_3 h^q + C_4 t^{1/2} h^{2q}, \tag{31}
$$

which guarantees that $\mathbb{E}|Q(Y_n) - Q(y_0)| = \mathcal{O}(h^q)$ for time intervals of length $t = \mathcal{O}(h^{-2q})$. Let us remark that this is the same accuracy that would be obtained employing a deterministic symplectic method. Nonetheless, the RTS-RK approximation is valid over polynomial time spans instead of exponential time spans as in the deterministic case.

4 Numerical experiments

In this section we illustrate by two numerical experiments the finding of Theorem 2. Moreover, we consider the inverse problem of retrieving the initial condition of an Hamiltonian system from a single observation of its state at some time $t > 0$, showing how a probabilistic approach can be employed in this framework.

4.1 Convergence

Let us consider the Hénon-Heiles system, which is given by the Hamiltonian $Q \colon \mathbb{R}^4 \to \mathbb{R}$ defined by

$$
Q(v, w) = \frac{1}{2}\|v\|^2 + \frac{1}{2}\|w\|^2 + w_1^2 w_2 - \frac{1}{3}w_2^3, \tag{32}
$$

where $y = (v, w)^\top \in \mathbb{R}^4$. We consider an initial condition such that $Q(y_0) = 0.13$ and integrate the equation employing the RTS-RK method using the two stage Gauss collocation method, which is symplectic and of order $q = 4$, and the noise scale $p = \{2, 4\}$. The time steps are drawn from a uniform distribution $\mathcal{U}(h - h^p, h - h^p)$, so that Assumption 2 is satisfied. We vary the mean time step $h_i = 0.2 \cdot 2^{-i}$ for $i = 0, \ldots, 7$ and consider the final time $T = 10^4$ for both values of p. We then compute the value of Q at final time and compare it with $Q(y_0)$ to check numerically the validity of Theorem 2. Results are shown in Figure 1, where the dashed and dotted lines are given by the theoretical estimation of Theorem 2.
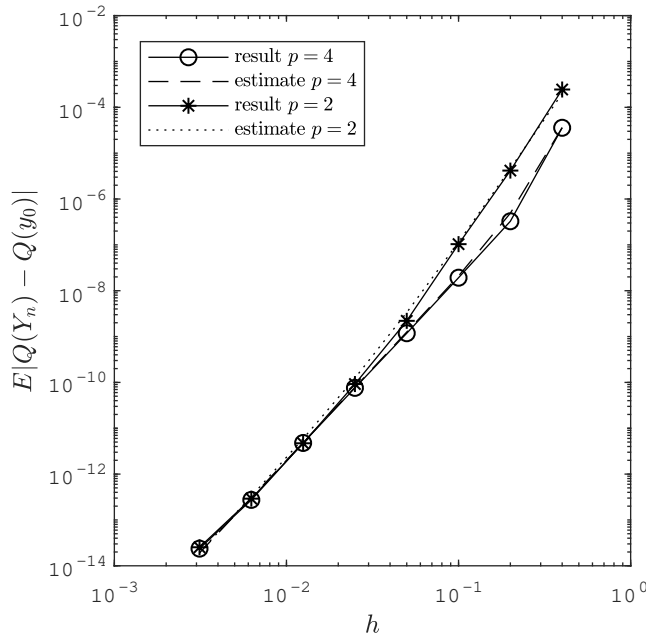
Figure 1: Convergence of the mean error on the Hamiltonian for the Hénon-Heiles problem. The dashed lines represent theoretical estimates given by Theorem 2, while the solid lines represent the experimental results. The mean was computed averaging 100 realisations of the numerical solution.

4.2 Error growth

Let us consider the pendulum problem, which is given by the Hamiltonian $Q \colon \mathbb{R}^2 \to \mathbb{R}$ defined by

$$Q(v, w) = \frac{v^2}{2} - \cos w, \tag{33}$$

where $y = (v, w)^\top \in \mathbb{R}^2$. We consider the initial condition $(v_0, w_0) = (1.5, -\pi)$ and integrate the equation employing RTS-RK based on the implicit midpoint method ($q = 2$) and $p = q = 2$, which is the optimal scaling of the noise. We wish to study the validity of (31), i.e., show that the error on the mean Hamiltonian grows as the square root of time after when the term of order $\mathcal{O}(t^{1/2}h^{2q})$ becomes dominant with respect to the term of order $\mathcal{O}(h^q)$. Hence, we vary the mean time step $h \in \{0.2, 0.1, 0.05, 0.025\}$, integrate the dynamical system up to the final time $T = 10^6$ and study the time evolution of the numerical error on the Hamiltonian Q. Results are shown in Figure 2, where it is possible to notice that the error is growing with a rate $1/2$ in time after a stationary phase, where the term in h^q in (31) is dominating the term in $t^{1/2}h^{2q}$ due to the lower power of h. The oscillations of the error which are shown in Figure 3 are due to the term of order h^q and are present even when integrating the pendulum system deterministically with a symplectic scheme. Moreover, we show in Figure 3 the time evolution up to final time $T = 10^4$ of the mean and of the standard deviation of the energy, which is as well proportional to $t^{1/2}h^{2q}$, with the final distribution of the samples which ressembles a Gaussian distribution.

4.3 Bayesian inverse problem

Let us consider a Hamiltonian system with energy $Q(v, w)$ and exact flow φ_t. We are interested in recovering the true value of the initial condition y_0 through a single observations y_{obs} of the solution (v, w) at time $t = t_{\text{obs}}$. The initial condition is chosen as in Section 4.1 to be such that $Q(y_0) = 0.13$. We assume that the observation is corrupted by an additive source of noise that we describe as a Gaussian random variable $\varepsilon \sim \mathcal{N}(0, \sigma_\varepsilon^2 I)$, i.e., $y_{\text{obs}} = \varphi_{t_{\text{obs}}}(y_0) + \varepsilon$. We interpret this inverse problem in the Bayesian sense and thus specify a Gaussian prior $\pi_{\text{prior}} = \mathcal{N}(0, I)$ on the
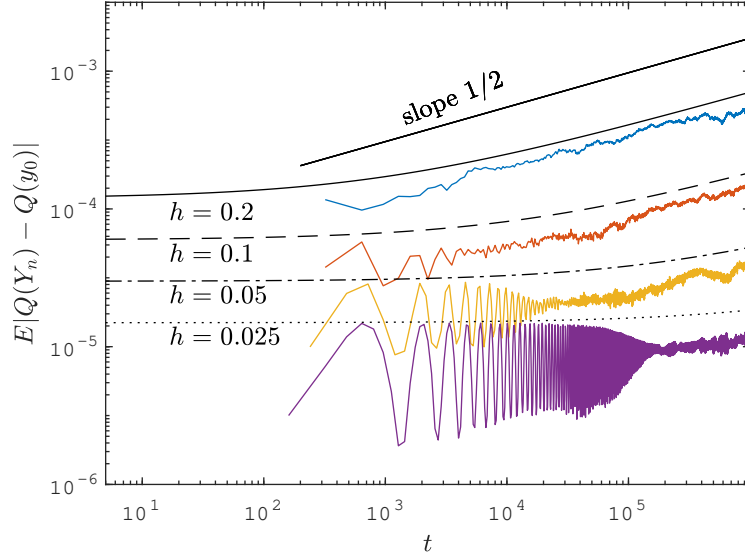
Figure 2: Time evolution of the mean error for the pendulum problem and different values of the time step h. The black lines represent the theoretical estimate given by the left hand side of equation (31), while the colored lines represent the experimental results. The mean was computed averaging 20 realisations of the numerical solution.

initial condition, so that we have by Bayes' theorem that the true posterior distribution reads

$$
\begin{aligned}
\pi(\tilde{y}_0 \mid y_{\mathrm{obs}}) &\propto \pi_{\mathrm{prior}}(\tilde{y}_0)\pi(y_{\mathrm{obs}} \mid \tilde{y}_0) \\
&\propto \exp\left(-\frac{\tilde{y}_0^\top \tilde{y}_0}{2} - \frac{(\varphi_{t_{\mathrm{obs}}}(\tilde{y}_0) - y_{\mathrm{obs}})^\top (\varphi_{t_{\mathrm{obs}}}(\tilde{y}_0) - y_{\mathrm{obs}})}{2\sigma_\varepsilon^2} \right),
\end{aligned}
\tag{34}
$$

where $\pi(y_{\mathrm{obs}} \mid \tilde{y}_0)$ is the likelihood of the observation with respect to an initial condition $\tilde{y}_0$. In order to compute the posterior distribution, it is necessary to introduce a numerical approximation of the equation's flow. If we denote by $\Psi_h^{n_{\mathrm{obs}}}(\tilde{y}_0)$ with $n_{\mathrm{obs}} = t_{\mathrm{obs}}/h$ a deterministic approximation obtained applying n times the numerical flow Ψ_h to the initial condition $\tilde{y}_0$, the approximation of the posterior reads

$$
\pi_{\mathrm{det}}^h(\tilde{y}_0 \mid y_{\mathrm{obs}}) \propto \exp\left(-\frac{\tilde{y}_0^\top \tilde{y}_0}{2} - \frac{(\Psi_h^{n_{\mathrm{obs}}}(\tilde{y}_0) - y_{\mathrm{obs}})^\top (\Psi_h^{n_{\mathrm{obs}}}(\tilde{y}_0) - y_{\mathrm{obs}})}{2\sigma_\varepsilon^2} \right).
\tag{35}
$$

Let us now denote by $\mathbf{H} = (H_0, H_1, \ldots, H_{n_{\mathrm{obs}}})^\top$ the set of random time steps employed for numerical integration with the RTS-RK method. We can compute the posterior distribution marginalising over the random variable $\mathbf{H}$, thus obtaining

$$
\pi_{\mathrm{prob}}^h(\tilde{y}_0 \mid y_{\mathrm{obs}}) \propto \exp\left(-\frac{\tilde{y}_0^\top \tilde{y}_0}{2} \right) \mathbb{E}^{\mathbf{H}} \exp\left(-\frac{(\Psi_{\mathbf{H}}(\tilde{y}_0) - y_{\mathrm{obs}})^\top (\Psi_{\mathbf{H}}(\tilde{y}_0) - y_{\mathrm{obs}})}{2\sigma_\varepsilon^2} \right),
\tag{36}
$$

where $\Psi_{\mathbf{H}}(\tilde{y}_0) = (\Psi_{H_{n_{\mathrm{obs}}-1}} \circ \Psi_{H_{n_{\mathrm{obs}}-2}} \circ \ldots \circ \Psi_{H_0})(\tilde{y}_0)$ is the probabilistic numerical approximation of $\varphi_{t_{\mathrm{obs}}}(\tilde{y}_0)$ given by the RTS-RK method and $\mathbb{E}^{\mathbf{H}}$ is the expectation with respect to the random variable $\mathbf{H}$. Let us remark that the convergence with respect to $h \to 0$ of the distributions defined in (35) and (36) to the true posterior distribution (34) is guaranteed under some regularity assumptions over the exact forward model, i.e., the exact flow $\varphi_{t_{\mathrm{obs}}}$ (see e.g., [15] Section 4 for the deterministic approximation and [11] for a general framework with random posterior distributions). Let us consider the Hénon-Heiles equation, given by the Hamiltonian (32). A classical second-order ($q = 2$) symplectic method is the Störmer-Verlet scheme [6, 14, 16], for which one step is defined in the
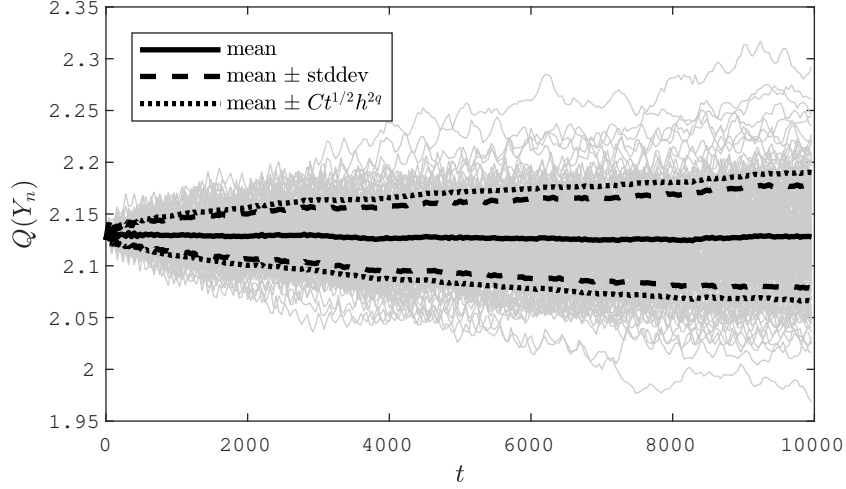
8

Figure 3: Samples from the numerical solution and time evolution of the standard deviation. Solid gray lines represent the energy over all the 200 considered trajectories, the solid black line is the mean energy, the dashed line is the standard deviation and the dotted line is the theoretical estimation of the energy deviation.

general case as

$$
v_{n+1/2} = v_n - \frac{h}{2}\nabla_w Q(v_n, w_n),
$$
$$
w_{n+1} = w_n + \frac{h}{2}\big(\nabla_v Q(v_{n+1/2}, w_n) + \nabla_v Q(v_{n+1/2}, w_{n+1})\big), \tag{37}
$$
$$
v_{n+1} = v_{n+1/2} - \frac{h}{2}\nabla_w Q(v_{n+1/2}, w_{n+1}).
$$

As the Hamiltonian Q given by (32) is separable, i.e., $Q(v, w) = Q_1(v) + Q_2(w)$, where $Q_1, Q_2 \colon \mathbb{R}^2 \to \mathbb{R}$, the Störmer-Verlet scheme simplifies to

$$
v_{n+1/2} = v_n - \frac{h}{2}\nabla_w Q_2(w_n),
$$
$$
w_{n+1} = w_n + h\nabla_v Q_1(v_{n+1/2}), \tag{38}
$$
$$
v_{n+1} = v_{n+1/2} - \frac{h}{2}\nabla_w Q_2(w_{n+1}).
$$

Hence, the evaluation of the flow of the RTS-RK method consists only of three evaluations of the derivatives of Q, instead of the solution of a nonlinear system.

In order to sample from the posterior distributions π_{det}^h and π_{prob}^h we employ the Metropolis-Hastings (MH) and the pseudo-marginal Metropolis-Hastings (PMMH) algorithms. Given an initial guess $\tilde{y}_0^{(0)}$, these algorithms fix a proposal distribution $q \colon \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ which is measurable in its first argument and a probability density function in its second argument and produce M samples from the posterior distribution by the following procedure

 (i) Compute $\pi(\tilde{y}_0^{(0)} \mid y_{\mathrm{obs}})$ and set $i = 0$;

 (ii) Propose a new value $\tilde{y}_0^\dagger$ sampling from $q(\tilde{y}_0^{(i)}, \cdot)$;

 (iii) With probability

$$
\alpha(\tilde{y}_0^{(i)}, \tilde{y}_0^\dagger) = \min\left\{ \frac{\pi(\tilde{y}_0^\dagger \mid y_{\mathrm{obs}})q(\tilde{y}_0^\dagger, \tilde{y}_0^{(i)})}{\pi(\tilde{y}_0^{(i)} \mid y_{\mathrm{obs}})q(\tilde{y}_0^{(i)}, \tilde{y}_0^\dagger)}, 1 \right\} \tag{39}
$$

set $\tilde{y}_0^{(i+1)} = \tilde{y}_0^\dagger$, otherwise set $\tilde{y}_0^{(i+1)} = \tilde{y}_0^{(i)}$;

 (iv) if $i < M$, set $i \leftarrow i + 1$ and return to step (ii).

9

Further details on how to employ these algorithms for RTS-RK methods can be found in [1]. We then consider $t_{\text{obs}} = 10$, the standard deviation of the observational noise $\sigma_\varepsilon = 5 \cdot 10^{-4}$ and the time step $h = \{0.2, 0.1, 0.05, 0.025\}$ for numerical integration. The approximated marginal posterior distributions of $y_0 = (v_{0,1}, v_{0,2}, w_{0,1}, w_{0,2})$ are shown in Figure 4. In particular, we have

- Figure 4a: Ψ_h is approximated via Heun's method [8] ($q = 2$), which is given by the numerical flow
$$\Psi_h(y_n) = y_n + \frac{h}{2}\Big(f(y_n) + f\big(y_n + hf(y_n)\big) \Big),$$

- Figure 4b: Ψ_h is approximated via the Störmer-Verlet method,

- Figure 4c: Ψ_h is approximated via the RTS-RK Störmer-Verlet method with the natural choice $p = q + 1/2 = 5/2$.

We can observe that the posterior distributions given by Heun's method are concentrated away from the true value of the initial condition for the larger values of the time step. In fact, Heun's method is not symplectic, and a deviation on the energy Q is produced when integrating in time the dynamical system. Hence, initial conditions with a different energy level with respect to the observation are endowed with a high value of likelihood and the resulting posterior distribution is concentrated far from the true value. This behaviour is corrected using the Störmer-Verlet method thanks to its symplecticity. However, we remark in Figure 4b that the posterior distribution for $h = 0.2$ is still concentrated on a biased value of the initial condition, without any indication of this bias given by the posterior's variance. Applying the RTS-RK method together with PMMH instead gives nested posterior distributions whose variance quantifies the uncertainty of the numerical solver. This favourable behaviour is possible thanks to the numerical error quantification of probabilistic methods, which has been already shown in [1, 4, 5], together with the good energy conservation properties of the RTS-RK method when a symplectic integrator is used as its deterministic component as proved in Theorem 2.

References

[1] A. ABDULLE AND G. GAREGNANI, *Random time step probabilistic methods for uncertainty quantification in chaotic and geometric numerical integration*, arXiv preprint arXiv:1801.01340, (2018).

[2] D. BROUWER, *On the accumulation of errors in numerical integration*, Astronomical Journal, 46 (1937), pp. 149–153.

[3] O. A. CHKREBTII, D. A. CAMPBELL, B. CALDERHEAD, AND M. A. GIROLAMI, *Bayesian solution uncertainty quantification for differential equations*, Bayesian Anal., 11 (2016), pp. 1239–1267.

[4] J. COCKAYNE, C. OATES, T. SULLIVAN, AND M. GIROLAMI, *Probabilistic numerical methods for PDE-constrained Bayesian inverse problems*, AIP Conference Proceedings, 1853 (2017), p. 060001.

[5] P. R. CONRAD, M. GIROLAMI, S. SÄRKKÄ, A. STUART, AND K. ZYGALAKIS, *Statistical analysis of differential equations: introducing probability measures on numerical solutions*, Stat. Comput., (2016).

[6] E. HAIRER, C. LUBICH, AND G. WANNER, *Geometric Numerical Integration. Structure-Preserving Algorithms for Ordinary Differential Equations*, Springer Series in Computational Mathematics 31, Springer-Verlag, Berlin, second ed., 2006.

[7] E. HAIRER, R. MCLACHLAN, AND A. RAZAKARIVONY, *Achieving Brouwer's law with implicit Runge-Kutta methods*, BIT, 48 (2008), pp. 231–243.

(a) $\pi_{\text{det}}^h(y_0 \mid y_{\text{obs}})$, Heun

(b) $\pi_{\text{det}}^h(y_0 \mid y_{\text{obs}})$, Störmer-Verlet

(c) $\pi_{\text{prob}}^h(y_0 \mid y_{\text{obs}})$, RTS-RK Störmer-Verlet
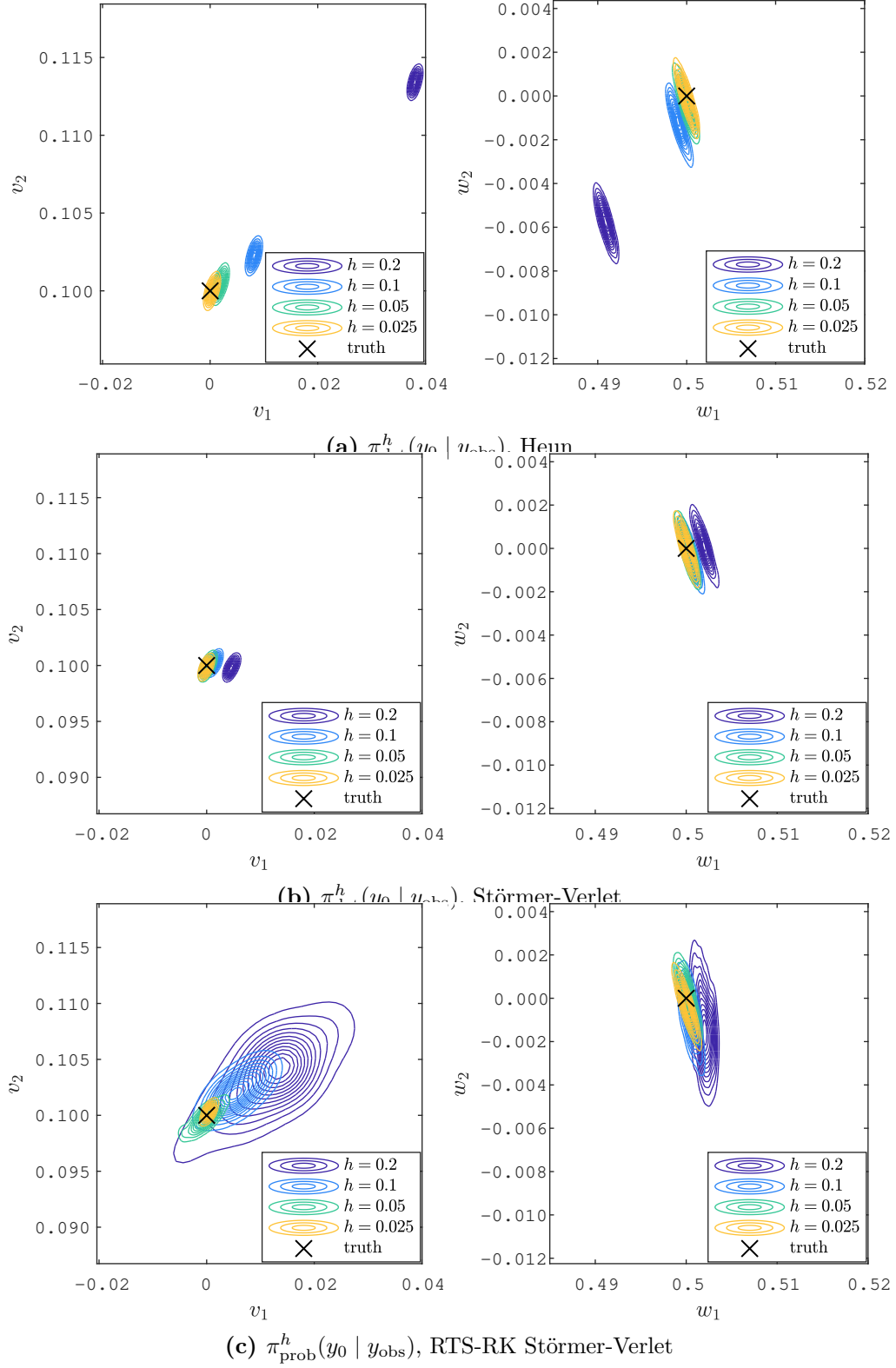
Figure 4: Deterministic posterior distribution for Heun's and Störmer-Verlet schemes and probabilistic posterior distribution for the RTS-RK Störmer-Verlet scheme for various values of the time step h. The cross corresponds to the true value of the initial condition that is employed to generate the observation.

11

[8] K. Heun, *Neue Methode zur approximativen Integration der differentialgleichungen einer unabhängigen veränderlichen*, Zeitschr. für Math. u. Phys., 45 (1900), pp. 23–38.

[9] H. Kersting and P. Hennig, *Active uncertainty calibration in Bayesian ODE solvers*, in Proceedings of the 32nd Conference on Uncertainty in Artificial Intelligence (UAI 2016), AUAI Press, 2016, pp. 309–318.

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

[11] H. C. Lie, T. Sullivan, and A. Teckentrup, *Random forward models and log-likelihoods in bayesian inverse problems*, arXiv preprint arXiv:1712.05717, (2017).

[12] J. M. Sanz-Serna and M. P. Calvo, *Numerical Hamiltonian Problems*, Chapman & Hall, London, 1994.

[13] M. Schober, S. Särkkä, and P. Hennig, *A probabilistic model for the numerical solution of initial value problems*, Stat. Comput., (2018).

[14] C. Störmer, *Sur les trajectoires des corpuscules électrisés*, Arch. sci. phys. nat. Genève, 24 (1907), pp. 5–18, 113–158, 221–247.

[15] A. M. Stuart, *Inverse problems: A Bayesian perspective*, Acta Numer., 19 (2010), pp. 451–559.

[16] L. Verlet, *Computer "experiments" on classical fluids. i. thermodynamical properties of Lennard-Jones molecules*, Physical Review, 159 (1967), pp. 98–103.

[17] G. Vilmart, *Reducing round-off errors in rigid body dynamics*, J. Comput. Phys., 227 (2008), pp. 7083–7088.