ESAIM: M2AN 59 (2025) 3041–3067
https://doi.org/10.1051/m2an/2025087

ESAIM: Mathematical Modelling and Numerical Analysis
www.esaim-m2an.org

STRONG ORDER-ONE CONVERGENCE OF THE EULER METHOD FOR RANDOM ORDINARY DIFFERENTIAL EQUATIONS DRIVEN BY SEMI-MARTINGALE NOISES

Peter E. Kloeden[1] and Ricardo M. S. Rosa[2,*]

Abstract. It is well known that the Euler method for a random ordinary differential equation $\mathrm{d}X_t/\mathrm{d}t = f(t, X_t, Y_t)$ driven by a stochastic process $\{Y_t\}_{t\in I}$, on a time interval I, with θ-Hölder sample paths is of strong order θ with respect to the time step, provided $f = f(t, x, y)$ is sufficiently regular and with suitable bounds. This order is known to increase to 1 in some special cases. Here, it is proved that, in many more typical cases, further structures on the noise can be exploited so that the strong convergence is of order 1. In fact, we prove so for any semi-martingale noise. This includes Itô diffusion processes, point-process noises, transport-type processes with sample paths of bounded variation, and time-changed Brownian motion. The result follows from estimating the global error as an iterated integral over both large and small mesh scales, and by switching the order of integration to move the critical regularity to the large scale. The work is complemented with numerical simulations showing the optimality of the strong order 1 convergence in those cases, and with an example with fractional Brownian motion noise with Hurst parameter $0 < H < 1/2$, which is not a semi-martingale and for which the order of convergence is $H + 1/2$, hence lower than the attained order 1 in the semi-martingale case, but still higher than the order H of convergence expected from previous works.

Mathematics Subject Classification. 60H35, 65C20, 60H25, 35R60.

Received May 27, 2025. Accepted October 8, 2025.

1. Introduction

A *random ordinary differential equation (RODE or Random ODE)* is a differential equation of the form

$$
\begin{cases} \dfrac{\mathrm{d}X_t}{\mathrm{d}t} = f(t, X_t, Y_t), \\ X_t|_{t=0} = X_0, \end{cases} \tag{1.1}
$$

for t on a time interval which we take to be $I = [0, T]$, with $T > 0$, where the initial condition X_0 is a given random variable, and the noise $\{Y_t\}_{t\in I}$ is a given stochastic process, both defined on a probability space

Keywords and phrases. Random ordinary differential equations, Euler method, strong convergence, Itô process, finite-variation process, semi-martingale, fractional Brownian motion.

[1] Mathematics Department, University of Tübingen, Tübingen, Germany.

[2] Instituto de Matemática, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil.

[*]Corresponding author: rrosa@im.ufrj.br

$(\Omega, \mathcal{F}, \mathbb{P})$. The function $f = f(t, x, y)$ is such that the differential equation in (1.1) can be either a scalar or a system of equations and the noise can also be either a scalar or a vector-valued process.

Random ODEs appear in applications where the parameters either vary randomly or are not fully known. As such, they have been playing an important role in modeling practical problems, including uncertainty quantification (see *e.g.* [3,5,10,13,26,28,31,32]). They are less known than the more popular Itô-diffusion stochastic differential equations (SDE), but, depending on the situation, they may actually lead to more realistic models.

As in many other models, numerical approximations are of crucial importance. One popular method for approximating Random ODEs is the Euler method. It consists in approximating the solution on a uniform time mesh $t_j = t_j^N = j\Delta t_N$, $j = 0, \ldots, N$, with fixed time step $\Delta t_N = T/N$, for a given $N \in \mathbb{N}$, *via*

$$\begin{cases} X_{t_j}^N = X_{t_{j-1}}^N + \Delta t_N f\Big(t_{j-1}, X_{t_{j-1}}^N, Y_{t_{j-1}}\Big), & j = 1, \ldots, N, \\ X_0^N = X_0. \end{cases} \tag{1.2}$$

An approximation $\{X_{t_j}^N\}_{j=0,1,\ldots,N}$ is said to converge to $\{X_t\}_{t\in I}$ with *strong order* $\theta > 0$ when there exists a constant $C \geq 0$ such that

$$\max_{j=0,\ldots,N} \mathbb{E}\left[\left\|X_{t_j} - X_{t_j}^N\right\|\right] \leq C\Delta t_N^\theta, \qquad \forall N \in \mathbb{N}, \tag{1.3}$$

where $\mathbb{E}[\cdot]$ indicates the expectation of a random variable on the underlying probability space, and $\|\cdot\|$ is a suitable norm in the appropriate phase space.

There are other important notions of convergence, such as pathwise convergence, weak convergence, mean-square convergence, and p-th mean convergence (see *e.g.* [4,12,13,16,18,21]), but here we focus only on the strong convergence.

Under certain regularity conditions on f and when the noise $\{Y_t\}_{t\in I}$ has θ-Hölder continuous sample paths, the result in Theorem 3 of [35] proves the order of convergence of the Euler scheme, in the mean square sense, to be of order θ with respect to the time step. This implies the strong convergence (1.3) with the same order θ. In the case of fractional Brownian motion noise with Hurst parameter H, Theorem 2 of [35] proves the mean square convergence to be of order H, for any $0 < H \leq 1$.

In some special cases, the order of convergence of the Euler method is known to be higher than the Hölder exponent of the noise. For example, for Itô diffusion noises, the RODE can be seen as part of a system of SDEs, with the Euler method for the RODE component coinciding with the Milstein method for the SDE, which is known to be of strong order 1 (see [35], Sect. 5, in particular [35], Sect. 5.2, Example 12 and Rem. 13).

In many more classical examples, though, the actual order of convergence seems to be higher than what the theory predicts. Our aim here is to show that, indeed, it is possible to exploit further properties of the noise that yield a strong order 1 convergence for essentially any noise that is a semi-martingale (see Thm. 7.1 for the precise statement). This includes noises such as point processes, transport processes, Itô processes, and time-changed Brownian motions, some of which are not even continuous and for which the previous theories did not apply.

This strong order 1 convergence is not only a major improvement from previous results, but it is also remarkable that it follows with a unified proof for all semi-martingale noises. It also has potential extensions to pathwise convergence for RODEs in the case of non-globally Lipschitz functions, at least when driven by finite-variation process. The result should also extend to random partial differential equations. It is not clear, though, whether this technique could also yield improvements in the order of convergence for higher order methods.

We numerically verify, in a series of examples, that the strong order 1 convergence is in fact optimal. Indeed, this work was motived by the fact that the observed convergence in these and many other examples was higher than what the theory predicted. Below is a summary of these examples:

(i) Two types of non-homogeneous linear equations of the form

$$\frac{\mathrm{d}X_t}{\mathrm{d}t} = -\|Y_t\|^2 X_t + Y_t, \tag{1.4}$$

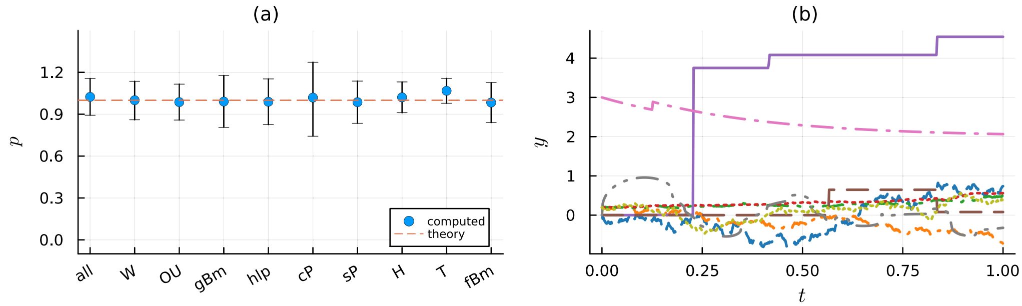

FIGURE 1. (a) Estimated order of convergence of the strong error of the Euler method for the linear system (1.4) as it depends on the noise, starting with all the noises combined (all) and continuing with the scalar equation, each with a different type of noise process, namely Wiener (W), Ornstein–Uhlenbeck (OU), geometric Brownian motion (gBm), a homogeneous linear Itô diffusion (hlp), compound Poisson process (cP), step point process (sP), Hawkes process (H), transport process (T), and also a fractional Brownian motion with Hurst parameter $1/2 < H < 1$. (fBm); (b) Sample paths of all the noises used in the linear system (1.4).

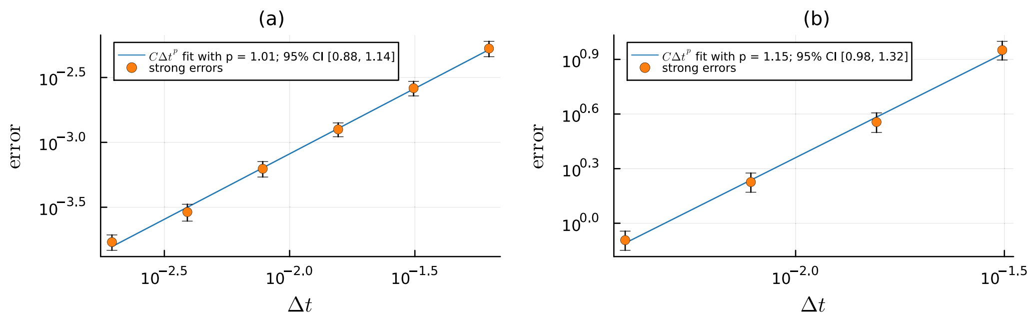

FIGURE 2. Estimated order of convergence of the strong error of the Euler method for (a) the population dynamics model (1.5) and (b) the seismic model (1.7).

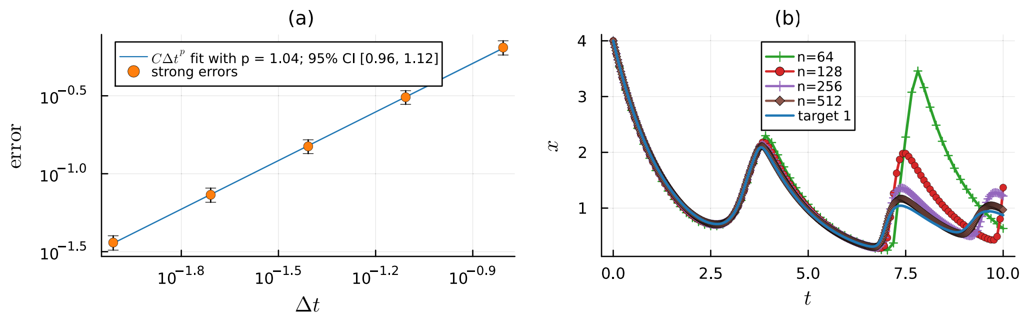

FIGURE 3. (a) Estimated order of convergence of the strong error of the Euler method for the toggle-switch model (1.6); (b) Euler approximations of the component $\{X_t\}_{t \in I}$ of a sample path.

where $\{Y_t\}_{t\in I}$ is either (a) a univariate process with one of many types of semi-martingale processes or (b) a vector-valued process combining all these noises as independent components.

(ii) A logistic model of population dynamics with random coefficients, loosely inspired by Section 15.2 of [13], and where the specific growth is the sine of a geometric Brownian motion process $\{G_t\}_{t\in I}$ and with an extra point-process random term $\{H_t\}_{t\in I}$ representing harvest:

$$\frac{\mathrm{d}X_t}{\mathrm{d}t} = \gamma(1+\varepsilon\sin(G_t))X_t\left(1-\frac{X_t}{r}\right) - \alpha H_t\frac{X_t}{r+X_t}. \tag{1.5}$$

(iii) A toggle-switch model of gene expression (similar to [3], Sect. 7.8, originated from [34], see also [32]) driven by a combination of a compound Poisson point process $\{A_t\}_{t\in I}$ and an Itô diffusion process $\{B_t\}_{t\in I}$,

$$\begin{cases} \dfrac{\mathrm{d}X_t}{\mathrm{d}t} = \left(A_t + \dfrac{X_t^4}{a^4+X_t^4}\right)\left(\dfrac{b^4}{b^4+Y_t^4}\right) - \mu X_t, \\ \dfrac{\mathrm{d}Y_t}{\mathrm{d}t} = \left(B_t + \dfrac{Y_t^4}{c^4+Y_t^4}\right)\left(\dfrac{d^4}{d^4+X_t^4}\right) - \nu Y_t. \end{cases} \tag{1.6}$$

(iv) A mechanical structure seismic model

$$\ddot{X}_t + 2\zeta_0\omega_0\dot{X}_t + \omega_0^2 X_t = -\ddot{M}_t, \tag{1.7}$$

driven by a random disturbance in the form of a transport process $M_t = \sum_{i=1}^k \gamma_i(t-\tau_i)_+^2 e^{-\delta_i(t-\tau_i)}\cos(\omega_i(t-\tau_i))$, simulating seismic ground-motion excitations as inspired by the Bogdanoff–Goldberg–Bernard model in [5].

(v) An actuarial risk model

$$\frac{\mathrm{d}X_t}{\mathrm{d}t} = R_t X_t + R_t(C_t + O_t) + \nu O_t + \gamma \tag{1.8}$$

for the surplus $U_t = X_t + C_t + O_t$ of an insurance company, inspired by Gerber and Shiu [11] and Brigo and Mercurio [6], where $C_t = \sum_{i=1}^{N_t} C_i$ defines a compound Poisson process modeling the insurance claims, $\{R_t\}_{t\in I}$ is a geometric Brownian motion process modeling the interest rate, and $\{O_t\}_{t\in I}$ is an Ornstein–Uhlenbeck process, used to transform an additive SDE into a RODE.

(vi) A Fisher-KPP partial differential equation

$$\frac{\partial u}{\partial t} = \mu\frac{\partial^2 u}{\partial x^2} + \lambda u\left(1-\frac{u}{u_m}\right), \quad (t,x)\in(0,\infty)\times(0,1), \tag{1.9}$$

endowed with random boundary conditions representing migration,

$$\frac{\partial u}{\partial x}(t,0) = -Y_t, \quad \frac{\partial u}{\partial x}(t,1) = 0, \tag{1.10}$$

as inspired by the works [10, 31] (see also [9, 24]), and where the noise $\{Y_t\}_{t\in I}$ is an Ornstein–Uhlenbeck process modulated by a decaying self-exciting Hawkes point process.

(vii) The linear equation

$$\frac{\mathrm{d}X_t}{\mathrm{d}t} = -X_t + B_t^H, \tag{1.11}$$

where $\{B_t^H\}_{t\in I}$ is a fractional Brownian motion (fBm) with Hurst parameter $0 < H < 1$, which is a Wiener process when $H = 1/2$ and which is not a semi-martingale for $H \neq 1/2$.

In all but the last example, the numerically obtained convergence was found to be of strong order approximately 1, while previous results would either not be applicable or yield a much lower order of convergence (see Figs. 1–5).

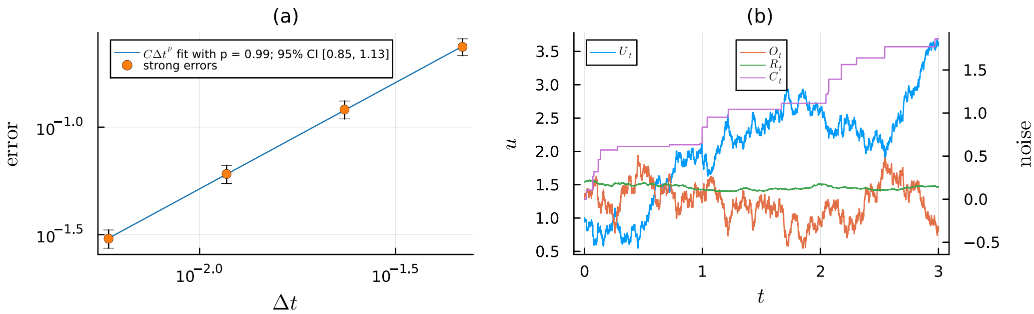

FIGURE 4. (a) Estimated order of convergence of the strong error of the Euler method for the insurance risk model (1.8); (b) Sample paths of the surplus $\{U_t\}_{t\in I}$, the Ornstein–Uhlenbeck process $\{O_t\}_{t\in I}$, the interest rate $\{I_t\}_{t\in I}$, and the insurance claims $\{C_t\}_{t\in I}$.

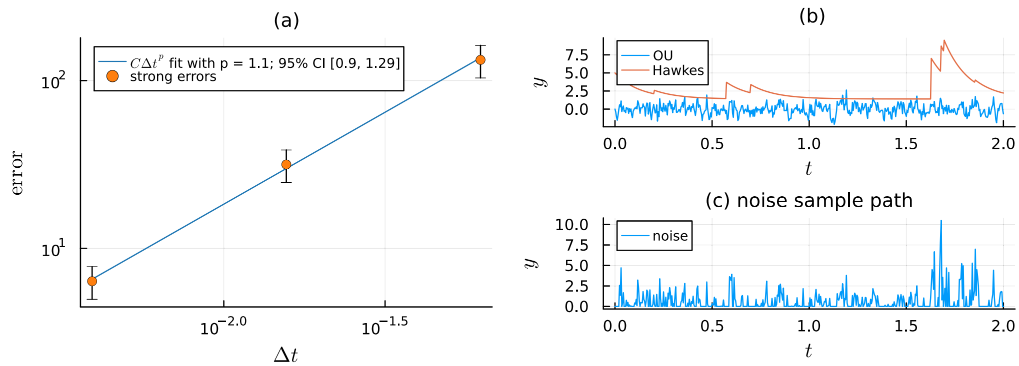

FIGURE 5. (a) Estimated order of convergence of the strong error of the Euler method for the Fisher-KPP model (1.9) and (1.10); Sample paths of (b) the Ornstein–Uhlenbeck and Hawkes processes, and of (c) the noise obtained as the product of these processes.

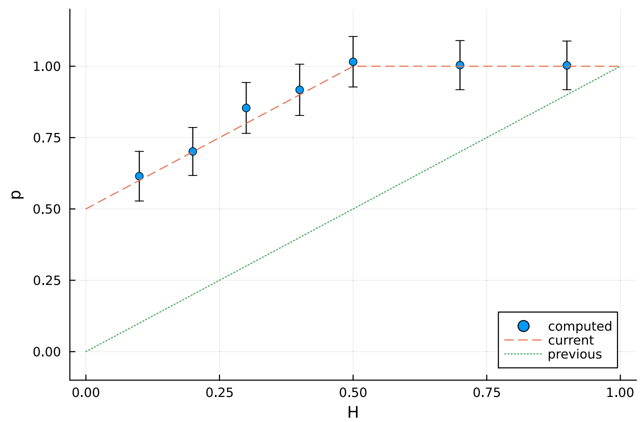

FIGURE 6. Estimated order p of strong convergence of the Euler method for (1.11) at different values of the Hurst parameter H (scattered plot) along with the theoretical value $p = \min\{H + 1/2, 1\}$ proved here (dashed line), and the previously known theoretical value $p = H$ (dotted).

In the last example, with the fractional Brownian motion, the order was found to be $\min\{H+1/2, 1\}$, as illustraded in Figure 6, well above the previously expected order of convergence H.

Our aim is not only to report on these discrepancies but also to prove and explain why we have such better convergence rates.

Details about the numerical setup for estimating the order of strong convergence are given in Section 8. Details concerning the first example (1.4) are given in Section 8.2. A more detailed description of the numerical setup and the other examples can be found in the open-access github repository [22].

Concerning the proof of strong order 1 convergence, the starting point is not to bound the local error and, instead, work with an explicit formula for the global error (see Lem. 5.1), similarly to the way it is done for approximations of stochastic differential equations. In this way, we obtain the bound (see (6.4))

$$\begin{aligned}\left\|X_{t_j} - X_{t_j}^N\right\| \leq \Bigg(&\left\|X_0 - X_0^N\right\| + c_L \int_0^{t_j} \left\|X_s - X_{\tau^N(s)}\right\| \mathrm{d}s \\ &\left\|\int_0^{t_j} \left(f\left(s, X_{\tau^N(s)}^N, Y_s\right) - f\left(\tau^N(s), X_{\tau^N(s)}^N, Y_{\tau^N(s)}\right)\right) \mathrm{d}s\right\|\Bigg) e^{c_L t_j},\end{aligned} \tag{1.12}$$

for $j = 1, \ldots, N$, where $\tau^N(t) = \max_{t_j \leq t} t_j$ is the running time of the Euler method, *i.e.* a piecewise constant function jumping along the mesh points t_j (see (5.2)). The constant c_L is the Lipschitz constant of $f = f(t, x, y)$ with respect to x, uniformly in (t, y), a typical assumption for this kind of strong convergence result.

The only problematic, noise-sensitive term in (1.12) is the last one. The classical analysis is to use an assumed θ-Hölder regularity of the noise sample paths and bound the local strong error by $C\Delta t_N^\theta$. Instead, we look at the global noise error and show that the steps of the process given by $F_t = f(t, X_{\tau^N(t)}^N, Y_t)$ can be controlled in a suitable global way. In order to give the main idea, let us consider a scalar equation with a scalar noise and assume that the sample paths of $\{F_t\}_{t\in I}$ satisfy

$$F_s - F_\tau = \int_\tau^s \mathrm{d}F_\xi.$$

For a semi-martingale noise, the integral above can be a combination of a Lebesgue–Stieltjes integral, a stochastic integral, and an integral with respect to a jump measure, but we leave it like that just for the sake of explaining the main idea. With that, we bound the global error term using the Fubini Theorem, taking into consideration that, despite the inner integral ranging from ξ to the "future" time $\tau^N(\xi) + \Delta t_N$, the integrand is still non-anticipative, hence

$$\begin{aligned}\int_0^{t_j} \left(f\left(s, X_{\tau^N(s)}^N, Y_s\right) - f\left(\tau^N(s), X_{\tau^N(s)}^N, Y_{\tau^N(s)}\right)\right) \mathrm{d}s &= \int_0^{t_j} \int_{\tau^N(s)}^s \mathrm{d}F_\xi \,\mathrm{d}s \\ &= \int_0^{t_j} \int_\xi^{\tau^N(\xi)+\Delta t_N} \mathrm{d}s \,\mathrm{d}F_\xi = \int_0^{t_j} \left(\tau^N(\xi) + \Delta t_N - \xi\right) \mathrm{d}F_\xi.\end{aligned}$$

Then, since $0 \leq \tau^N(\xi) + \Delta t_N - \xi \leq \Delta t_N$, we find that

$$\mathbb{E}\left[\left\|\int_0^{t_j} \left(f\left(s, X_{\tau^N(s)}^N, Y_s\right) - f\left(\tau^N(s), X_{\tau^N(s)}^N, Y_{\tau^N(s)}\right)\right) \mathrm{d}s\right\|\right] \leq \Delta t_N \mathbb{E}\left[\left\|\int_0^{t_j} \mathrm{d}F_\xi\right\|\right],$$

which yields the strong order 1 convergence. The main work is to actually show that $\int_0^{t_j} \mathrm{d}F_\xi$ is strongly bounded.

In the fractional Brownian motion case, which is not a semi-martingale, the argument above does not apply. Instead, we essentially have, in the case $0 < H < 1/2$,

$$F_s - F_\tau = \int_\tau^s g(\xi)(s-\tau)^{H-1/2} \,\mathrm{d}W_\xi + \text{higher order term},$$

for a suitable function $g = g(\xi)$. Notice the kernel $(s-\tau)^{H-1/2}$ is unbounded when $0 < H < 1/2$. In this case, disregarding the higher order term,

$$\int_0^{t_j} \left(f\left(s, X^N_{\tau^N(s)}, Y_s\right) - f\left(\tau^N(s), X^N_{\tau^N(s)}, Y_{\tau^N(s)}\right)\right) \mathrm{d}s = \int_0^{t_j} \int_{\tau^N(s)}^s g(\xi)(s-\tau^N(s))^{H-1/2} \,\mathrm{d}W_\xi \,\mathrm{d}s$$
$$= (\Delta t_N)^{H+1/2} \int_0^{t_j} g(\xi)\,\mathrm{d}W_\xi,$$

which yields a strong convergence of order $H+1/2$, better than the previously known rate of H. For $1/2 \leq H < 1$, there are other terms of order 1, so that we obtain order $\min\{H+1/2, 1\}$, for $0 < H < 1$.

2. Pathwise solutions and main assumptions

For the notion and main results on pathwise solutions for RODEs, we refer the reader to Section 2.1 of [13] and Section 3.3 of [28]. We start with a fundamental set of conditions that imply the existence and uniqueness of pathwise solutions of the RODE (1.1) in the sense of Carathéodory. In what follows, $\|\cdot\|$ denotes de Euclidian norm for vectors and the corresponding operator norm for matrices.

Standing Hypothesis 2.1. *Let $d, k \in \mathbb{N}$ and $I = [0, T]$, where $T > 0$. We consider a function $f = f(t, x, y)$ defined on $I \times \mathbb{R}^d \times \mathbb{R}^k$ and with values in $\mathbb{R}^d$, a initial condition X_0, and an $\mathbb{R}^k$-valued stochastic process $\{Y_t\}_{t\in I}$. We make the following standing assumptions:*

(i) *The function $f = f(t, x, y)$ is continuous in $(t, x, y) \in I \times \mathbb{R}^d \times \mathbb{R}^k$.*
(ii) *The mapping $x \mapsto f(t, x, y)$ is globally Lipschitz continuous in x, uniformly in $t \in I$ and locally in $y \in \mathbb{R}^k$, i.e.*
$$\|f(t, x_1, y) - f(t, x_2, y)\| \leq L(\|y\|)\|x_1 - x_2\|, \tag{2.1}$$
for every $t \in I$, $y \in \mathbb{R}^k$, and $x_1, x_2 \in \mathbb{R}^d$, where $L = L(\eta)$ is a monotonic nondecreasing real-valued function defined for $\eta \geq 0$.
(iii) *The stochastic vector process $\{Y_t\}_{t\in I}$ is an adapted semi-martingale on a right-continuous, filtered, and complete probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$, i.e. where $\mathcal{F}_0$ contains all the null sets of probability distribution $\mathbb{P}$, and with a right-continuous filtration $\mathcal{F}_t = \cap_{s>t}\mathcal{F}_s$ in the σ-algebra $\mathcal{F}$.*
(iv) *The initial condition is a random variable X_0 on $(\Omega, \mathcal{F}, \mathbb{P})$.*

Under these assumptions, for almost every sample value ω in Ω, the integral equation

$$X_t(\omega) = X_0(\omega) + \int_0^t f(s, X_s(\omega), Y_s(\omega))\,\mathrm{d}s \tag{2.2}$$

has a unique solution $t \mapsto X_t(\omega)$, for the realizations $X_0 = X_0(\omega)$ of the initial condition and $t \mapsto Y_t(\omega)$ of the noise process (see [7], Thm. 1.1). Each pathwise solution of (2.2) is absolutely continuous. Moreover, the mapping $(t, \omega) \mapsto X_t(\omega)$ is (jointly) measurable (see [13], Sect. 2.1.2 and also [2], Lem. 4.51) and, hence, gives rise to a well-defined stochastic process $\{X_t\}_{t\in I}$. Equation (2.2) is the integral form of equation (1.1), so that $\{X_t\}_{t\in I}$ is a solution of (1.1) in the pathwise sense.

Remark 2.1. Since $\{Y_t\}_{t\in I}$ is a semi-martingale, the sample paths $t \mapsto Y_t(\omega)$ are almost surely continuous except for a set of measure zero (see Sect. 7). Then, since $f = f(t, x, y)$ is continuous, the integrand $s \mapsto f(s, X_s(\omega), Y_s(\omega))$ is (Riemann) integrable, and each pathwise solution is absolutely continuous.

Remark 2.2. Since the paths of a semi-martingale are càdlàg (see the definition in Sect. 4), the noise process is in fact progressively measurable with respect to the filtration $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$. In turn, the solution process $\{X_t\}_{t\in I}$ is adapted and hence also progressively measurable to the filtration $(\Omega, \mathcal{F}, (\tilde{\mathcal{F}}_t)_{t\geq 0}, \mathbb{P})$, where each $\tilde{\mathcal{F}}_t$ is the smallest σ-algebra generated by the union of $\mathcal{F}_t$ and the smallest σ-algebra that makes the initial condition X_0 measurable.

Adding and subtracting $f(s,0,Y_s)$ and using assumption (ii) of the Standing Hypothesis 2.1 yield the pathwise bound

$$\|X_t\| \leq \|X_0\| + \int_0^t (\|f(s,0,Y_s)\| + L(\|Y_s\|)\|X_s\|)\,\mathrm{d}s,$$

almost surely. Using Gronwall's inequality, we obtain the following result.

Lemma 2.1. *Under the Standing Hypothesis 2.1, almost every solution $\{X_t\}_{t\in I}$ of (2.2) satisfies*

$$\|X_t\| \leq \left(\|X_0\| + \int_0^t \|f(s,0,Y_s)\|\,\mathrm{d}s\right)e^{\int_0^t L(\|Y_s\|)\,\mathrm{d}s}, \quad \forall t \in I. \tag{2.3}$$

This is a useful pathwise bound on its own and is particularly useful for proving pathwise convergence. For the strong convergence, however, we need further assumptions. We assume that X_0 has finite expectation, that $\|f(t,0,Y_t)\|$ has integrable expectation, and that L is bounded, *i.e.*

$$\mathbb{E}[\|X_0\|] < \infty, \tag{2.4}$$

$$\int_0^T \mathbb{E}[\|f(t,0,Y_t)\|]\,\mathrm{d}t < \infty, \tag{2.5}$$

and there exists a constant $c_L \geq 0$ such that

$$\sup_{\eta\geq 0} L(\eta) \leq c_L < \infty, \tag{2.6}$$

so that

$$\|f(t,x_1,y) - f(t,x_2,y)\| \leq c_L\|x_1 - x_2\|, \tag{2.7}$$

for every $t \in I$, $y \in \mathbb{R}^k$, and $x_1, x_2 \in \mathbb{R}^d$. With such assumptions, it is straightforward, from taking the expectation of (2.3), to obtain the following result.

Lemma 2.2. *Under the Standing Hypothesis 2.1, assume that (2.4)–(2.6) hold. Then*

$$\mathbb{E}[\|X_t\|] \leq \left(\mathbb{E}[\|X_0\|] + \int_0^t \mathbb{E}[\|f(s,0,Y_s)\|]\,\mathrm{d}s\right)e^{c_L t}, \quad t \in I. \tag{2.8}$$

Remark 2.3. In a number of cases, depending on the structure of the equation, we might not need all the previous assumptions, especially the restrictive assumption (2.6), corresponding to a global Lipschitz condition in x, uniformly in t and y. Indeed, when some bounded, positively invariant region exists and is of interest, we may truncate the nonlinear term to achieve the desired global conditions. Within the invariant region, the equation with the truncated term coincides with the original equation, and so do their Euler approximations, for a sufficiently refined mesh. In other situations, higher order moments of the noise and/or of the initial condition may compensate the lack of some of these assumptions and still yield bounds for the expectation of the norm of the solution. In the examples described in the Introduction, the linear models (iv) and (vii) are the only ones that satisfy the global condition (2.6). All the other examples fall within the framework of having a bounded, positively invariant region. An example in which neither applies but bounds on higher order moments compensate is, for example, a linear equation with an unbounded multiplicative noise, which can be found in the github repository [22]. Finally, it is worth mentioning that the global condition (2.6) is the classical assumption used for the strong convergence of approximatinos of stochastic differential equations [4, 12, 13, 16, 18, 21].

3. Bound on the pathwise approximations

We also need bounds for the Euler approximation (1.2) analogous to (2.3) and (2.8). Notice that

$$\begin{aligned}\left\|X_{t_j}^N\right\| &\leq \left\|X_{t_{j-1}}^N\right\| + \Delta t_N \left\|f\left(t_{j-1}, X_{t_{j-1}}^N, Y_{t_{j-1}}\right)\right\| \\ &\leq \left\|X_{t_{j-1}}^N\right\| + \Delta t_N \left(\left\|f\left(t_{j-1}, 0, Y_{t_{j-1}}\right)\right\| + L(\|Y_{t_{j-1}}\|)\|X_{t_{j-1}}\|\right) \\ &\leq \left(1 + L(\|Y_{t_{j-1}}\|)\Delta t_N\right)\left\|X_{t_{j-1}}^N\right\| + \Delta t_N \left\|f\left(t_{j-1}, 0, Y_{t_{j-1}}\right)\right\|.\end{aligned}$$

We bound $(1+a) \leq e^a$, so that

$$\left\|X_{t_j}^N\right\| \leq e^{L(\|Y_{t_{j-1}}\|)\Delta t_N}\left\|X_{t_{j-1}}^N\right\| + \Delta t_N \left\|f\left(t_{j-1}, 0, Y_{t_{j-1}}\right)\right\|.$$

Iterating this relation we obtain the following bound.

$$\left\|X_{t_j}^N\right\| \leq \left(\left\|X_0^N\right\| + \Delta t_N \sum_{i=0}^{j-1}\|f(t_i, 0, Y_{t_i})\|\right) e^{\Delta t_N \sum_{i=0}^{j-1} L(\|Y_{t_i}\|)}, \tag{3.1}$$

for $j = 0, \ldots, N$. The right hand side above depends on the time mesh. A straighforward bound independent of the time mesh is obtained from (3.1) and the fact that semi-martingales are pathwise locally bounded, almost surely:

$$\left\|X_{t_j}^N\right\| \leq \left(\left\|X_0^N\right\| + t_j \sup_{0\leq s\leq t_j}\|f(t_s, 0, Y_s)\|\right) e^{t_j \sup_{0\leq s\leq t_j} L(\|Y_s\|)}, \tag{3.2}$$

for $j = 0, \ldots, N$. With that, we obtain the following result, upon taking the expectation and using the assumed bounds.

Lemma 3.1. *Under the Standing Hypothesis 2.1, assume that (2.4)–(2.6) hold. Then, the Euler approximation (1.2) for a pathwise solution of the random ordinary differential equation (1.1) satisfies*

$$\mathbb{E}\left[\left\|X_{t_j}^N\right\|\right] \leq \left(\mathbb{E}\left[\left\|X_0^N\right\|\right] + \mathbb{E}\left[\sum_{i=0}^{j-1}\|f(t_i, 0, Y_{t_i})\|\Delta t_N\right]\right) e^{c_L t_j}, \tag{3.3}$$

for all $j = 0, \ldots, N$.

4. Semi-martingale noises

For the theory of semi-martingales, we refer the reader to [1,15,20,27,30]. Starting with the approach of [30], the main idea is that *semi-martingales* are stochastic processes $\{Y_t\}_{t\in I}$ which are càdlàg, adapted, and "good integrators", in the sense that the sums

$$I_Y(H) = H_0 Y_0 + \sum_{i=1}^n H_i\left(Y_{t_i} - Y_{t_{i-1}}\right)$$

are continuous functions from the space of *simple predictable* processes

$$H_t = H_0\chi_0(t) + \sum_{i=1}^n H_j \chi_{(T_{i-1}, T_i]},$$

to the space of a.s. finite-valued random variables, where $0 = T_0 < T_1 < \ldots < T_n$, are stopping times, $H_j \in \mathcal{F}_{T_j}$, and $|H_j| < \infty$ almost surely, with the former space endowed with the topology of uniform convergence in

$(t, \omega) \in [0, \infty) \times \Omega$, and the latter, with the topology of convergence in probability. We recall the definitions and some facts about (scalar or vector-valued) semi-martingale processes.

A function defined on a real interval is called *càdlàg* when it is right continuous and with left limits ("continue à droite, limite à gauche", in French). A *càdlàg process* is a stochastic process $\{Y_t\}_{t\in I}$ whose sample paths are almost surely càdlàg paths. Thus, $\{Y_t\}_{t\in I}$ is a càdlàg process when for almost every $\omega \in \Omega$ and for all $t \in I$, we have that $\lim_{s\to t^+} Y_s(\omega) = Y_t(\omega)$ (right continuous) and the limit $\lim_{s\to t^-} Y_s(\omega)$ exists (left limit).

For a càdlàg process such as semi-martingales, we define the left-limit process $Y_{t^-} = \lim_{t\to t^-} Y_s$ and the jump process $\Delta Y_t = Y_t - Y_{t^-}$, with the usual convention that $Y_{0^-} = 0$, and thus $\Delta Y_0 = Y_0$.

Almost surely, the sample paths of the jump process are zero except possibly at a countable number of points, denoted J_Y, *i.e.* for almost every $\omega \in \Omega$, there exists a countable set $J_Y(\omega) \subset I$ for which $\Delta Y_t(\omega)$ is zero on each connected component of $I \setminus J_Y(\omega)$. In fact, for any given $\varepsilon > 0$, the set of times for which the jump $\|\Delta Y_t\|$ is larger than or equal to ε is finite. As a consequence, each sample path is continuous except on the countable null set J_Y.

Another important process associated with a semi-martingale is the *quadratic variation* $\{[Y, Y]_t\}_{t\in I}$. It is defined either as a suitable limit of the summations of quadratic steps $(Y_{T_j^n} - Y_{T_{j-1}^n})^2$ over random partitions made of stopping times T_j^n or, equivalently, as a suitable integral ([30], Sect. II.6). In the vector-valued case, denoting $Y_t = (Y_t^{(l)})_{l=1,\ldots,k}$, with the vectorial components $Y_t^{(l)}$, the quadratic process becomes matrix-valued, *i.e.* $[Y, Y]_t = ([Y^{(l_1)}, Y^{(l_2)}]_t)_{l_1,l_2=1,\ldots,k}$, with the appearance of the quadratic variations and covariations. The equivalent integral definition takes, more precisely, the form of an "integration by parts" formula for two semi-martingales,

$$\left[Y^{(l_1)}, Y^{(l_2)}\right]_t = Y_t^{(l_1)} Y_t^{(l_2)} - \int_0^t Y_{s^-}^{(l_1)}\, \mathrm{d}Y_s^{(l_2)} - \int_0^t Y_{s^-}^{(l_2)}\, \mathrm{d}Y_s^{(l_1)}.$$

The quadratic variation is a càdlàg, increasing, and adapted process ([30], Thm. II.22). The *continuous part* $\{[Y, Y]_t^c\}_{t\in I}$ of the quadratic variation is defined by

$$[Y, Y]_t^c = [Y, Y]_t - \sum_{0\le s\le t} \Delta Y_s \Delta Y_s^{\mathrm{tr}},$$

which itself is a semi-martingale of finite-variation type, and where Y_t^{tr} is the transpose of Y_t. Notice that $[Y, Y]_0 = Y_0^2$, while, for the continuous part, $[Y, Y]_0^c = [Y, Y]_0 - \Delta Y_0 \Delta Y_0^{\mathrm{tr}} = Y_0 Y_0^{\mathrm{tr}} - Y_0 Y_0^{\mathrm{tr}} = 0$.

The Bichteler–Dellacherie Theorem ([30], Thm. III.47) yields a characterization of a semi-martingale as a càdlàg process $\{Y_t\}_{t\in I}$ decomposable as a sum of a finite-variation (FV) process and a local martingale process, which we can write in standard form as $Y_t = Y_0 + F_t + Z_t$, with an FV process $\{F_t\}_{t\in I}$ and a local martingale process $\{Z_t\}_{t\in I}$ with $F_0 = Z_0 = 0$.

A *Finite-Variation process* (*FV process*) is a càdlàg process $\{F_t\}_{t\in I}$ whose sample paths are almost surely of finite variation $V(\{F_t\}_{t\in I}; K)$ on compact subsets $K \subset I$, *i.e.*

$$V(\{F_t\}_{t\in I}; K) = \sup_{t_0<\ldots<t_n\in K, n\in\mathbb{N}} \sum_{j=1}^n \left\|F_{t_j} - F_{t_{j-1}}\right\| < \infty.$$

This total variation induces a positive measure $\|\mathrm{d}F_t\|$, which is the infinitesimal of this total variation, so that $\|\mathrm{d}F_t\|([a, b]) = V(\{F_t\}_{t\in I}; [a, b])$, for every $[a, b] \subset I$.

The class of FV processes includes the compound Poisson process, renewal-reward processes, inhomogeneous Poisson process, Hawkes self-exciting process, and many more. It also includes transport processes $Y_t = g(t, Y_0)$ where Y_0 is a random variable and $t \mapsto g(t, y_0)$ is a càdlàg function of finite variation, for every sample value $y_0 = Y_0(\omega)$.

A *local martingale* is an adapted, càdlàg process $\{Z_t\}_{t\in I}$ such that there exists a sequence of increasing stopping times T_n with $T_n \to \infty$ almost surely, for which $\{Z_{t\wedge T_n} 1_{T_n>0}\}_{t\in I}$ is a uniformly integrable martingale, for each $n \in \mathbb{N}$. Here, $t \wedge T = \min\{t, T\}$, the function $1_{T_n>0}$ is the indicator function of the set $\{\omega;\ T_n(\omega) > 0\}$, and uniform integrability means

$$\lim_{k\to\infty} \sup_{t\in I} \int_{\{T_n>0\}\cap\{\|Z_{t\wedge T_n}\|\geq k\}} \|Z_{t\wedge T_n}\| \,\mathrm{d}\mathbb{P} = 0.$$

The class of local martingales includes time-changed Brownian motions and Itô diffusion processes. For more information on Itô diffusion, see also [29].

With the decomposition $Y_t = Y_0 + F_t + Z_t$ described above, we recall from Lemma 18.7 of [27] that

$$[Y, Y]_t = [Z, Z]_t + \sum_{0\leq s\leq t} \Delta F_s \Delta F_s^{\mathrm{tr}}.$$

Each step is bounded by the variation of $\{F_t\}_{t\in I}$, *i.e.*

$$\sum_{0\leq s\leq t} \|\Delta F_s\|^2 = \sum_{0\leq s\leq t} \|F_s - F_{s^-}\|^2 \leq \sum_{0\leq s\leq t} (\|F_s\| + \|F_{s^-}\|)\|F_s - F_{s^-}\| \leq 2\sup_{t\in I}\|F_t\| \sum_{0\leq s\leq t} \|F_s - F_{s^-}\|.$$

Since the number of jumps is countable, we find, by an approximation procedure (taking sequences s_j converging from below to each jump point s), that

$$\sum_{0\leq s\leq t} \|\Delta F_s\|^2 \leq \sup_{t\in I} \|F_t\| \|V(\{F_t\}_{t\in I}; I)\| \leq \|V(\{F_t\}_{t\in I}; I)\|^2.$$

We also have

$$\sum_{s\in J_Y} \|\Delta F_s\| \leq V(\{F_t\}_{t\in I}; I) < \infty. \tag{4.1}$$

In the more general case of a semi-martingale, the process $\{Y_t\}_{t\in I}$ has finite quadratic variation ([30], Sect. II.6) on the bounded interval $I = [0, T]$, and we find, using that $[Y, Y]_0 = Y_0 Y_0^{\mathrm{tr}}$ and $\Delta[Y, Y]_t = \Delta Y_t \Delta Y_t^{\mathrm{tr}}$ (see [30], Thm. II.22), that

$$\sum_{s\in J_Y} \|\Delta Y_s\|^2 \leq \|[Y, Y]_T\| < \infty. \tag{4.2}$$

By Doob's maximal quadratic inequality (see *e.g.* [30], Thm. I.20), if $Z_t = Z_{t\wedge T}$, $0 \leq t \leq T$, is a square-integrable martingale, then

$$\mathbb{E}\left[\sup_{t\in I} \|Z_t\|^2\right] \leq 4\mathbb{E}[\|Z_T\|^2] < \infty. \tag{4.3}$$

Since $F_0 = 0$, the finite variation part is bounded by

$$\sup_{t\in I}\|F_t\| \leq V\big(\{F_t\}_{t\in I}; I\big),$$

which implies, in particular,

$$\mathbb{E}\left[\sup_{t\in I} \|F_t\|\right] \leq \mathbb{E}[V(\{F_t\}_{t\in I}; I)], \qquad \mathbb{E}\left[\sup_{t\in I} \|F_t\|^2\right] \leq \mathbb{E}\Big[V(\{F_t\}_{t\in I}; I)^2\Big].$$

For $Y_t = Y_0 + Z_t + F_t$, we find

$$\mathbb{E}\left[\sup_{t\in I}\|Y_t\|\right] \leq \mathbb{E}[\|Y_0\|] + \mathbb{E}\big[V\big(\{F_t\}_{t\in I}; I\big)\big] + 2\mathbb{E}\Big[\|Z_T\|^2\Big]^{1/2}, \tag{4.4}$$

and

$$\mathbb{E}\left[\sup_{t\in I}\|Y_t\|^2\right] \leq 3\mathbb{E}\left[\|Y_0\|^2\right] + 3\mathbb{E}\left[V\left(\{F_t\}_{t\in I}; I\right)^2\right] + 12\mathbb{E}\left[\|Z_T\|^2\right]. \tag{4.5}$$

For the quadratic variation part of a local martingale $\{Z_t\}_{t\in I}$, the Burkholder–David–Gundy inequalities (see *e.g.* [30], Thm. IV.48 and [25]) relate the moments of the maximum of the local martingale with moments of its quadratic variation,

$$\frac{1}{\gamma_p}\mathbb{E}\left[\|[Z,Z]_t\|^{p/2}\right] \leq \mathbb{E}\left[\max_{0\leq s\leq t}\|Z_s\|^p\right] \leq \tilde{\gamma}_p\mathbb{E}\left[\|[Z,Z]_t\|^{p/2}\right],$$

for $p \geq 1$ and suitable constants $\gamma_p, \tilde{\gamma}_p > 0$. Thus, we obtain the bound

$$\mathbb{E}[\|[Z,Z]_t\|] \leq 4\gamma_2\mathbb{E}\left[\|Z_T\|^2\right] < \infty, \tag{4.6}$$

for any $t \in I$, and then

$$\mathbb{E}[\|[Y,Y]_t\|] \leq 4\gamma_2\mathbb{E}\left[\|Z_T\|^2\right] + \mathbb{E}\left[\|V(\{F_t\}_{t\in I}; I)\|^2\right]. \tag{4.7}$$

This, together with (4.2), also yields

$$\mathbb{E}\left[\sum_{s\in J_Y}\|\Delta Y_s\|^2\right] \leq 4\gamma_2\mathbb{E}\left[\|Z_T\|^2\right] + \mathbb{E}\left[\|V(\{F_t\}_{t\in I}; I)\|^2\right]. \tag{4.8}$$

The polarization identity

$$\left[Y^{(l_1)}, Y^{(l_2)}\right] = \frac{1}{2}\left(\left[Y^{(l_1)} + Y^{(l_2)}, Y^{(l_1)} + Y^{(l_2)}\right] - \left[Y^{(l_1)}, Y^{(l_1)}\right] - \left[Y^{(l_2)}, Y^{(l_2)}\right]\right)$$

implies that each quadratic covariation is a difference of quadratic variations, which are increasing processes, so that each quadratic covariation is also of finite variation (see [30], Cor. 1 of Thm. II.22). In particular, the continuous part of the quadratic local martingale term $\{Z_t\}_{t\in I}$ is of finite variation and it follows from the definition of the continuous part of the quadratic variation, (4.7), and (4.6) that

$$\mathbb{E}\left[V\left(\{[Z,Z]_t^c\}_{t\in I}; I\right)\right] \leq 5\gamma_2\mathbb{E}\left[\|Z_T\|^2\right] < \infty. \tag{4.9}$$

A fundamental tool for integration and differentiation is the chain rule. In the case of an FV process, given a continuously differentiable function $f : \mathbb{R}^k \to \mathbb{R}^d$, the following change of variables formula holds (see [30], Thms. II.31 and II.33):

$$f(Y_t) - f(Y_0) = \int_{0^+}^t Df(Y_{s^-})\,\mathrm{d}Y_s + \sum_{0<s\leq t}(f(Y_s) - f(Y_{s^-}) - Df(Y_{s^-})\Delta Y_s). \tag{4.10}$$

The summation on the right hand side is an at most countable summation since the summand vanishes for $s \notin J_Y$. The summation can be written as an integral with respect to a jump measure (see (4.12) for the full expression), but we keep it as a summation for simplicity.

The formula above is written in vectorial form, with f and Y_t of the form $f(y) = (f_i(y_1, \ldots, y_k))_{i=1,\ldots,d}$ and $Y_t = (Y_t^{(l)})_{l=1,\ldots,k}$, and

$$Df(y)\Delta Y_t = (\nabla f_i(y)\Delta Y_t)_{i=1,\ldots,d} = \left(\sum_{l=1,\ldots,k}\partial y_l f_i(y)\Delta Y_t^{(l)}\right)_{i=1,\ldots,d}.$$

When $f = f(t, y)$ depends also on t, one can apply the above formula to the extended process $\{(t, Y_t)\}_{t\in I}$. The first component is continuous, hence it has no jumps, and the corresponding summation term vanishes. This leads to the following change of variables formula for FV processes:

$$\begin{aligned} f(t, Y_t) - f(0, Y_0) = \int_0^t \partial_s f(s, Y_{s^-})\,\mathrm{d}s + \int_{0^+}^t D_y f(s, Y_{s^-})\,\mathrm{d}Y_s \\ + \sum_{0<s\leq t} (f(s, Y_s) - f(s, Y_{s^-}) - D_y f(s, Y_{s^-})\Delta Y_s). \end{aligned} \tag{4.11}$$

In the more general case of a semi-martingale noise $\{Y_t\}_{t\in I}$, the change of variables formula (4.11) takes the form (see *e.g.* [30], Thms. II.32 and II.33, [1], Sect. 1.3.3, [20], Thm. 3.6, or [15], Thm. 9.35)

$$\begin{aligned} f(t, Y_t) - f(0, Y_0) = \int_0^t \partial_s f(s, Y_{s^-})\,\mathrm{d}s + \int_{0^+}^t D_y f(s, Y_{s^-})\,\mathrm{d}Y_s \\ + \sum_{0<s\leq t} (f(s, Y_s) - f(s, Y_{s^-}) - D_y f(s, Y_{s^-})\Delta Y_s) + \frac{1}{2}\int_0^t D_{yy} f(s, Y_{s^-})\,\mathrm{d}[Y, Y]^c_s, \end{aligned} \tag{4.12}$$

where each component of $D_{yy} f(s, Y_{s^-})\,\mathrm{d}[Y, Y]^c_s$ is given by

$$(D_{yy} f(s, Y_{s^-})\,\mathrm{d}[Y, Y]^c_s)^{(i)} = \sum_{l_1,l_2=1,\ldots,d} \partial_{y_{l_1}}\partial_{y_{l_2}} f_i(s, Y_{s^-})\,\mathrm{d}\Big[Y^{(l_1)}, Y^{(l_2)}\Big]^c_s.$$

Since $[Y, Y]^c_t$ is of finite variation, the last integral in (4.12) is a Lebesgue–Stieltjes integral.

We actually apply the formula (4.12) to a function of the form $f = f(t, x, y)$, but with x constant, so the above formula suffices.

Finally, a fundamental identity is the Itô isometry, which, in the case of a vector-valued local martingale $\{Z_t\}_{t\in I}$ and an adapted matrix-valued process $\{A_t\}_{t\in I}$, takes the form (*e.g.* apply Chapt. 3, Prop. 2.17 of [20] to the expanded coordinate terms)

$$\mathbb{E}\left[\left\|\int_0^t A_s\,\mathrm{d}Z_s\right\|^2\right] = \mathbb{E}\left[\int_0^t A_s^{\mathrm{tr}} A_s \odot \mathrm{d}[Z, Z]_s\right] = \mathbb{E}\left[\int_0^t \sum_{l_1,l_2=1}^k \sum_{i=1}^d A_s^{(il_1)} A_s^{(il_2)}\,\mathrm{d}\Big[Z^{(l_1)}, Z^{(l_2)}\Big]_s\right], \tag{4.13}$$

provided each component $\{A^{il}_t\}_{t\in I}$ is square integrable for the integration with respect to $\{[Z^{(l)}, Z^{(l)}]_t\}_{t\in I}$, and where $\odot$ denotes the Hadamard product of two matrices of same dimensions. Recall that the steps of the cross variation matrix $[Z, Z]_s$ are positive semi-definite, so the right hand side is nonnegative.

5. Integral formula for the global pathwise error

In this section, we derive the following integral formula for the global error:

Lemma 5.1. *Under the Standing Hypothesis 2.1, the Euler approximation* (1.2) *for a pathwise solution of the random ordinary differential equation* (1.1) *satisfies almost surely the global error formula*

$$\begin{aligned} X_{t_j} - X^N_{t_j} = X_0 - X^N_0 + \int_0^{t_j} \big(f(s, X_s, Y_s) - f\big(s, X_{\tau^N(s)}, Y_s\big)\big)\,\mathrm{d}s \\ + \int_0^{t_j} \Big(f\big(s, X_{\tau^N(s)}, Y_s\big) - f\Big(s, X^N_{\tau^N(s)}, Y_s\Big)\Big)\,\mathrm{d}s \\ + \int_0^{t_j} \Big(f\Big(s, X^N_{\tau^N(s)}, Y_s\Big) - f\Big(\tau^N(s), X^N_{\tau^N(s)}, Y_{\tau^N(s)}\Big)\Big)\,\mathrm{d}s, \end{aligned} \tag{5.1}$$

for $j = 1, \ldots, N$, where τ^N is the piecewise constant function along the time mesh which gives the running time of the Euler method:

$$\tau^N(t) = \max_{t_j \le t}\{t_j\} = \left\lfloor \frac{t}{\Delta t_N} \right\rfloor \Delta t_N = \left\lfloor \frac{tN}{T} \right\rfloor \frac{T}{N}. \tag{5.2}$$

Proof. Under the Standing Hypothesis 2.1, the solutions of (1.1) are pathwise solutions in the sense of (2.2). With that in mind, we first obtain an expression for a single time step, from time t_{j-1} to $t_j = t_{j-1} + \Delta t_N$.

Subtracting the Euler step from the pathwise solution on a single time step yields

$$X_{t_j} - X_{t_j}^N = X_{t_{j-1}} - X_{t_{j-1}}^N + \int_{t_{j-1}}^{t_j} \Big(f(s, X_s, Y_s) - f\Big(t_{j-1}, X_{t_{j-1}}^N, Y_{t_{j-1}}\Big)\Big)\,\mathrm{d}s.$$

Adding and subtracting appropriate terms yield

$$\begin{aligned} X_{t_j} - X_{t_j}^N = X_{t_{j-1}} - X_{t_{j-1}}^N &+ \int_{t_{j-1}}^{t_j} \left(f(s, X_s, Y_s) - f\left(s, X_{t_{j-1}}, Y_s\right)\right)\mathrm{d}s \\ &+ \int_{t_{j-1}}^{t_j} \Big(f\big(s, X_{t_{j-1}}, Y_s\big) - f\Big(s, X_{t_{j-1}}^N, Y_s\Big)\Big)\,\mathrm{d}s \\ &+ \int_{t_{j-1}}^{t_j} \Big(f\Big(s, X_{t_{j-1}}^N, Y_s\Big) - f\Big(t_{j-1}, X_{t_{j-1}}^N, Y_{t_{j-1}}\Big)\Big)\,\mathrm{d}s. \end{aligned} \tag{5.3}$$

Now we iterate the time steps (5.3) to find that

$$\begin{aligned} X_{t_j} - X_{t_j}^N = X_0 - X_0^N + \sum_{i=1}^{j} \Bigg(&\int_{t_{i-1}}^{t_i} \left(f(s, X_s, Y_s) - f\left(s, X_{t_i}, Y_s\right)\right)\mathrm{d}s \\ &+ \int_{t_{i-1}}^{t_i} \Big(f\big(s, X_{t_{i-1}}, Y_s\big) - f\Big(s, X_{t_{i-1}}^N, Y_s\Big)\Big)\,\mathrm{d}s \\ &+ \int_{t_{i-1}}^{t_i} \Big(f\Big(s, X_{t_{i-1}}^N, Y_s\Big) - f\Big(t_{i-1}, X_{t_{i-1}}^N, Y_{t_{i-1}}\Big)\Big)\,\mathrm{d}s \Bigg). \end{aligned}$$

Using the running time τ^N defined by (5.2), the above expression becomes (5.1), as desired. □

6. Basic bound for the global error

Here we derive a bound for the global error which is the basis for the convergence bound. For that, we use the following classical discrete version of the Gronwall Lemma, which can be deduced from *e.g.* Lemma 1 of [33].

Lemma 6.1 (Discrete Gronwall Lemma)**.** *Let $(e_j)_j$ be a (finite or infinite) sequence of positive numbers starting at $j = 0$ and satisfying*

$$e_j \le \sum_{i=0}^{j-1} a_i e_i + b_i, \tag{6.1}$$

for every j, and where $a_i, b_i \ge 0$, with b_i nondecreasing. Then, for all j, we have

$$e_j \le b_j e^{\sum_{i=0}^{j-1} a_i}. \tag{6.2}$$

With that, we are ready to prove our basic bound for the global error.

Proposition 6.1 (Basic strong bound). *Under the Standing Hypothesis 2.1, assume that (2.6) holds. Then, for every $j = 0, \ldots, N$,*

$$\mathbb{E}\Big[\Big\|X_{t_j} - X_{t_j}^N\Big\|\Big] \leq \Bigg(\mathbb{E}\big[\big\|X_0 - X_0^N\big\|\big] + \int_0^{t_j} c_L \mathbb{E}\big[\big\|X_s - X_{\tau^N(s)}\big\|\big]\,\mathrm{d}s \\ + \mathbb{E}\bigg[\bigg\|\int_0^{t_j}\Big(f\Big(s, X_{\tau^N(s)}^N, Y_s\Big) - f\Big(\tau^N(s), X_{\tau^N(s)}^N, Y_{\tau^N(s)}\Big)\Big)\,\mathrm{d}s\bigg\|\bigg]\Bigg)e^{c_L t_j}. \tag{6.3}$$

Proof. We bound the first two integrals in (5.1). For the first one, we use (2.1) together with the assumed (2.6), so that

$$\|f(s, X_s, Y_s) - f(s, X_t, Y_s)\| \leq c_L\|X_s - X_t\|,$$

for $t, s \in I$, and, in particular, for $t = \tau^N(s)$. Hence,

$$\bigg\|\int_0^{t_j}\big(f(s, X_s, Y_s) - f\big(s, X_{\tau^N(s)}, Y_s\big)\big)\,\mathrm{d}s\bigg\| \leq c_L\int_0^{t_j}\big\|X_s - X_{\tau^N(s)}\big\|\,\mathrm{d}s.$$

For the second term, we use again (2.1), and find that

$$\bigg\|\int_0^{t_j}\Big(f\big(s, X_{\tau^N(s)}, Y_s\big) - f\Big(s, X_{\tau^N(s)}^N, Y_s\Big)\Big)\,\mathrm{d}s\bigg\| \leq c_L\int_0^{t_j}\Big\|X_{\tau^N(s)} - X_{\tau^N(s)}^N\Big\|\,\mathrm{d}s \\ \leq c_L\sum_{i=0}^{j-1}\big\|X_{t_i} - X_{t_i}^N\big\|\Delta t.$$

Using these two bounds in (5.1) yields

$$\Big\|X_{t_j} - X_{t_j}^N\Big\| \leq \big\|X_0 - X_0^N\big\| + c_L\int_0^{t_j}\big\|X_s - X_{\tau^N(s)}\big\|\,\mathrm{d}s + c_L\sum_{i=0}^{j-1}\big\|X_{t_i} - X_{t_i}^N\big\|\Delta t \\ + \bigg\|\int_0^{t_j}\Big(f\Big(s, X_{\tau^N(s)}^N, Y_s\Big) - f\Big(\tau^N(s), X_{\tau^N(s)}^N, Y_{\tau^N(s)}\Big)\Big)\,\mathrm{d}s\bigg\|.$$

This can be cast in the form of (6.1), with a constant coefficient $a_i = c_L\Delta t$. Then, using the discrete Gronwall Lemma 6.1, we obtain from (6.2) that

$$\Big\|X_{t_j} - X_{t_j}^N\Big\| \leq \Bigg(\big\|X_0 - X_0^N\big\| + c_L\int_0^{t_j}\big\|X_s - X_{\tau^N(s)}\big\|\,\mathrm{d}s \\ + \bigg\|\int_0^{t_j}\Big(f\Big(s, X_{\tau^N(s)}^N, Y_s\Big) - f\Big(\tau^N(s), X_{\tau^N(s)}^N, Y_{\tau^N(s)}\Big)\Big)\,\mathrm{d}s\bigg\|\Bigg)e^{c_L t_j}, \tag{6.4}$$

for $j = 1, \ldots, N$, where τ^N is given by (5.2). Taking the expectation of (6.4), we obtain the desired result (6.3). ☐

The first term in the right hand side of (6.3) usually vanishes since in general we take $X_0^N = X_0$, but it suffices to assume that X_0^N approximates X_0 to first order Δt_N, which is useful for lower order approximations or for discretizations of (random) partial differential equations.

The third term in (6.3) is the more delicate one that will be handled differently in the next section.

As for the second term, which only concerns the solution itself, not the approximation, we use the following crude but useful general result.

Lemma 6.2. *Under the Standing Hypothesis 2.1, assume that (2.4)–(2.6) hold. Then, we find that, for $j = 1, \ldots, N$,*

$$\mathbb{E}\left[\int_0^{t_j} \left\|X_s - X_{\tau^N(s)}\right\| \mathrm{d}s\right] \leq \Delta t_N \left(\mathbb{E}[\|X_0\|] + \int_0^{t_j} \mathbb{E}[\|f(s,0,Y_s)\|] \,\mathrm{d}s\right) e^{c_L t_j}. \tag{6.5}$$

Proof. Using that the paths of $\{X_t\}_{t\in I}$ are almost surely absolutely continuous, adding and subtracting $f(\xi, 0, Y_\xi)$ to the integrand, and using assumption (ii) inside the Standing Hypothesis 2.1 along with (2.6), we have almost surely that

$$\left\|X_s - X_{\tau^N(s)}\right\| = \left\|\int_{\tau^N(s)}^s f(\xi, X_\xi, Y_\xi)\,\mathrm{d}\xi\right\| \leq \int_{\tau^N(s)}^s (\|f(\xi,0,Y_\xi)\| + c_L\|X_\xi\|)\,\mathrm{d}\xi.$$

Integrating over $[0, t_j]$ and using Fubini's theorem to exchange the order of integration,

$$\begin{aligned}\int_0^{t_j} \left\|X_s - X_{\tau^N(s)}\right\| \mathrm{d}s &\leq \int_0^{t_j}\int_{\tau^N(s)}^s (\|f(\xi,0,Y_\xi)\| + c_L\|X_\xi\|)\,\mathrm{d}\xi\,\mathrm{d}s \\ &\leq \int_0^{t_j} \left(\tau^N(\xi) + \Delta t_N - \xi\right)(\|f(\xi,0,Y_\xi)\| + c_L\|X_\xi\|)\,\mathrm{d}\xi.\end{aligned}$$

Using that $\tau^N(\xi) \leq \xi$, we have $\tau^N(\xi) + \Delta t_N - \xi \leq \Delta t_N$, and we obtain, upon taking the expectation

$$\mathbb{E}\left[\int_0^{t_j} \left\|X_s - X_{\tau^N(s)}\right\| \mathrm{d}s\right] \leq \Delta t_N \int_0^{t_j} (\mathbb{E}[\|f(s,0,Y_s)\|] + c_L\mathbb{E}[\|X_s\|])\,\mathrm{d}s.$$

Thanks to (2.4)–(2.6), Lemma 2.2 applies and inequality (2.8) holds, so that we obtain

$$\begin{aligned}\mathbb{E}\left[\int_0^{t_j} \left\|X_s - X_{\tau^N(s)}\right\| \mathrm{d}s\right] &\leq \Delta t_N \bigg(\int_0^{t_j} \mathbb{E}[\|f(s,0,Y_s)\|]\,\mathrm{d}s \\ &\qquad + c_L \int_0^{t_j}\left(\mathbb{E}[\|X_0\|] + \int_0^s \mathbb{E}[\|f(\xi,0,Y_\xi)\|]\,\mathrm{d}\xi\right) e^{c_L s}\,\mathrm{d}s\bigg) \\ &\leq \Delta t_N \bigg(\int_0^{t_j} \mathbb{E}[\|f(s,0,Y_s)\|]\,\mathrm{d}s \\ &\qquad + \left(\mathbb{E}[\|X_0\|] + \int_0^{t_j} \mathbb{E}[\|f(\xi,0,Y_\xi)\|]\,\mathrm{d}\xi\right)\left(e^{c_L t_j} - 1\right)\bigg).\end{aligned}$$

This simplifies to (6.5). □

7. Strong convergence for semi-martingale noises

We now prove the desired strong order 1 convergence.

Theorem 7.1. *Consider the Random ODE (1.1) and the Euler approximation (1.2), with the function $f = f(t,x,y)$, the initial condition X_0, and the noise $\{Y_t\}_{t\in I}$ satisfying the Standing Hypothesis 2.1. Suppose that (2.4)–(2.6) hold; that X_0^N is a first order approximation of X_0, in the sense that*

$$\mathbb{E}\left[\left\|X_0 - X_0^N\right\|\right] \leq c_0\Delta t_N, \qquad \forall N \in \mathbb{N}, \tag{7.1}$$

for some constant $c_0 \geq 0$; and that $f = f(t,x,y)$ is uniformly globally Lipschitz continuous in x, continuously differentiable with respect to t and twice continuously differentiable with respect to y, with partial differentials $\partial_t f$, $D_y f$ and $D_{yy} f$ satisfying the bounds

$$\|\partial_t f(t,x,y)\| \leq c_1 + c_2\|x\| + c_3\|y\|, \tag{7.2}$$

$$\|D_y f(t,x,y)\| \leq c_4, \tag{7.3}$$
$$\|D_{yy} f(t,x,y)\| \leq c_7, \tag{7.4}$$

in $(t,x,y) \in I \times \mathbb{R}^d \times \mathbb{R}^k$*, for suitable constants* $c_1, c_2, c_3, c_4, c_7 \geq 0$*. Assume the semi-martingale noise* $\{Y_t\}_{t\in I}$ *is decomposed as* $Y_t = Y_0 + F_t + Z_t$*, where* $\{F_t\}_{t\in I}$ *is an FV process,* $\{Z_{t\wedge T}\}_{t\geq 0}$ *is a martingale, and* $F_0 = Z_0 = 0$*, and assume they satisfy*

$$\mathbb{E}[\|Y_0\|] < \infty, \tag{7.5}$$
$$\mathbb{E}\Big[V(\{F_t\}_{t\in I}; I)^2\Big] < \infty, \tag{7.6}$$
$$\mathbb{E}\big[\|Z_T\|^2\big] < \infty. \tag{7.7}$$

Then, the Euler scheme is strongly convergent of order 1, i.e.

$$\max_{j=0,\ldots,N} \mathbb{E}\Big[\Big\|X_{t_j} - X_{t_j}^N\Big\|\Big] \leq c\Delta t_N, \qquad \forall N \in \mathbb{N}, \tag{7.8}$$

for a suitable constant $c \geq 0$ *independent of* N*.*

Proof. We apply Proposition 6.1. The troublesome term is the last one in (6.3). Using (4.12), the integral in this last term reads

$$\begin{aligned}\int_0^{t_j} &\Big(f\Big(s, X_{\tau^N(s)}^N, Y_s\Big) - f\Big(\tau^N(s), X_{\tau^N(s)}^N, Y_{\tau^N(s)}\Big)\Big)\,\mathrm{d}s \\ &= \int_0^{t_j}\int_{\tau^N(s)}^s \partial_\xi f\Big(\xi, X_{\tau^N(s)}^N, Y_{\xi-}\Big)\,\mathrm{d}\xi\,\mathrm{d}s \\ &\quad + \int_0^{t_j}\int_{\tau^N(s)^+}^s D_y f\Big(\xi, X_{\tau^N(s)}^N, Y_{\xi-}\Big)\,\mathrm{d}Y_\xi\,\mathrm{d}s \\ &\quad + \int_0^{t_j}\sum_{\tau^N(s)<\xi\leq s}\Big(f\Big(\xi, X_{\tau^N(s)}^N, Y_\xi\Big) - f\Big(\xi, X_{\tau^N(s)}^N, Y_{\xi-}\Big) \\ &\qquad\qquad - D_y f\Big(\xi, X_{\tau^N(s)}^N, Y_{\xi-}\Big)\Delta Y_\xi\Big)\,\mathrm{d}s \\ &\quad + \frac{1}{2}\int_0^{t_j}\int_{\tau^N(s)}^s D_{yy} f\Big(\xi, X_{\tau^N(s)}^N, Y_{\xi-}\Big)\,\mathrm{d}[Y,Y]_\xi^c\,\mathrm{d}s.\end{aligned} \tag{7.9}$$

The bound on the expectation of the first term in (7.9) is obtained using (7.2), so that

$$\begin{aligned}\mathbb{E}&\Big[\Big\|\int_0^{t_j}\int_{\tau^N(s)}^s \partial_\xi f\Big(\xi, X_{\tau^N(s)}^N, Y_{\xi-}\Big)\,\mathrm{d}\xi\,\mathrm{d}s\Big\|\Big] \\ &\leq \mathbb{E}\bigg[\int_0^{t_j}\int_{\tau^N(s)}^s \Big(c_1 + c_2\Big\|X_{\tau^N(s)}^N\Big\| + c_3\big\|Y_{\xi-}\big\|\Big)\,\mathrm{d}\xi\,\mathrm{d}s\bigg] \\ &\leq t_j\Delta t_N\bigg(c_1 + c_2\sup_{0\leq s\leq t_j}\mathbb{E}\Big[\Big\|X_{\tau^N(s)}^N\Big\|\Big] + c_3\sup_{0\leq s\leq t_j}\mathbb{E}[\|Y_s\|]\bigg).\end{aligned} \tag{7.10}$$

For the second term in (7.9), we use the decomposition $Y_t = Y_0 + F_t + Z_t$. With that, the second term can be written as

$$\begin{aligned}\int_0^{t_j}\int_{\tau^N(s)^+}^s D_y f\Big(\xi, X_{\tau^N(s)}^N, Y_{\xi-}\Big)\,\mathrm{d}Y_\xi\,\mathrm{d}s &= \int_0^{t_j}\int_{\tau^N(s)^+}^s D_y f\Big(\xi, X_{\tau^N(s)}^N, Y_{\xi-}\Big)\,\mathrm{d}F_\xi\,\mathrm{d}s \\ &\quad + \int_0^{t_j}\int_{\tau^N(s)^+}^s D_y f\Big(\xi, X_{\tau^N(s)}^N, Y_{\xi-}\Big)\,\mathrm{d}Z_\xi\,\mathrm{d}s,\end{aligned} \tag{7.11}$$

where the first inner integral is a Lebesgue–Stieltjes integral and the second inner integral is a stochastic integral with respect to a local martingale. Since $\{F_t\}_{t\in I}$ is of finite variation, the bound on the first integral on the right hand side of (7.11) is obtained using (7.3) and switching the order of integration, so that

$$\left\|\int_0^{t_j}\int_{\tau^N(s)+}^{s} D_y f\left(\xi, X^N_{\tau^N(s)}, Y_{\xi-}\right) \mathrm{d}F_\xi \,\mathrm{d}s\right\| \leq \int_0^{t_j}\int_{\tau^N(s)}^{s} c_4 \,\|\mathrm{d}F_\xi\|\,\mathrm{d}s$$
$$\leq c_4\int_0^{t_j}\int_{\xi}^{\tau^N(\xi)+\Delta t_N} \mathrm{d}s\,\|\mathrm{d}F_\xi\| \leq c_4\Delta t_N\int_0^{t_j}\|\mathrm{d}F_\xi\| \leq c_4\Delta t_N V(\{F_t\}_{t\in I};[0,t_j]).$$

Taking the expectation and using (7.6), we obtain

$$\mathbb{E}\left[\left\|\int_0^{t_j}\int_{\tau^N(s)+}^{s} D_y f\left(\xi, X^N_{\tau^N(s)}, Y_{\xi-}\right) \mathrm{d}F_\xi \,\mathrm{d}s\right\|\right] \leq c_4\Delta t_N\mathbb{E}[V(\{F_t\}_{t\in I};[0,t_j])]. \tag{7.12}$$

For the second integral on the right hand side of (7.11), we once again switch the order of integration, then use the Lyapunov inequality, the Itô isometry (4.13) for local martingales, and then again the bound (7.3), so that

$$\begin{aligned}
&\mathbb{E}\left[\left\|\int_0^{t_j}\int_{\tau^N(s)+}^{s} D_y f\left(\xi, X^N_{\tau^N(s)}, Y_{\xi-}\right) \mathrm{d}Z_\xi \,\mathrm{d}s\right\|\right]\\
&\qquad= \mathbb{E}\left[\left\|\int_0^{t_j}\int_{\xi}^{\tau^N(\xi)+\Delta t_N} D_y f\left(\xi, X^N_{\tau^N(s)}, Y_{\xi-}\right) \mathrm{d}s\,\mathrm{d}Z_\xi\right\|\right]\\
&\qquad\leq \mathbb{E}\left[\left\|\int_0^{t_j}\int_{\xi}^{\tau^N(\xi)+\Delta t_N} D_y f\left(\xi, X^N_{\tau^N(\xi)}, Y_{\xi-}\right) \mathrm{d}s\,\mathrm{d}Z_\xi\right\|^2\right]^{1/2}\\
&\qquad\leq c_4\Delta t_N\mathbb{E}\left[\sum_{l_1,l_2=1}^{k}\int_0^{t_j}\left|\mathrm{d}\left[Z^{(l_1)}, Z^{(l_2)}\right]_\xi\right|\right]^{1/2},
\end{aligned}$$

where we used that, for any $0\leq\xi\leq T$, and for $\xi\leq s\leq\tau^N(\xi)+\Delta t_N$, we have $\tau^N(s)=\tau^N(\xi)$, so we can write $X^N_{\tau^N(s)} = X^N_{\tau^N(\xi)}$. Now, using that each quadratic variation and covariation is of finite variation, we obtain

$$\begin{aligned}
&\mathbb{E}\left[\left\|\int_0^{t_j}\int_{\tau^N(s)+}^{s} D_y f\left(\xi, X^N_{\tau^N(s)}, Y_{\xi-}\right) \mathrm{d}Z_\xi \,\mathrm{d}s\right\|\right]\\
&\qquad\leq c_4\Delta t_N\mathbb{E}\left[\sum_{l_1,l_2=1}^{k} V\left(\left\{\left[Z^{(l_1)}, Z^{(l_2)}\right]_t\right\}_{t\in I};[0,t_j]\right)\right]^{1/2}.
\end{aligned} \tag{7.13}$$

Now we move to the third term in (7.9). We use the bound (7.4) on the second derivative of $D_{yy}f$ to write

$$\left\|f\left(\xi, X^N_{\tau^N(s)}, Y_\xi\right) - f\left(\xi, X^N_{\tau^N(s)}, Y_{\xi-}\right) - D_y f\left(\xi, X^N_{\tau^N(s)}, Y_{\xi-}\right)\Delta Y_\xi\right\| \leq \frac{c_7}{2}\|\Delta Y_\xi\|^2.$$

Recall $\tau^N(s)=\tau^N(\xi)$ for $\xi\leq s<\tau^N(\xi)+\Delta t_N$ and $\tau^N(\xi)+\Delta t_N-\xi\leq\Delta t_N$. Thus,

$$\left\|\int_0^{t_j}\sum_{\tau^N(s)<\xi\leq s}\left(f\left(\xi, X^N_{\tau^N(s)}, Y_\xi\right) - f\left(\xi, X^N_{\tau^N(s)}, Y_{\xi-}\right) - D_y f\left(\xi, X^N_{\tau^N(s)}, Y_{\xi-}\right)\Delta Y_\xi\right)\mathrm{d}s\right\|$$

$$
\begin{aligned}
&\leq \frac{c_7}{2}\int_0^{t_j} \sum_{\tau^N(s)<\xi\leq s} \|\Delta Y_\xi\|^2 \,\mathrm{d}s \\
&\leq \frac{c_7}{2}\sum_{0<\xi\leq t_j} \int_\xi^{\tau^N(\xi)+\Delta t_N} \|\Delta Y_\xi\|^2 \,\mathrm{d}s \\
&\leq \frac{c_7}{2}\Delta t_N \sum_{0<\xi\leq t_j} \|\Delta Y_\xi\|^2.
\end{aligned}
$$

Taking the expectation,

$$
\begin{aligned}
\mathbb{E}\Bigg[\Bigg\|\int_0^{t_j} \sum_{\tau^N(s)<\xi\leq s} \bigg(f\Big(\xi, X^N_{\tau^N(s)}, Y_\xi\Big) - f\Big(\xi, X^N_{\tau^N(s)}, Y_{\xi^-}\Big) \\
- D_y f\Big(\xi, X^N_{\tau^N(s)}, Y_{\xi^-}\Big)\Delta Y_\xi\bigg)\,\mathrm{d}s\Bigg\|\Bigg] \leq \frac{c_7}{2}\Delta t_N \mathbb{E}\Bigg[\Bigg\|\sum_{0<\xi\leq t_j} \|\Delta Y_\xi\|^2\Bigg\|\Bigg]. \qquad (7.14)
\end{aligned}
$$

The final term in (7.9) involves a Lebesgue–Stieltjes integral and is handled in a similar way, except that instead of the Itô isometry we just bound the Euclidian norm by the sum norm and obtain

$$
\begin{aligned}
&\left\|\frac{1}{2}\int_0^{t_j}\int_{\tau^N(s)}^s D_{yy} f\Big(\xi, X^N_{\tau^N(s)}, Y_{\xi^-}\Big)\,\mathrm{d}[Y,Y]^c_\xi \,\mathrm{d}s\right\| \\
&\quad = \left\|\frac{1}{2}\int_0^{t_j}\int_\xi^{\tau^N(\xi)+\Delta t_N} D_{yy} f\Big(\xi, X^N_{\tau^N(s)}, Y_{\xi^-}\Big)\,\mathrm{d}s\,\mathrm{d}[Y,Y]^c_\xi\right\| \\
&\quad \leq \frac{1}{2}\sum_{i=1}^n \left|\int_0^{t_j}\int_\xi^{\tau^N(\xi)+\Delta t_N} \Big(D_{yy} f\Big(\xi, X^N_{\tau^N(s)}, Y_{\xi^-}\Big)\,\mathrm{d}s\,\mathrm{d}[Y,Y]^c_\xi\Big)^{(i)}\right| \\
&\quad \leq \frac{1}{2}\sum_{i=1}^n \left|\int_0^{t_j}\int_\xi^{\tau^N(\xi)+\Delta t_N} \sum_{l_1,l_2=1,\ldots,d} \partial_{y_{l_1}}\partial_{y_{l_2}} f_i(s, Y_{s^-})\,\mathrm{d}s\,\mathrm{d}\Big[Y^{(l_1)}, Y^{(l_2)}\Big]^c_\xi\right| \\
&\quad \leq \frac{1}{2}\sum_{i=1}^n \sum_{l_1,l_2=1,\ldots,d} \int_0^{t_j}\left(\int_\xi^{\tau^N(\xi)+\Delta t_N} \left|\partial_{y_{l_1}}\partial_{y_{l_2}} f_i(s, Y_{s^-})\right|\mathrm{d}s\right)\left|\mathrm{d}\Big[Y^{(l_1)}, Y^{(l_2)}\Big]^c_\xi\right| \\
&\quad \leq \frac{1}{2}\sum_{i=1}^n \sum_{l_1,l_2=1,\ldots,d} \int_0^{t_j}\left(\int_\xi^{\tau^N(\xi)+\Delta t_N} c_7\,\mathrm{d}s\right)\left|\mathrm{d}\Big[Y^{(l_1)}, Y^{(l_2)}\Big]^c_\xi\right| \\
&\quad \leq \frac{c_7}{2}\Delta t_N \sum_{i=1}^n \sum_{l_1,l_2=1,\ldots,d} V\left(\left\{\Big[Y^{(l_1)}, Y^{(l_2)}\Big]_t\right\}_{t\in I}; [0,t_j]\right).
\end{aligned}
$$

Taking the expectation,

$$
\begin{aligned}
\mathbb{E}\left[\left\|\frac{1}{2}\int_0^{t_j}\int_{\tau^N(s)}^s D_{yy} f\Big(\xi, X^N_{\tau^N(s)}, Y_{\xi^-}\Big)\,\mathrm{d}[Y,Y]^c_\xi\,\mathrm{d}s\right\|\right] \\
\leq \frac{c_7}{2}\Delta t_N \sum_{i=1}^n \sum_{l_1,l_2=1,\ldots,d} \mathbb{E}\left[V\left(\left\{\Big[Y^{(l_1)}, Y^{(l_2)}\Big]_t\right\}_{t\in I}; [0,t_j]\right)\right]. \qquad (7.15)
\end{aligned}
$$

Putting the bounds (7.10)–(7.15) together, we obtain, from (7.9), that

$$
\mathbb{E}\left[\left\|\int_0^{t_j} \Big(f\Big(s, X^N_{\tau^N(s)}, Y_s\Big) - f\Big(\tau^N(s), X^N_{\tau^N(s)}, Y_{\tau^N(s)}\Big)\Big)\,\mathrm{d}s\right\|\right] \leq C_T \Delta t_N, \qquad (7.16)
$$

where

$$
\begin{aligned}
C_T = {} & c_1 T + c_2 T \sup_{0\leq s\leq T} \mathbb{E}\left[\left\|X^N_{\tau^N(s)}\right\|\right] + c_3 T \sup_{0\leq s\leq T} \mathbb{E}[\|Y_s\|] \\
& + c_4 \mathbb{E}\left[V\left(\{F_t\}_{t\in I}; I\right)\right] + c_4 \mathbb{E}\left[\sum_{l_1,l_2=1}^{k} V\left(\left\{\left[Z^{(l_1)}, Z^{(l_2)}\right]_t\right\}_{t\in I}; I\right)\right]^{1/2} \\
& + \frac{c_7}{2}\mathbb{E}\left[\sum_{0<s\leq T} \|\Delta Y_s\|^2\right] + \frac{c_7}{2}\sum_{i=1}^{n} \sum_{l_1,l_2=1,\ldots,d} V\left(\left\{\left[Y^{(l_1)}, Y^{(l_2)}\right]_t\right\}_{t\in I}; [0,t_j]\right).
\end{aligned}
$$

Thanks to (2.4), (2.5), (2.6), (3.3), (7.5), (7.6), (7.7), (4.4), (4.3), (4.9), (4.7), and (4.8), the constant C_T above is finite and independent of N. Plugging (7.16) into (6.3) and applying (6.5) and (7.1) yield the desired strong order 1 convergence (7.8). □

Remark 7.1. If the initial condition X_0 is bounded almost surely, then (2.3) shows that X_t is also bounded almost surely, uniformly on a bounded interval $t \in [0,T]$. In this case, the conditions on the derivative of $f = f(t,x,y)$ can be relaxed to have an arbitrary growth on x. Similarly, if the noise has bounded support almost surely, uniformly on $t \in [0,T]$, then the derivatives are allowed to grow with y. If the initial condition has no bounded support but higher moments are bounded, then appropriate power growths on x are allowed. Even the uniform Lipschitz condition (2.6) may be relaxed, provided higher-order moments of X_0 are finite.

8. Numerical example

We have briefly mentioned, in the Introduction, several numerical examples confirming the strong order one convergence and therefore showing this rate is optimal. Here, we describe how the order of convergence is estimated numerically and then we give more details on one of the examples, namely the non-homogeneous linear example with different types of noises. The remaining examples can be found in the github repository [22].

8.1. Estimating the order of convergence

For estimating the order of convergence, we use the Monte Carlo method, computing a number of numerical approximations $\{X_{t_j}^{N_i}(\omega_m)\}_{j=0,\ldots,N_i}$, of sample path solutions $\{X_t(\omega_m)\}_{t\in I}$, for samples ω_m, with $m = 1,\ldots,M$, and mesh resolutions N_i, with $i = 1,\ldots,i_{\max}$, and taking the maximum in time of the average of their absolute differences at the mesh points:

$$
\epsilon^{N_i} = \max_{j=0,\ldots,N_i} \frac{1}{M}\sum_{m=1}^{M}\left|X_{t_j}(\omega_m) - X_{t_j}^{N_i}(\omega_m)\right|. \tag{8.1}
$$

Then we fit the errors ϵ^{N_i} to the power law $C\Delta t_{N_i}^p$, in order to find p, along with a 95% confidence interval. The fitting process and the estimation of the confidence intervals follow classical methods in Statistics [8,14,17,19]. Reporting confidence intervals in the context of approximations of stochastic differential equations, in particular, can also be seen in [23].

Here are the main parameters for the error estimate:

(i) The number $M \in \mathbb{N}$ of samples for the Monte Carlo estimate of the strong error.
(ii) The time interval $[0,T]$, $T > 0$, for the initial-value problem.
(iii) The distribution law for the random initial condition X_0.
(iv) A series of time steps $\Delta t_{N_i} = T/N_i$, with $N_i = 2^{n_i}$, for some $n_i \in \mathbb{N}$, $i = 1,\ldots,i_{\max}$, so that finer meshes are refinements of coarser meshes.

(v) A number $N_{\text{tgt}} = 2^{n_{\text{tgt}}}$ for a finer resolution to compute a target solution path, typically $N_{\text{tgt}} = \max_i\{N_i^2\}$, unless an exact pathwise solution is available, in which case a coarser mesh of the order of $\max_i\{N_i\}$ can be used.

With this setup, we proceed as follows.

(i) For each sample ω_m, $m = 1, \ldots, M$, we first generate a discretization of a sample path of the noise, $\{Y_{t_j}\}_{j=0,\ldots,N_{\text{tgt}}}$, on the finest grid $t_j = j\Delta t_{N_{\text{tgt}}}$, $j = 0, \ldots, N_{\text{tgt}}$, using an exact distribution for the noise.

(ii) Next, we use the values of the noise at the target time mesh to generate the target solution $\{X_{t_j}\}_{j=0,\ldots,N_{\text{tgt}}}$, still on the fine mesh. This is constructed either using the Euler approximation itself, keeping in mind that the mesh is sufficiently fine, or an exact distribution of the solution, when available.

(iii) Then, for each time resolution N_i, we compute the Euler approximation using the computed noise values at the corresponding coarser mesh $t_j = j\Delta t_{N_i}$, $j = 0, \ldots, N_i$.

(iv) We then compare each approximation $\{X_{t_j}^{N_i}\}_{j=0,\ldots,N_i}$ to the values of the target path on that coarse mesh and update the strong error

$$\epsilon_{t_j}^{N_i} = \frac{1}{M}\sum_{m=1}^{M}\left|X_{t_j}(\omega_m) - X_{t_j}^{N_i}(\omega_m)\right|,$$

at each mesh point.

(v) At the end of all the simulations, we take the maximum in time, on each corresponding coarse mesh, to obtain the error for each mesh,

$$\epsilon^{N_i} = \max_{j=0,\ldots,N_i} \epsilon_{t_j}^{N_i}.$$

(vi) We fit $(\Delta t_{N_i}, \epsilon^{N_i})$ to the power law $C\Delta t_{N_i}^p$, *via* linear least-square regression in log scale, so that $\ln \epsilon^{N_i} \approx \ln C + p \ln \Delta t_{N_i}$, for suitable C and p, with p giving the order of convergence. This amounts to solving the normal equation $(A^{\text{tr}}A)\mathbf{v} = A^{\text{tr}}\ln(\boldsymbol{\epsilon})$, where $\mathbf{v}$ is the vector $\mathbf{v} = (\ln(C), p)$, A is the Vandermonde matrix associated with the logarithm of the mesh steps $(\Delta t_{N_i})_i$, and $\ln(\boldsymbol{\epsilon})$ is the vector $\ln(\boldsymbol{\epsilon}) = (\ln(\epsilon^{N_i}))_i$.

(vii) We also compute the standard error of the Monte-Carlo sample at each time step,

$$s_{t_j}^{N_i} = \frac{\sigma_{t_j}^{N_i}}{\sqrt{M}},$$

where $\sigma_{t_j}^{N_i}$ is the sample standard deviation given by

$$\sigma_{t_j}^{N_i} = \sqrt{\frac{1}{M-1}\sum_{m=1}^{M}\left(\left|X_{t_j}(\omega_m) - X_{t_j}^{N_i}(\omega_m)\right| - \epsilon_{t_j}^{N_i}\right)^2},$$

and compute the 95% confidence interval $[\epsilon_{\min}^{N_i}, \epsilon_{\max}^{N_i}]$ for the strong error at each mesh resolution N_i with

$$\epsilon_{\min}^{N_i} = \max_{j=0,\ldots,N_i}\left(\epsilon_{t_j}^{N_i} - 1.96\sigma_{t_j}^{N_i}\right), \quad \epsilon_{\max}^{N_i} = \max_{j=0,\ldots,N_i}\left(\epsilon_{t_j}^{N_i} + 1.96\sigma_{t_j}^{N_i}\right).$$

(viii) Next we estimate the 95% confidence interval for the joint distribution $\boldsymbol{\epsilon} = (\epsilon^{N_i})_{i=1,\ldots,i_{\max}}$ using the Bonferroni inequality [19]. More precisely, we find the z-score value for the $\alpha = 1 - 0.05/i_{\max}$ confidence level for a univariate normal distribution and use that to obtain the α confidence intervals $[\epsilon_{\alpha,\min}^{N_i}, \epsilon_{\alpha,\max}^{N_i}]$ for the marginals ϵ^{N_i} with

$$\epsilon_{\alpha,\min}^{N_i} = \max_{j=0,\ldots,N_i}\left(\epsilon_{t_j}^{N_i} - z\sigma_{t_j}^{N_i}\right), \quad \epsilon_{\alpha,\max}^{N_i} = \max_{j=0,\ldots,N_i}\left(\epsilon_{t_j}^{N_i} + z\sigma_{t_j}^{N_i}\right).$$

Then, the probability of the region $\Pi_{i=1}^{i_{\max}}[\epsilon_{\alpha,\min}^{N_i}, \epsilon_{\alpha,\max}^{N_i}]$ is bounded from below by $1 - i_{\max}(1-\alpha) = 1 - 0.05 = 0.95$, which means this is the confidence region of confidence level (at least) 95%, for $\boldsymbol{\epsilon}$.

(ix) Finally, from the normal equations and the confidence region for $\boldsymbol{\epsilon}$, we compute the (at least) 95% confidence interval $[p_{\min}, p_{\max}]$ for p *via* the transformation method. We do so by computing the minimum and maximum values of p in the image of the confidence region $\Pi_{i=1}^{i_{\max}}[\epsilon_{\alpha,\min}^{N_i}, \epsilon_{\alpha,\max}^{N_i}]$ by the map $\mathbf{e} \mapsto (\ln(C), p) = (A^{\mathrm{tr}}A)^{-1}A^{\mathrm{tr}}\ln(\mathbf{e})$. The map $\mathbf{e} \mapsto \ln(\mathbf{e})$ takes the hyperrectangle $\Pi_{i=1}^{i_{\max}}[\epsilon_{\alpha,\min}^{N_i}, \epsilon_{\alpha,\max}^{N_i}]$ into the hyperrectangle $\Pi_{i=1}^{i_{\max}}[\ln(\epsilon_{\alpha,\min}^{N_i}), \ln(\epsilon_{\alpha,\max}^{N_i})]$. This hyperrectangle is taken onto a polygon in the plane (C, p) by the linear map $(A^{\mathrm{tr}}A)^{-1}A^{\mathrm{tr}}$, which is then projected onto an interval $[p_{\min}, p_{\max}]$ in the p axis. Since the pseudo-inverse $(A^{\mathrm{tr}}A)^{-1}A^{\mathrm{tr}}$ and the projection are not bijective, the corresponding confidence level of this interval is at least 95%, but probably higher. In order to find $p_{\min}$ and $p_{\max}$, we just need to look at the vertices of the polygon in (C, p), which are the set of points obtained by all combinations of the indices in $(\epsilon_{\alpha,\,\min}^{N_i})_i$ and $(\epsilon_{\alpha,\,\max}^{N_i})_i$. We take the minimum of the p-coordinate of such points to be $p_{\min}$ and the maximum to be $p_{\max}$.

8.2. Non-homogeneous linear system of RODEs with different types of noises

We examine a linear equation with noise present in both the homogeneous and nonhomogeneous terms, approached in two ways: first, as a series of scalar equations, each with a unique univariate noise in a family of different types of noises; and secondly, as a system of equations that combines all these noise types. For most of the noises, the current knowledge expects a lower order of strong convergence than the strong order 1 we prove here. The aim of this section is to illustrate this improvement for all such noises.

The equations take the form

$$\begin{cases} \dfrac{\mathrm{d}X_t}{\mathrm{d}t} = -\|Y_t\|^2 X_t + Y_t, \qquad 0 \leq t \leq T, \\ X_t|_{t=0} = X_0, \end{cases} \tag{8.2}$$

where $\{X_t\}_{t\in I}$ is either a scalar or a vector-valued process, and $\{Y_t\}_{t\in I}$ is, correspondingly, either a scalar or a vector-valued noise process with the same dimension as X_t. In the case of a system, each coordinate of $\{Y_t\}_{t\in I}$ is a scalar noise process independent of the noises in the other coordinates. The scalar noises used in the simulations are the following, in the order of coordinates of Y_t:

(i) A standard Wiener process (W);
(ii) An Ornstein–Uhlenbeck process (OU) with drift $\nu = 0.3$, diffusion $\sigma = 0.5$, and initial condition $y_0 = 0.2$;
(iii) A geometric Brownian motion process (gBm) with drift $\mu = 0.3$, diffusion coefficient $\sigma = 0.5$, and initial condition $y_0 = 0.2$;
(iv) A non-autonomous homogeneous linear Itô process (hlp) $\{H_t\}_{t\in I}$ given by $\mathrm{d}H_t = (\mu_1 + \mu_2\sin(\vartheta t))H_t\,\mathrm{d}t + \sigma\sin(\vartheta t)H_t\,\mathrm{d}W_t$ with $\mu_1 = 0.5$, $\mu_2 = 0.3$, $\sigma = 0.5$, $\vartheta = 3\pi$, and initial condition $H_0 = 0.2$;
(v) A compound Poisson process (cP) with rate $\lambda = 5.0$ and jump law following an exponential distribution with scale $\theta = 0.5$;
(vi) A Poisson step process (sP) with rate $\lambda = 5.0$ and step law following a uniform distribution within the unit interval;
(vii) An exponentially decaying Hawkes process with initial rate $\lambda_0 = 3.0$, base rate $a = 2.0$, exponential decay rate $\delta = 3.0$, and jump law following an exponential distribution with scale $\theta = 0.5$;
(viii) A transport process of the form $t \mapsto \sum_{i=1}^{6}\sin^{1/3}(\omega_i t)$, where the frequencies ω_i are independent random variables following a Gamma distribution with shape parameter $\alpha = 7.5$ and scale $\theta = 2.0$;
(ix) A fractional Brownian motion (fBm) process with Hurst parameter $H = 0.6$ and initial condition $y_0 = 0.2$.

Table 1 shows the estimated strong error obtained from the Monte Carlo simulations, with the target solution obtained *via* Euler approximation on a much finer mesh.

Figure 7 illustrates the obtained order of convergence, while Figure 8 illustrates some sample paths of all the noises used in this system.

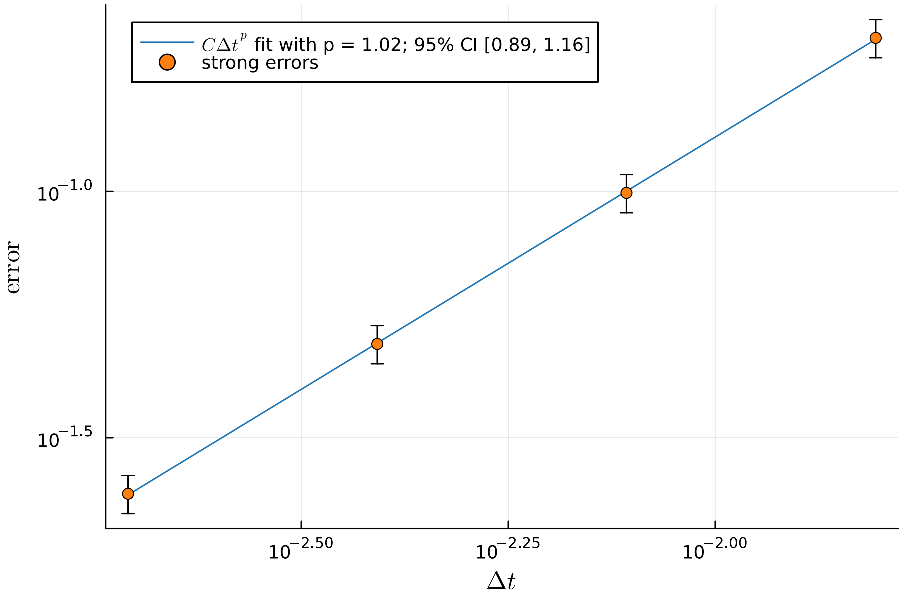

FIGURE 7. Estimated order of convergence of the strong error of the Euler method for $\mathrm{d}X_t/\mathrm{d}t = -\|Y_t\|^2 X_t + Y_t$, based on Table 1.

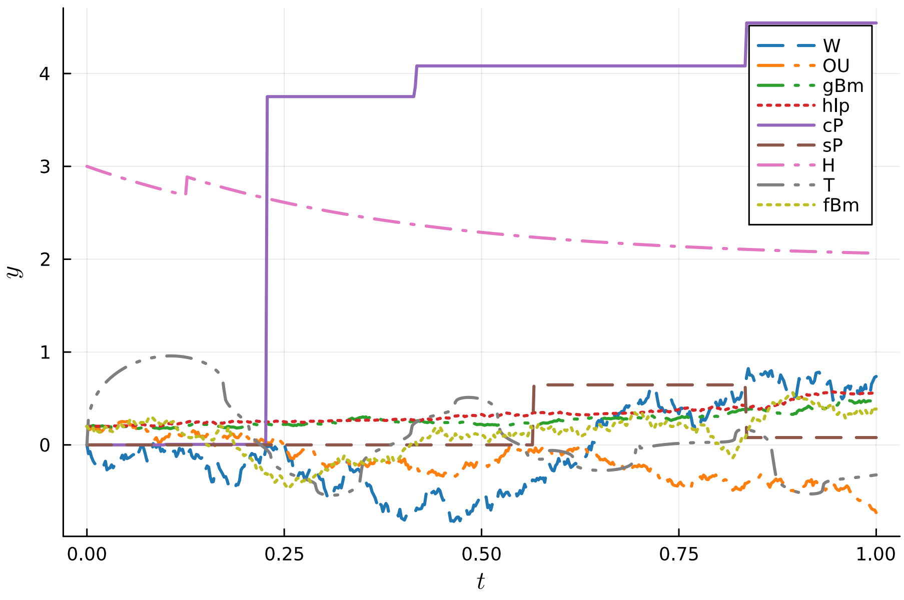

FIGURE 8. Sample paths of all the noises used in the linear system (8.2), mixing all different types of implemented noises.

Table 1. Mesh points (N), time steps (dt), strong error (error), and standard error (std err) of the Euler method for $\mathrm{d}X_t/\mathrm{d}t = -\|Y_t\|^2 X_t + Y_t$ for each mesh resolution N, with initial condition $X_0 \sim \mathcal{N}(0, \mathrm{I})$ and vector-valued noise $\{Y_t\}_t$ with all the implemented noises, on the time interval $I = [0.0, 1.0]$, based on $M = 400$ sample paths for each fixed time step, with the target solution calculated with 262 144 points. The order of strong convergence is estimated to be $p = 1.024$, with the 95% confidence interval $[0.8926, 1.1556]$.

N	dt	Error	Std err
64	0.0156	0.205	0.0093
128	0.00781	0.0993	0.00451
256	0.00391	0.049	0.00223
512	0.00195	0.0243	0.00111

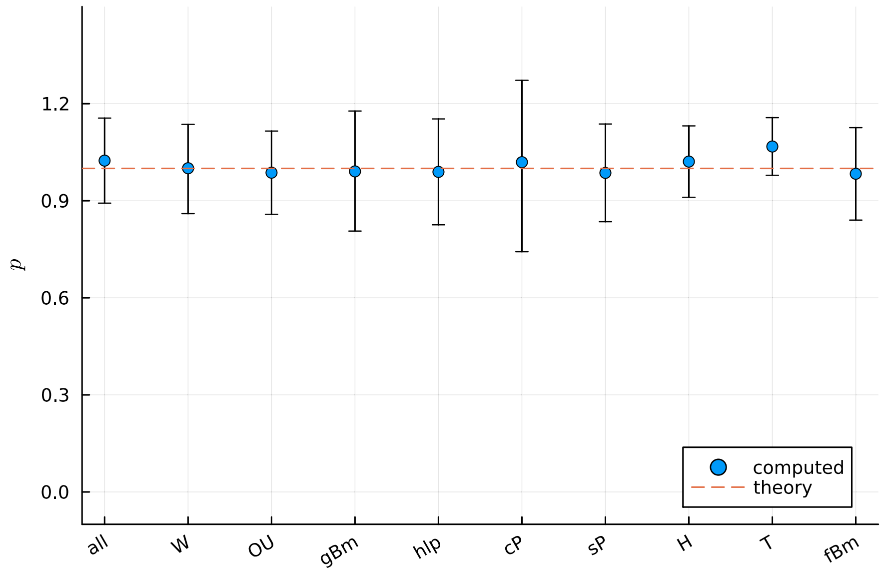

Figure 9. Estimated order of convergence for the equation (8.2) as it varies among different types of noises, starting with all the noises combined in a system of equations (all) and continuing with the scalar cases, each with a different type of noise, according to Table 2.

Table 2 and Figure 9 show the order of convergence as it varies among different types of noises, starting with all the noises combined in a system of equations and continuing with the scalar cases, each with a different type of noise, as described above.

The strong order 1 convergence is not a surprise in the case of the Wiener and Ornstein–Uhlenbeck process since the corresponding RODE can be turned into an SDE with an additive noise. In such a case, the Euler-Maruyama approximation for the noise part of the SDE is distributionally exact, and the Euler method for the RODE becomes equivalent to the Euler-Maruyama method for the SDE, and it is known that the Euler-Maruyama method for an SDE with additive noise is of strong order 1 [16]. The geometric Brownian motion is also expected to yield strong order 1 convergence thanks to the results in [35] and the fact that it is an Itô diffusion process. For the remaining noises, however, previous works would show the order of convergence to be below the order 1 attained here.

TABLE 2. Type of noise, order p of strong convergence, and 95% confidence interval endpoints $p_{\min}$ and $p_{\max}$, for the linear equation (8.2) with various types of noise.

Noise	p	$p_{\min}$	$p_{\max}$
all noises combined (all)	1.02409	0.892626	1.15557
Wiener (W)	1.0005	0.860131	1.1361
OrnsteinUhlenbeck (OU)	0.98683	0.858187	1.11547
GeometricBrownianMotion (gBm)	0.990724	0.806347	1.17738
HomogeneousLinearIto (hlp)	0.989188	0.825703	1.15296
CompoundPoisson (cP)	1.01923	0.742455	1.2723
PoissonStep (sP)	0.98615	0.835225	1.1373
ExponentialHawkes (H)	1.02111	0.910741	1.13141
Transport (T)	1.0677	0.97865	1.15692
FractionalBrownianMotion (fBm)	0.983487	0.840285	1.12595

Notice we chose the Hurst parameter of the fractional Brownian motion process to be between 1/2 and 1, so that our results for fBm indicate that the strong convergence should also be of order 1, just like for the other types of noises in $\{Y_t\}_{t\in I}$. Previous knowledge would expect a strong convergence of order H, with $1/2 < H < 1$, instead.

Acknowledgments

The authors thank Glauco Valle (IM/UFRJ), for helpful discussions about semi-martingales, Luan Lima Freitas (IM/UFRJ), for many helpful comments on the manuscript, and Ralph dos Santos Silva (IM/UFRJ) for helpful discussions about the estimates for the confidence intervals. We also like to thank the referees for a careful reading of the manuscript, with several comments and suggestions that helped improve the paper.

Funding

The second author was partly supported by Laboratório de Matemática Aplicada, Instituto de Matemática, Universidade Federal do Rio de Janeiro (LabMA/ IM/UFRJ), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Brasil, grant 001, and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Brasília, Brasil, grant no. 408751/2023-1.

Data availability statement

The research code associated with this article is available at the open-access, github repository `https://github.com/rmsrosa/rode_convergence_euler`, and detailed documentation for the repository is available at the website `https://rmsrosa.github.io/rode_convergence_euler` [22].

References

[1] Y. Aït-Sahalia and J. Jacod, Chapter 1. From Diffusions to Semimartingales, in High-Frequency Financial Econometrics. Princeton University Press (2014).

[2] C.D. Aliprantis and K.C. Border, Infinite Dimensional Analysis: A Hitchhiker's Guide. Springer, Berlin, Heidelberg (2006).

[3] Y. Asai, *Numerical methods for random ordinary differential equations and their applications in biology and medicine*. Doctoral thesis, Johann Wolfgang Goethe Universität, Frankfurt am Main (2016).

[4] Y. Asai and P.E. Kloeden, Numerical schemes for random odes with affine noise. *Numer. Algorithms* **72** (2016) 155–171.

[5] J.L. Bogdanoff, J.E. Goldberg and M.C. Bernard, Response of a simple structure to a random earthquake-type disturbance. *Bull. Seismol. Soc. Am.* **51** (1961) 293–310.

[6] D. Brigo and F. Mercurio, Interest-Rate Models: Theory and Practice. *Springer Finance*. Springer-Verlag, Berlin, Heidelberg (2001).

[7] A. Coddington and N. Levinson, Theory of Ordinary Differential Equations. *International Series in Pure and Applied Mathematics*. R.E. Krieger (1984).

[8] M.H. DeGroot and M.J. Schervish, Probability and Statistics, 4th edition. Pearson Education (2018).

[9] R.A. Fisher, The wave of advance of advantageous genes. *Ann. Eugen.* **7** (1937) 355–369.

[10] M.I. Freidlin and A.D. Wentzell, Reaction-diffusion equations with randomly perturbed boundary conditions. *Ann. Probab.* **20** (1992) 963–986.

[11] H.U. Gerber and E.S.W. Shiu, On the time value of ruin. *N. Am. Actuar. J.* **2** (1998) 48–72.

[12] L. Grüne and P.E. Kloeden, Higher order numerical schemes for affinely controlled nonlinear systems. *Numer. Math.* **89** (2001) 669–690.

[13] X. Han and P.E. Kloeden, Random Ordinary Differential Equations and Their Numerical Solution, 1st edition. Springer Publishing Company, Incorporated (2017).

[14] T. Hastie, R. Tibshirani and J. Friedman, The Elements of Statistical Learning. *Springer Series in Statistics*, 2nd edition. Springer, New York, NY (2009).

[15] S.-W. He, J.-G. Wang and J.-A. Yan, Semimartingale Theory and Stochastic Calculus. Taylor & Francis Group, New York (1992).

[16] D. Higham and P.E. Kloeden, An Introduction to the Numerical Simulation of Stochastic Differential Equations. Society for Industrial and Applied Mathematics, Philadelphia, PA (2021).

[17] G. James, D. Witten, T. Hastie and R. Tibshirani, An Introduction to Statistical Learning, 2nd edition. *Springer Texts in Statistics*. Springer, New York, NY (2021).

[18] A. Jentzen and P.E. Kloeden, Taylor Approximations for Stochastic Partial Differential Equations. Society for Industrial and Applied Mathematics (2011).

[19] R.A. Johnson and D.W. Wichern, Applied Multivariate Statistical Analysis, 6th edition. Pearson Prentice-Hall (2007).

[20] I. Karatzas and S.E. Shreve, Brownian Motion and Stochastic Calculus. *Graduate Texts in Mathematics*, 2nd edition. Springer, New York, NY (2014).

[21] P.E. Kloeden and A. Jentzen, Pathwise convergent higher order numerical schemes for random ordinary differential equations. *Proc. Roy. Soc. London Sect. A.* **463** (2007) 2229–2944.

[22] P. Kloeden and R. Rosa, Numerical examples of strong order of convergence of the euler method for random ordinary differential equations. https://github.com/rmsrosa/rode_convergence_euler and https://rmsrosa.github.io/rode_convergence_euler (2025).

[23] P.E. Kloeden, E. Platen and H. Schurz, Numerical Solution of SDE Through Computer Experiments. *Universitext*. Springer-Verlag, Berlin Heidelberg (2012).

[24] A. Kolmogorov, I. Petrovskii and N. Piscunov, A study of the equation of diffusion with increase in the quantity of matter, and its application to a biological problem. *Byul. Moskovskogo Gos. Univ.* **1** (1937) 1–25.

[25] C. Marinelli and M. Röckner, On the maximal inequalities of Burkholder, Davis and Gundy. *Expositiones Mathematicae* **34** (2016) 1–26.

[26] M. Merdan, Z. Bekiryazici, T. Kesemen and T. Khaniyev, Comparison of stochastic and random models for bacterial resistance. *Adv. Differ. Equ.* **2017** (2017) 133.

[27] M. Métivier, Semimartingales. De Gruyter, Berlin, New York (1982).

[28] T. Neckel and F. Rupp, Random Differential Equations in Scientific Computing. De Gruyter, London (2013).

[29] B. Øksendal, Stochastic Differential Equations: An Introduction with Applications. Springer-Verlag, Berlin, Heidelberg (2003).

[30] P.E. Protter, Stochastic Integration and Differential Equations, 2nd edition. Springer-Verlag, Berlin, Heidelberg (2005).

[31] R.B. Salako and W. Shen, Long time behavior of random and nonautonomous Fisher–KPP equations: Part I – Stability of equilibria and spreading speeds. *J. Dyn. Differ. Equ.* **33** (2021) 1035–1070.

[32] M. Strasser, F.J. Theis and C. Marr, Stability and multiattractor dynamics of a toggle switch based on a two-stage model of stochastic gene expression. *Biophys. J.* **102** (2012) 19–29.

[33] S. Sugiyama, Stability problems on difference and functional-differential equations. *Proc. Jpn. Acad.* **45** (1969) 526–529.

[34] B. Verd, A. Crombach and J. Jaeger, Classification of transient behaviours in a time-dependent toggle switch model. *BMC Syst. Biol.* **8** (2014) 43.

[35] P. Wang, Y. Cao, X. Han and P.E. Kloeden, Mean-square convergence of numerical methods for random ordinary differential equations. *Numer. Algorithms* **87** (2021) 299–333.