# A PROBABILISTIC INTEGRATOR OF ORDINARY DIFFERENTIAL EQUATIONS BASED ON RUNGE-KUTTA METHODS WITH RANDOM SELECTION OF THE TIME STEPS

ASSYR ABDULLE* AND GIACOMO GAREGNANI†

**Abstract.**

## 1. Introduction.

## 2. Random time stepping.
Let us consider a Lipschitz function $f \colon \mathbb{R}^d \to \mathbb{R}^d$ and the ODE

$$(1) \qquad y' = f(y), \quad y(0) = y_0 \in \mathbb{R}^d.$$

We can write the solution $y(t)$ in terms of the flow of the ODE, i.e., the family $\{\varphi_t\}_{t \geq 0}$ of functions $\varphi_t \colon \mathbb{R}^d \to \mathbb{R}^d$ such that

$$(2) \qquad y(t) = \varphi_t(y_0).$$

Let us consider a Runge-Kutta method for (1). Given a time step $h$, we can write the numerical solution at $y_n$ approximation of $y(t_n)$, with $t_n = nh$ in terms of the numerical flow $\{\Psi_t\}_{t \geq 0}$, with $\Psi_t \colon \mathbb{R}^d \to \mathbb{R}^d$, which is uniquely determined by the coefficients of the method, as

$$(3) \qquad y_{k+1} = \Psi_h(y_k).$$

If the ODE is chaotic, integrating the equation with different time steps leads to different numerical solutions. In order to describe the uncertainty of the numerical solution, we choose at each time $t_k$ the time step as the realization of a random variable $H_k$. Therefore, the numerical solution is given by a discrete stochastic process $\{Y_k\}_{k \geq 0}$ such that

$$(4) \qquad Y_{k+1} = \Psi_{H_k}(Y_k).$$

This probabilistic numerical method shares the basic idea with [1].

## 3. Strong order analysis.
We wish to deduce a result of strong convergence for the numerical method (4).

DEFINITION 3.1. *The numerical method (4) has strong local order $r$ for (1) if there exists a constant $C > 0$ such that*

$$(5) \qquad \mathbb{E}|Y_1 - y(h)| \leq Ch^{r+1}.$$

In the following we will present the assumptions needed to prove a result of strong convergence. First of all, the following assumption on the random variables $\{H_k\}_{k \geq 0}$ is needed for strong convergence.

*Assumption* 3.2. The i.i.d. random variables $H_k$ satisfy for all $k = 1, 2, \dots$

   (i) $H_k > 0$ a.s.,
   (ii) there exists $h > 0$ such that $\mathbb{E} H_k = h$,
   (iii) there exists $p \geq 1$ such that the scaled random variables $Z_k := H_k - h$ satisfy

$$(6) \qquad \mathbb{E} Z_k^2 = Ch^{2p},$$

   which is equivalent to $\mathbb{E} H_k^2 = h^2 + Ch^{2p}$.

---

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

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

The class of random variable satisfying the hypotheses above is general. However, it is practical for an implementation point of view to have examples of these variables.

*Example* 3.3. An example of random variables $H_k$ satisfying the properties above is given by

$$(7) \qquad\qquad H_k \overset{\text{i.i.d.}}{\sim} \mathcal{U}(h - h^p, h + h^p), \quad 0 < h \le 1, \quad p \ge 1.$$

Let us verify all the assumptions above for this choice of the time steps

(i) $H_k > 0$ a.s. trivially since $h \le 1$,

(ii) $\mathbb{E}\, H_k = h$ since

$$(8) \qquad\qquad \mathbb{E}\, H_k = \frac{1}{2}(h + h^p + h - h^p) = h.$$

(iii) The random variables $Z_k = H_k - h$ are $Z_k \sim \mathcal{U}(-h^p, h^p)$. Therefore

$$(9) \qquad\qquad \mathbb{E}\, Z_k^2 = \frac{4h^{2p}}{12} = \frac{1}{3} h^{2p}.$$

The following assumption on the numerical flow is needed.

*Assumption* 3.4. The Runge-Kutta method defined by the numerical flow $\{\Psi_t\}_{t \ge 0}$ satisfies the following properties.

(i) For $h$ small enough, there exists a constant $C > 0$ such that

$$(10) \qquad\qquad |\Psi_h(y) - \varphi_h(y)| \le Ch^{q+1}, \quad \forall y \in \mathbb{R}^d.$$

(ii) The map $t \mapsto \Psi_t(y)$ is of class $\mathcal{C}^2(\mathbb{R}^+, \mathbb{R}^d)$ and is Lipschitz continuous of constant $L_\Psi$, i.e.,

$$(11) \qquad\qquad |\Psi_t(y) - \Psi_s(y)| \le L_\Psi |t - s|, \quad \forall t, s > 0.$$

(iii) There exists a constant $C > 0$ independent of $h$ such that

$$(12) \qquad\qquad |\Psi_h(y) - \Psi_h(z)| \le (1 + Ch)|y - z|, \quad \forall y, z \in \mathbb{R}^d.$$

The local strong convergence of (4) can now be proved.

THEOREM 3.5 (Strong local order). *Under Assumptions 3.2 and 3.4 the numerical solution $Y_1$ given by one step of* (4) *satisfies*

$$(13) \qquad\qquad \mathbb{E}|Y_1 - y(h)| \le Ch^{\min\{q+1, p\}},$$

*where $C$ is a real positive constant independent of $h$ and the coefficients $p$, $q$ are given in the assumptions.*

*Proof.* By the triangular inequality we have for all $y \in \mathbb{R}^d$

$$(14) \qquad\qquad \mathbb{E}|\Psi_{H_0}(y) - \varphi_h(y)| \le \mathbb{E}|\Psi_{H_0}(y) - \Psi_h(y)| + \mathbb{E}|\Psi_h(y) - \varphi_h(y)|.$$

We now consider Assumption 3.4.(ii) and 3.4.(i), thus getting

$$(15) \qquad\qquad \mathbb{E}|\Psi_{H_0}(y) - \varphi_h(y)| \le L_\Psi \, \mathbb{E}|H_0 - h| + C_1 h^{q+1}.$$

We now apply Jensen's inequality and Assumption 3.2 to obtain

$$(16) \qquad\begin{aligned} \mathbb{E}|\Psi_{H_0}(y) - \varphi_h(y)| &\le L(\mathbb{E}|H_0 - h|^2)^{1/2} + C_1 h^{q+1} \\ &\le Lh^p + C_1 h^{q+1} \\ &\le Ch^{\min\{q+1, p\}}, \end{aligned}$$

which is the desired result. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

As a consequence of the one-step convergence, we can prove a result of strong global convergence.

THEOREM 3.6 (Strong global order). *If $t_k = kh$ for $k = 1, 2, \ldots, N$, where $Nh = T$, and under Assumptions 3.2 and 3.4 the numerical solution given by (4) satisfies*

$$(17) \qquad \sup_{k=1,2,\ldots,N} \mathbb{E}|Y_k - y(t_k)| \leq Ch^{\min\{q, p-1/2\}},$$

*where $C$ is a real positive constant independent of $h$.*

*Proof.* Let us define $e_k := \mathbb{E}|Y_k - y(t_k)|$. Thanks to the triangular inequality we have

$$(18) \qquad e_k \leq \mathbb{E}|\Psi_{H_{k-1}}(Y_{k-1}) - \Psi_{H_{k-1}}(y(t_{k-1}))| + \mathbb{E}|\Psi_{H_{k-1}}(y(t_{k-1})) - \varphi_h(y(t_{k-1}))|.$$

We then apply Assumption 3.4.(iii) for the first term and Theorem 3.5 for the second term, thus obtaining

$$(19) \qquad e_k \leq (1 + Ch)e_{k-1} + Ch^{\min\{q+1, p\}}.$$

Hence, iterating over $k$ and noticing that $e_0 = 0$, we get

$$(20) \qquad e_k \leq Ch^{\min\{q+1, p\}} \sum_{i=0}^{k-1}(1 + Ch)^i.$$

We then remark that

$$(21) \qquad \sum_{i=0}^{k-1}(1 + Ch)^i \leq Te^{CT}h^{-1},$$

which proves the desired result. $\square$

**4. Weak order analysis.** We apply techniques of backwards error analysis. Let us introduce the operators $\mathcal{L}$ and $\mathcal{L}^h$ such that, if $Y_1$ is the first value produced by the numerical scheme described above,

$$(22) \qquad \begin{aligned} \Phi(\varphi_h(y)) &= e^{h\mathcal{L}}\Phi(y), \\ \mathbb{E}\,\Phi(Y_1 \mid Y_0 = y) &= e^{h\mathcal{L}^h}\Phi(y), \end{aligned}$$

for all functions $\Phi$ in $\mathcal{C}^\infty(\mathbb{R}^d, \mathbb{R})$. In particular, we can write explicitly $\mathcal{L} = f \cdot \nabla$, while no closed form expression for $\mathcal{L}^h$ is available. We now expand the functional of the numerical solution as follows

$$(23) \qquad \begin{aligned} \Phi(Y_1) &= \Phi(\Psi_{H_0}(Y_0)) \\ &= \Phi\Big(\Psi_h(Y_0) + (H_0 - h)\partial_t\Psi_h(Y_0) + \frac{1}{2}(H_0 - h)^2\partial_{tt}\Psi_h(Y_0) + \mathcal{O}(|H_0 - h|^3)\Big) \\ &= \Phi(\Psi_h(Y_0)) + \Big((H_0 - h)\partial_t\Psi_h(Y_0) + \frac{1}{2}(H_0 - h)^2\partial_{tt}\Psi_h(Y_0)\Big) \cdot \nabla\Phi(\Psi_h(Y_0)) \\ &\quad + \frac{1}{2}(H_0 - h)^2\partial_t\Psi_h(Y_0)\partial_t\Psi_h(Y_0)^T : \nabla^2\Phi(\Psi_h(Y_0)) + \mathcal{O}(|H_0 - h|^3), \end{aligned}$$

where we denote by $\nabla^2\Phi$ the Hessian matrix of $\Phi$, and by : the inner product on matrices induced by the Frobenius norm on $\mathbb{R}^d$, i.e., $A : B = \text{tr}(A^T B)$. Taking the conditional expectation with respect to $Y_0 = y$ we get

$$(24) \qquad \begin{aligned} e^{h\mathcal{L}^h}\Phi(y) - \Phi(\Psi_h(y)) &= \frac{1}{2}Ch^{2p}\partial_{tt}\Psi(h, y) \cdot \nabla\Phi(\Psi(h, y)) \\ &\quad + \frac{1}{2}Ch^{2p}\partial_t\Psi_h(y)\partial_t\Psi_h(y)^T : \nabla^2\Phi(\Psi_h(y)) + \mathcal{O}(h^{3p}), \end{aligned}$$

where we exploited Hölder inequality for the last term. Moreover, expanding $\Phi$ in $y$ we get

$$
\begin{aligned}
\Phi(\Psi(h,y)) &= \Phi\left(\Psi_0(y) + h\partial_t\Psi_0(y) + \mathcal{O}(h^2)\right) \\
&= \Phi(y) + \mathcal{O}(h).
\end{aligned}
\tag{25}
$$

which implies

$$
\begin{aligned}
e^{h\mathcal{L}^h}\Phi(y) - \Phi(\Psi_h(y)) &= \frac{1}{2}Ch^{2p}\partial_{tt}\Psi_h(y)\cdot\nabla\Phi(y) \\
&\quad + \frac{1}{2}Ch^{2p}\partial_t\Psi_h(y)\partial_t\Psi_h(y)^T : \nabla^2\Phi(y) + \mathcal{O}(h^{2p+1}).
\end{aligned}
\tag{26}
$$

Consider now the following modified ODE and modified SDE

$$
\hat{y}' = f^h(\hat{y}),
\tag{27}
$$

$$
\mathrm{d}\tilde{y} = \left(f^h(\tilde{y}) + \frac{1}{2}Ch^{2p-1}\partial_{tt}\Psi_h(\tilde{y})\right)\mathrm{d}t + \sqrt{Ch^{2p-1}\partial_t\Psi_h(\tilde{y})\partial_t\Psi_h(\tilde{y})^T}\mathrm{d}W_t,
\tag{28}
$$

where $W_t$, with $t \geq 0$, is a standard $d$-dimensional Wiener process. These equations define the differential operators $\hat{\mathcal{L}}$ and $\tilde{\mathcal{L}}$ such that

$$
\Phi(\hat{y}(h) \mid \hat{y}(0) = y) = (e^{h\hat{\mathcal{L}}}\Phi)(y), \quad \mathbb{E}\,\Phi(\tilde{y}(h) \mid \tilde{y}(0) = y) = (e^{h\tilde{\mathcal{L}}}\Phi)(y).
\tag{29}
$$

The operators $\hat{\mathcal{L}}$ and $\tilde{\mathcal{L}}$ can be written explicitly from (27) and (28) as

$$
\hat{\mathcal{L}} = f^h\cdot\nabla, \quad \tilde{\mathcal{L}} = \left(f^h + \frac{1}{2}Ch^{2p-1}\partial_{tt}\Psi_h\right)\cdot\nabla + \frac{1}{2}Ch^{2p-1}\partial_t\Psi_h\partial_t\Psi_h^T : \nabla^2.
\tag{30}
$$

Let us remark that the function $f^h$ is of the form

$$
f^h = f + \sum_{i=q}^{q+l} h^i f_i,
\tag{31}
$$

where $q$ is given in Assumption 3.4.(i) and the terms $f_i$ are chosen such that

$$
e^{h\hat{\mathcal{L}}}\Phi(y) - \Phi(\Psi_h(y)) = \mathcal{O}(h^{q+2+l}).
\tag{32}
$$

It has been shown [2] that such a function always exists for Runge-Kutta methods. Let us compute the distance between the generator of the SDE and the modified ODE

$$
e^{h\tilde{\mathcal{L}}}\Phi(y) - e^{h\hat{\mathcal{L}}}\Phi(y) = e^{hf^h\cdot\nabla}(e^{\frac{1}{2}Ch^{2p}(\partial_{tt}\Psi_h\cdot\nabla + \partial_t\Psi_h\partial_t\Psi_h^T : \nabla^2)} - I)\Phi(y).
\tag{33}
$$

Expanding with Taylor the two factors we get

$$
\begin{aligned}
e^{h\tilde{\mathcal{L}}}\Phi(y) - e^{h\hat{\mathcal{L}}}\Phi(y) &= (I + \mathcal{O}(h))\left(\frac{1}{2}Ch^{2p}\partial_{tt}\Psi_h\cdot\nabla + \frac{1}{2}Ch^{2p}\partial_t\Psi_h\partial_t\Psi_h^T : \nabla^2 + \mathcal{O}(h^{4p})\right)\Phi(y) \\
&= \frac{1}{2}Ch^{2p}\partial_{tt}\Psi_h(y)\cdot\nabla\Phi(y) + \frac{1}{2}Ch^{2p}\partial_t\Psi_h(y)\partial_t\Psi_h(y)^T : \nabla^2\Phi(y) + \mathcal{O}(h^{2p+1})
\end{aligned}
\tag{34}
$$

Now considering (26) and (32), we have

$$
\begin{aligned}
e^{h\mathcal{L}^h}\Phi(y) - e^{h\hat{\mathcal{L}}}\Phi(y) &= \frac{1}{2}Ch^{2p}\partial_{tt}\Psi_h(y)\cdot\nabla\Phi(y) + \frac{1}{2}Ch^{2p}\partial_t\Psi_h(y)\partial_t\Psi_h(y)^T : \nabla^2\Phi(y) \\
&\quad + \mathcal{O}(h^{2p+1}) + \mathcal{O}(h^{q+2+l}).
\end{aligned}
\tag{35}
$$

Then, combining (34) and (35) we then get the distance between the probabilistic method and the solution of the SDE as

$$(36) \qquad e^{h\tilde{\mathcal{L}}}\Phi(y) - e^{h\mathcal{L}^h}\Phi(y) = \mathcal{O}(h^{2p+1}) + \mathcal{O}(h^{q+2+l}).$$

Choosing $l = 2p - 1 - q$ we have the balance between the two terms and

$$(37) \qquad e^{h\tilde{\mathcal{L}}}\Phi(y) - e^{h\mathcal{L}^h}\Phi(y) = \mathcal{O}(h^{2p+1}),$$

which is the one-step weak error between the probabilistic numerical method and the modified SDE. For the original equation, let us remark that thanks to Assumption 3.4.(i) we have

$$(38) \qquad e^{h\mathcal{L}}\Phi(y) - \Phi(\Psi(h, y)) = \mathcal{O}(h^{q+1}).$$

Combining (38) and (26) we have the one-step weak error of the probabilistic method on the original ODE, i.e.,

$$(39) \qquad e^{h\mathcal{L}}\Phi(y) - e^{h\mathcal{L}^h}\Phi(y) = \mathcal{O}(h^{\min\{2p,q+1\}}).$$

In order to obtain a result on the global order of convergence we need a further stability assumption, which is the same as Assumption 3 in [1].

*Assumption* 4.1. The function $f$ is in $\mathcal{C}^\infty(\mathbb{R}^d, \mathbb{R}^d)$ and all its derivatives are uniformly bounded on $\mathbb{R}^d$. Furthermore, $f$ is such that the operators $e^{h\mathcal{L}}$ and $e^{h\mathcal{L}^h}$ satisfy, for all functions $\Phi \in \mathbb{C}^\infty(\mathbb{R}^d, \mathbb{R})$ and a positive constant $L$

$$(40) \qquad \begin{aligned} \sup_{u\in\mathbb{R}^d} |e^{h\mathcal{L}}\Phi(u)| &\leq (1 + Lh) \sup_{u\in\mathbb{R}^d} |\Phi(u)|, \\ \sup_{u\in\mathbb{R}^d} |e^{h\mathcal{L}^h}\Phi(u)| &\leq (1 + Lh) \sup_{u\in\mathbb{R}^d} |\Phi(u)|. \end{aligned}$$

We can now state the main result on weak convergence. Let us remark that the theorem and its proof are similar to Theorem 2.4 in [1].

THEOREM 4.2. *Under Assumptions 3.2, 3.4 and 4.1 there exists a constant $C > 0$ such that for all functions $\varphi \colon \mathbb{R}^d \to \mathbb{R}$ in $\mathcal{C}^\infty(\mathbb{R}^d, \mathbb{R})$*

$$(41) \qquad |\varphi(u(T)) - \mathbb{E}\,\varphi(U_N)| \leq Ch^{\min\{2p-1,q\}},$$

*where $u(t)$ is the solution of (1) and $T = Nh$. Moreover,*

$$(42) \qquad |\varphi(\tilde{u}(T)) - \mathbb{E}\,\varphi(U_N)| \leq Ch^{2p},$$

*where $\tilde{u}(t)$ is the solution of (28).*

*Proof.* Let us introduce the following notation

$$(43) \qquad \begin{aligned} w_k &= \Phi(y(t_k) \mid y(0) = y_0) \\ W_k &= \mathbb{E}\,\Phi(Y_k \mid Y_0 = y_0). \end{aligned}$$

Then by the triangular inequality and the Markov property we have

$$(44) \qquad |w_k - W_k| \leq |e^{h\mathcal{L}}w_{k-1} - e^{h\mathcal{L}^h}w_{k-1}| + |e^{h\mathcal{L}^h}w_{k-1} - e^{h\mathcal{L}^h}W_{k-1}|.$$

Applying (39) to the first term and Assumption 4.1 to the second, we have

$$(45) \qquad |w_k - W_k| \leq Ch^{\min\{2p,q+1\}} + (1 + Lh)|w_{k-1} - W_{k-1}|.$$

Proceeding iteratively on the index $k$ and noticing that $w_0 = W_0$, we obtain

$$
(46) \qquad \begin{aligned} |w_k - W_k| &\leq C k h^{\min\{2p, q+1\}} \\ &\leq C T h^{\min\{2p-1, q\}}, \end{aligned}
$$

which proves the first inequality. The proof of the second inequality follows the same steps as above considering $w_k = \mathbb{E}\,\Phi(\tilde{y} \mid \tilde{y}(0) = y_0)$ and applying (37). $\qquad\square$

**5. Monte Carlo estimators.** We are interested in understanding the behavior of Monte Carlo estimator drawn from the numerical solution (4). Given a function $\Phi\colon \mathbb{R}^d \to \mathbb{R}$ which is for simplicity $\mathcal{C}^\infty(\mathbb{R}^d, \mathbb{R})$ with Lipschitz constant $L_\Phi$, we consider the mean square error (MSE) of the estimator $\hat{Z}$ defined as

$$
(47) \qquad \hat{Z} = M^{-1} \sum_{i=1}^M \Phi(Y_N^{(i)}),
$$

where $T = hN$ is the final time, $M$ is the number of trajectories and we denote by $Y_N^{(i)}$ the realizations of the solution for $i = 1, 2, \ldots, M$. It is trivial to remark that

$$
(48) \qquad \begin{aligned} \mathrm{MSE}(\hat{Z}) &= \mathbb{E}(\hat{Z} - Z)^2 \\ &= \mathrm{Var}(\hat{Z}) + \big(\mathbb{E}(\hat{Z} - Z)\big)^2. \end{aligned}
$$

Hence, thanks to the result of Theorem 4.2, we have

$$
(49) \qquad \mathrm{MSE}(\hat{Z}) = \mathrm{Var}(\hat{Z}) + \mathcal{O}(h^{2\min\{q, 2p-1\}}).
$$

The variance of the estimator can be trivially bounded exploiting the Lipschitz continuity of $\Phi$ and the independence of the samples by

$$
(50) \qquad \mathrm{Var}\,\hat{Z} \leq M^{-1} L_\Phi^2 \,\mathrm{Var}\,Y_N.
$$

We can now prove a bound for the MSE of the Monte Carlo estimator.

THEOREM 5.1. *Under Assumptions 3.2 and 3.4, the MSE of the Monte Carlo estimator satisfies*

$$
(51) \qquad \mathrm{MSE}(\hat{Z}) \leq C h^{\min\{2q, 2p-1\}} + \text{h.o.t.},
$$

*where $C$ is a positive constant independent of $h$.*

*Proof.* Thanks to (49) and (50), we just have to show

$$
(52) \qquad \mathrm{Var}(Y_k \mid Y_0 = y) \leq \hat{C} h^{2p-1},
$$

for a positive constant $\hat{C}$. From the definition of variance, of the operator $e^{h\mathcal{L}^h}$ and from the Markov property we have

$$
(53) \qquad \mathrm{Var}(Y_k \mid Y_0 = y) = e^{(k-1)h\mathcal{L}^h}\,\mathbb{E}(Y_1^2 \mid Y_0 = y) - \big(e^{(k-1)h\mathcal{L}^h}\,\mathbb{E}(Y_1 \mid Y_0 = y)\big)^2.
$$

Using the definition of the numerical method (4) we have

$$
(54) \qquad \begin{aligned} \mathbb{E}(Y_1^2 \mid Y_0 = y) &= \mathbb{E}(\Psi_{H_0}(y))^2 \\ &= \mathbb{E}(\Psi_{H_0}(y) - \Psi_h(y))^2 + \Psi_h(y)^2 \\ &\leq L_\Psi^2 C h^{2p} + \Psi_h(y)^2, \end{aligned}
$$

where we exploited Assumption 3.2 and 3.4.(ii). Moreover, it is trivial that

$$(55) \qquad \mathbb{E}(Y_1 \mid Y_0 = y) = \Psi_h(y).$$

Hence, we can bound the variance at the $N$-th step as

$$(56) \qquad \begin{aligned} \mathrm{Var}(Y_k \mid Y_0 = y) &\leq L_\Psi^2 C h^{2p} + \mathcal{O}(h^{2p+1}) + e^{(k-1)h\mathcal{L}^h} \Psi_h(y)^2 - \left(e^{(k-1)h\mathcal{L}^h} \Psi_h(y)\right)^2 \\ &= L_\Psi^2 C h^{2p} + \mathcal{O}(h^{2p+1}) + \mathrm{Var}(\Psi_h(Y_k) \mid Y_0 = y). \end{aligned}$$

Thanks to Assumption 3.4.(iii) we now have

$$(57) \qquad \mathrm{Var}(Y_k \mid Y_0 = y) \leq L_\Psi^2 C h^{2p} + \mathcal{O}(h^{2p+1}) + (1 + \hat{C}h)^2 \, \mathrm{Var}(Y_k \mid Y_0 = y),$$

for some constant $\hat{C} > 0$ independent of $h$. Hence,

$$(58) \qquad \mathrm{Var}(Y_k \mid Y_0 = y) \leq \left(L_\Psi^2 C h^{2p} + \mathcal{O}(h^{2p+1})\right) \sum_{i=0}^{k-1}(1 + \hat{C}h)^{2i}.$$

We now remark that the sum is bounded as

$$(59) \qquad \sum_{i=0}^{k-1}(1 + \hat{C}h)^{2i} \leq k e^{2\hat{C}hk} \leq T e^{2\hat{C}T} h^{-1},$$

where $T$ is the final time of (1). The desired result is then proved setting $C = T e^{2\hat{C}T}$. □

*Remark* 5.2. This result is of critical importance as it implies that the Monte Carlo estimators drawn from (4) converge in the mean square sense independently of the number of samples $M$ in (47). Hence, we can set $M = \mathcal{O}(1)$ when computing Monte Carlo averages without losing accuracy in the result if $h$ is small enough. In fact, introducing as a measure of error the square root of the MSE, and imposing a fixed tolerance $\varepsilon$, i.e.,

$$(60) \qquad \mathrm{MSE}(\hat{Z})^{1/2} = \mathcal{O}(\varepsilon),$$

we have thanks to Theorem 5.1 that the time step has to satisfy

$$(61) \qquad h = \mathcal{O}\left(\varepsilon^{1/\min\{q, p-1/2\}}\right),$$

regardless of the number of trajectories $M$. Therefore, the computational cost to attain such a tolerance $\varepsilon$ is given by

$$(62) \qquad \mathrm{cost} = \mathcal{O}(Mh^{-1}) = \mathcal{O}(\varepsilon^{\min\{q, p-1/2\}}).$$

**6. Numerical experiments.** We show numerical experiments verifying the theoretical results presented above.

**6.1. Strong order of convergence.** We now verify the weak order of convergence predicted in Theorem 4.2. We consider the FitzHug-Nagumo equation, which is defined as

$$(63) \qquad \begin{aligned} y_1' &= c\left(y_1 - \frac{y_1^3}{3} + y_2\right), & y_1(0) &= -1, \\ y_2' &= -\frac{1}{c}(y_1 - a + b y_2), & y_2(0) &= 1, \end{aligned}$$

where $a, b, c$ are real parameters with values $a = 0.2$, $b = 0.2$, $c = 3$. We integrate the equation from time $t_0 = 0$ to final time $T = 1$. The reference solution is generated with an high order method on a fine time scale. We consider as deterministic solvers the explicit trapezoidal rule and the classic fourth order Runge-Kutta method, which verify Assumption 3.4 with $q = 2$ and $q = 4$ respectively. Moreover, we consider random time steps as in Example 3.3, where we vary $p$ in order to verify the order of convergence predicted in Theorem 3.6. We vary the mean time step $h$ taken by the random time steps $H_n$ in the range $h_i = 0.01^i$, with $i = 0, 1, \ldots, 4$. Then, we simulate $10^4$ realizations of the numerical solution $Y_{N_i}$, with $N_i = T/h_i$ for $i = 0, 1, \ldots, 4$, and compute the approximate strong order of convergence for each value of $h$ with a Monte Carlo mean. Results (Table 1) show that the orders predicted theoretically by Proposition 3.6 are confirmed numerically.

| Method | ET | | | | | RK4 | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| $q$ | | | 2 | | | | | 4 | | |
| $p$ | 1 | 1.5 | 2 | 2.5 | 3 | 3 | 3.5 | 4 | 4.5 | 5 |
| $\min\{q, p-1/2\}$ | 0.5 | 1 | 1.5 | 2 | 2 | 2.5 | 3 | 3.5 | 4 | 4 |
| strong order | 0.52 | 1.01 | 1.52 | 2.02 | 2.01 | 2.50 | 2.99 | 3.55 | 3.99 | 3.98 |

**Table 1:** Strong order of convergence for the random time-stepping explicit trapezoidal (ET) and fourth-order Runge-Kutta (RK4) as a function of the value of $p$ of Assumption 3.2.

| Method | ET | | | RK4 | | | | |
|---|---|---|---|---|---|---|---|---|
| $q$ | | 2 | | | | 4 | | |
| $p$ | 1 | 1.5 | 2 | 1 | 1.5 | 2 | 3 | 4 |
| $\min\{q, 2p-1\}$ | 1 | 2 | 2 | 1 | 2 | 3 | 4 | 4 |
| weak order | 0.98 | 2.06 | 2.12 | 0.90 | 1.96 | 3.01 | 3.97 | 4.08 |

**Table 2:** Weak order of convergence for the random time-stepping explicit trapezoidal (ET) and fourth-order Runge-Kutta (RK4) as a function of the value of $p$ of Assumption 3.2.

| Method | ET | | RK4 | | | |
|---|---|---|---|---|---|---|
| $q$ | | 2 | | | 4 | |
| $p$ | 2 | 3 | 2 | 3 | 4 | 5 |
| $\min\{2q, 2p-1\}$ | 3 | 4 | 3 | 5 | 7 | 8 |
| MSE order | 3.01 | 4.05 | 3.04 | 5.02 | 7.08 | 8.06 |

**Table 3:** Convergence of the MSE of the Monte Carlo estimator for the random time-stepping explicit trapezoidal (ET) and fourth-order Runge-Kutta (RK4) with respect to $p$ of Assumption 3.2.

**6.2. Weak order of convergence.** We now verify the weak order of convergence predicted in Theorem 4.2. For this experiment we consider the ODE (63) as well, with the same time scale and parameters as above. The reference solution at final time is generated in this case as well with an high-order method on a fine time scale. The deterministic integrators we choose in this experiment are the explicit trapezoidal rule and the classic fourth-order Runge-Kutta method. The mean time step varies in the range $h_i = 0.1 \cdot 2^{-i}$ with $i = 0, 1, \ldots, 5$, and we vary the value of $p$ in Assumption 3.2 in order to verify the theoretical result of Theorem 4.2. The function $\Phi \colon \mathbb{R}^d \to \mathbb{R}$ of the solution we consider is defined as $\varphi(x) = x^T x$. Finally, we consider $10^6$ trajectories of the numerical solution in order to approximate the expectation with a Monte Carlo sum. Results (Table 2) show that the order of convergence predicted theoretically is confirmed by numerical experiments.

**6.3. Monte Carlo estimator.** In the same spirit of the previous numerical experiments, we now verify numerically the validity of Theorem 5.1. We consider the ODE (63), with final time $T = 10$ and the same parameters as above. In this case as well, we consider the explicit trapezoidal rule and the fourth-order explicit Runge-Kutta method with random time steps having mean $h_i = 0.1 \cdot 2^{-i}$ with $i = 0, 1, \ldots, 5$. We thus vary the value of $p$ of Assumption 3.2 and compute the mean order of convergence over 300 repetitions of the experiment. For each repetition, we consider only one trajectory to approximate the Monte Carlo estimator. We then compute the error with respect to a

reference solution given by a high-order method with a small time step. Results (Table 3) show that the convergence order given in Theorem 5.1 is respected in practice.

## REFERENCES

[1] 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).
[2] E. HAIRER, S. P. NØRSETT, AND G. WANNER, *Solving Ordinary Differential Equations I. Nonstiff Problems*, vol. 8, Springer Verlag Series in Comput. Math., Berlin, 1993.