.

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

PETER E. KLOEDEN AND RICARDO M. S. ROSA

ABSTRACT. It is well known that the Euler method for a random ordinary differential equation $dX_t/dt = f(t, X_t, Y_t)$ driven by a stochastic process $\{Y_t\}_t$ with θ-Hölder sample paths is estimated to be 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.

CONTENTS

1. Introduction

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

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

on a time interval $I = [0, T]$, 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 sample space Ω with probability distribution $\mathbb{P}$. 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 and in uncertainty quantification (see e.g. [3, 9, 15, 18, 27, 30, 34, 35]). They are, in some cases, milder versions of the more popular stochastic diffusion equations (SDE), but, depending on the situation, they may actually lead to more realistic models, where, for example, the white noise is replaced by a more structured noise.

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

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

An approximation $\{X_{t_j}^N\}_j$ is said to converge to $\{X_t\}_t$ with *strong order* $\theta > 0$ when there exists a constant $C \ge 0$ such that

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

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. [17, 19, 24, 21, 2, 18]), 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 [39, Theorem 3] estimates 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, [39, Theorem 2] estimates 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 [39, Section 5], in particular [39, Section 5.2, Example 12 and Remark 13]).
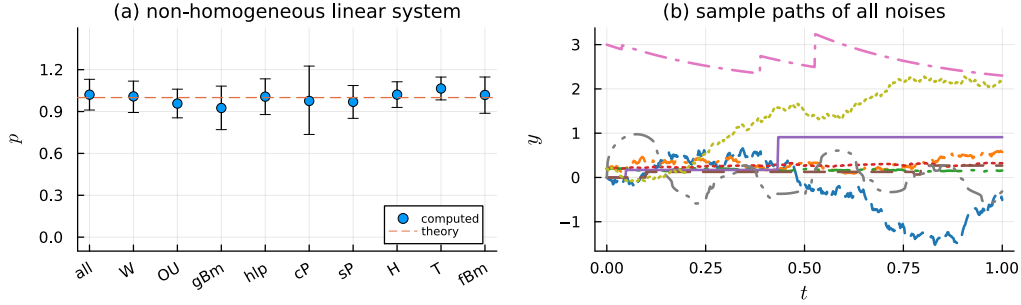

FIGURE 1. (a) Estimated order of convergence of the strong error of the Euler method for the linear system (1.4) as it depends on the noise, starting with all the noises combined (all) and continuing with the scalar equation, each with a different type of noise process, namely Wiener (W), Ornstein-Uhlenbeck (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).

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 Theorem 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. The technique may also yield improvements in the results concerning higher order methods and or other types of noises, although this requires a more thorough investigation.
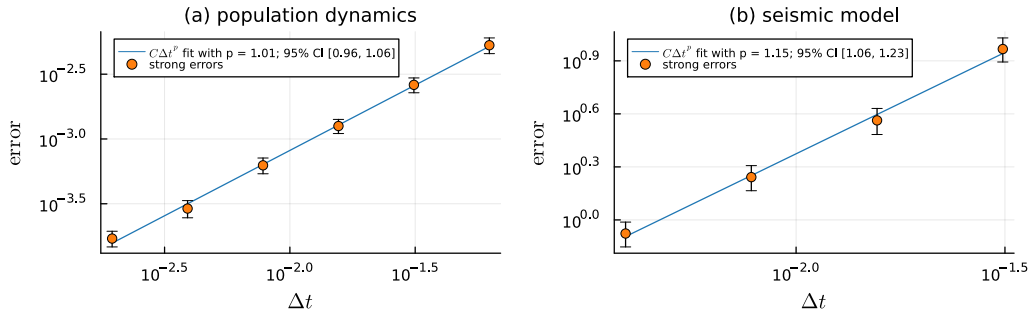
FIGURE 2. Estimated order of convergence of the strong error of the Euler method for (a) the population dynamics model (1.5) and (b) the seismic model (1.7).

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 $\{X_t\}_t$ of a sample path.

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. The examples are presented in more details in the supplementary material in Section 8 and in the github repository [23] and include the following:

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

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

where $\{Y_t\}_t$ is either (i) a univariate process of one of many types of semi-martingale processes or (ii) a vector-valued process combining all these noises as independent components.

(ii) A logistic model of population dynamics with random coefficients, loosely inspired by [18, Section 15.2], where the specific growth is the sine of a geometric Brownian motion process $\{G_t\}_t$ and with an extra point-process random term $\{H_t\}_t$ representing harvest:

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

(iii) A toggle-switch model of gene expression (similar to [3, Section 7.8], originated from [38], see also [35]) driven by a combination of a compound Poisson point process $\{A_t\}_t$ and an Itô diffusion process $\{B_t\}_t$,

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

(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 [9].

(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 [16] and [10], where $C_t = \sum_{i=1}^{N_t} C_i$ defines a compound Poisson process modeling the insurance claims, $\{R_t\}_t$ is a geometric Brownian motion process modeling the interest rate, and $\{O_t\}_t$ 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 [34] and [15] (see also [14] and [25]), and where the noise $\{Y_t\}_t$ 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 $\{B_t^H\}_t$ 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 $H \neq 1/2$.

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 $\{U_t\}_t$, the Ornstein-Uhlenbeck process $\{O_t\}_t$, the interest rate $\{I_t\}_t$, and the insurance claims $\{C_t\}_t$.

FIGURE 5. (a) Estimated order of convergence of the strong error of the Euler method for the Fisher-KPP model Equations (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.

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

FIGURE 6. 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.

The starting point of the proof is to not estimate the local error and, instead, work with an explicit formula for the global error (see Lemma 5.1), similarly to the way it is done for approximations of stochastic differential equations. In this way, we obtain the estimate (see (6.4))

$$\|X_{t_j} - X_{t_j}^N\| \le \left(\|X_0 - X_0^N\| + c_L \int_0^{t_j} \|X_s - X_{\tau^N(s)}\| \, \mathrm{d}s \right.$$
$$\left. \left\| \int_0^{t_j} \left(f(s, X_{\tau^N(s)}^N, Y_s) - f(\tau^N(s), X_{\tau^N(s)}^N, Y_{\tau^N(s)}) \right) \, \mathrm{d}s \right\| \right) e^{c_L t_j}, \quad (1.12)$$

for $j = 1, \ldots, N$, where $\tau^N(t) = \max_{t_j \le t} t_j$ 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 estimate the local strong error as $C\Delta t_N^\theta$. 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 being from ξ to the "future" time $\tau^N(\xi) + \Delta t_N$, the integrand is still non-anticipative, hence

$$\int_0^{t_j} \left(f(s, X^N_{\tau^N(s)}, Y_s) - f(\tau^N(s), X^N_{\tau^N(s)}, Y_{\tau^N(s)}) \right) \, \mathrm{d}s = \int_0^{t_j} \int_{\tau^N(s)}^s \mathrm{d}F_\xi \, \mathrm{d}s$$

$$= \int_0^{t_j} \int_\xi^{\tau^N(\xi) + \Delta t_N} \mathrm{d}s \, \mathrm{d}F_\xi = \int_0^{t_j} (\tau^N(\xi) + \Delta t_N - \xi) \, \mathrm{d}F_\xi.$$

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

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

$$\leq \Delta t_N \mathbb{E}\left[\left\| \int_0^{t_j} \mathrm{d}F_\xi \right\| \right],$$

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 above estimate does not apply. Instead, we essentially have (see (8.12) and (8.14))

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

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

$$\int_0^{t_j} \left(f(s, X^N_{\tau^N(s)}, Y_s) - f(\tau^N(s), X^N_{\tau^N(s)}, Y_{\tau^N(s)}) \right) \, \mathrm{d}s$$

$$\sim \int_0^{t_j} \int_{\tau^N(s)}^s (s - \tau^N(s))^{H - 1/2} \, \mathrm{d}W_\xi \, \mathrm{d}s \sim (\Delta t_N)^{H + 1/2} \int_0^{t_j} \mathrm{d}W_\xi.$$

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 [18, Section 2.1] and [30, Section 3.3]. 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:

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

   (i) *The function $f = f(t, x, y)$ is continous on $(t, x, y) \in I \times \mathbb{R}^d \times \mathbb{R}^k$.*

   (ii) *The mapping $x \mapsto f(t, x, y)$ is globally Lipschitz continuous on x, uniformly on $t \in I$ and locally on $y \in \mathbb{R}^k$, 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 $\{Y_t\}_{t \geq 0}$ 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 [12, Theorem 1.1]). Each pathwise solution of (2.2) is absolutely continuous. Moreover, the mapping $(t, \omega) \mapsto X_t(\omega)$ is (jointly) measurable (see [18, Section 2.1.2] and also [1, Lemma 4.51]) and, hence, give rise to a well-defined stochastic process $\{X_t\}_{t \in I}$.

Remark 2.1. Since $\{Y_t\}_{t \geq 0}$ is a semi-martingale, the sample paths $t \mapsto Y_t(\omega)$ are almost surely continuous except for a set of measure zero (see Section 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 Section 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, 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

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

almost surely. Using Grownwall'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 estimate on its own and is useful for 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.

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

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

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), (2.5) and (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. Indeed, when some bounded, positively invariant region exists and is of interest, we may truncate the nonlinear term to achieve the desired global conditions for the equation with the truncated term, but which coincides with the original equation in the region of interest. 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. We leave these cases to be handled in the applications.

3. Bound on the pathwise approximations

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

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

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

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

Iterating this relation we obtain the following bound.

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

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

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

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 moreover that (2.4), (2.5) and (2.6) hold. Then, the Euler approximation (1.2) for a pathwise solution of the random ordinary differential equation (1.1) satisfies*

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

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

4. Semi-martingale noises

For the theory of semi-martingales, we refer the reader to [32] and [28]. The main idea is that *semi-martingales* are stochastic processes $\{Y_t\}_{t \geq 0}$ which are càdlàg, adapted, and "good integrators," in the sense that the sums

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

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 finitely-valued random variables, where $0 = T_1 < \ldots < T_n$, 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 $\{Y_t\}_{t\geq 0}$ whose sample paths are almost surely càdlàg paths. Thus, $\{Y_t\}_{t\geq 0}$ is càdlàg process when for almost every $\omega \in \Omega$ and for all $t \geq 0$, 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* $\{[Y,Y]_t\}_{t\geq 0}$ defined as the limit of the summations of quadratic increments as the mesh is refined. In the scalar case, this is given by

$$[Y,Y]_t = \lim_{\substack{0\leq t_0<\ldots<t_n\leq t,\, n\in\mathbb{N} \\ \max_{j=1,\ldots,n}|t_j-t_{j-1}|\to 0}} \sum_{j=1}^n (Y_{t_j} - Y_{t_{j-1}})^2,$$

or, equivalently, by the formula (which is a particular case of an "integration by parts" formula for the quadratic covariation process between two semi-martingales)

$$[Y,Y]_t = Y_t^2 - 2\int_0^t Y_{s^-}\, \mathrm{d}Y_s.$$

In the vector-valued case, denoting $Y_t = (Y_t^{(l)})_{l=1,\ldots,k}$, 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 appearence of the quadratic cross-variations

$$[Y^{(l_1)},Y^{(l_2)}] = \lim_{\substack{0\leq t_0<\ldots<t_n\leq t,\, n\in\mathbb{N} \\ \max_{j=1,\ldots,n}|t_j-t_{j-1}|\to 0}} \sum_{j=1}^n (Y_{t_j}^{(l_1)} - Y_{t_{j-1}}^{(l_1)})(Y_{t_j}^{(l_2)} - Y_{t_{j-1}}^{(l_2)}).$$

The quadratic process can be written in matrix form as

$$[Y,Y]_t = \lim_{\substack{0 \leq t_0 < \ldots < t_n \leq t,\, n \in \mathbb{N} \\ \max_{j=1,\ldots,n} |t_j - t_{j-1}| \to 0}} \sum_{j=1}^{n} (Y_{t_j} - Y_{t_{j-1}})(Y_{t_j} - Y_{t_{j-1}})^{\mathrm{tr}},$$

or, equivalently,

$$[Y,Y]_t = Y_t Y_t^{\mathrm{tr}} - 2 \int_0^t Y_{s^-} \, \mathrm{d}Y_s^{\mathrm{tr}},$$

where Y_t^{tr} is the transpose of Y_t.

The quadratic variation is a càdlàg, increasing, and adapted process [32, Theorem II.22]. The *continuous part* $\{[Y,Y]_t^c\}_{t \geq 0}$ of the quadratic variation is defined by

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

which itself is a semi-martingale of finite-variation type. 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 polarization identity

$$[Y^{(l_1)}, Y^{(l_2)}] = \frac{1}{4} \left([Y^{(l_1)} + Y^{(l_2)}, Y^{(l_1)} + Y^{(l_2)}] - [Y^{(l_1)} - Y^{(l_2)}, Y^{(l_1)} - Y^{(l_2)}] \right)$$

implies that each quadratic cross-variation is a difference of quadratic variations, which are non-decreasing processes, so that each quadratic cross-variation is of finite variation.

The Bichteler-Dellacherie Theorem [32, Theorem III.47] yields a characterization of a semi-martingale as a càdlàg process $\{Y_t\}_{t \geq 0}$ decomposable as a sum of a finite-variation (FV) process and a local martingale process, which we can write more precisely as $Y_t = Y_0 + F_t + Z_t$, with an FV process $\{F_t\}_{t \geq 0}$ and a local martingale process $\{Z_t\}_{t \geq 0}$ with $F_0 = Z_0 = 0$.

A *Finite-Variation process* (*FV process*) is a càdlàg process $\{F_t\}_{t \geq 0}$ whose sample paths are almost surely of finite variation, $V(\{F_t\}_t; K)$, on compact subsets $K \subset [0, \infty)$, i.e.

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

The class of FV processes include 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 $\{Z_t\}_{t \geq 0}$ such that there exists a sequence of increasing stopping times with $T_n \to \infty$ almost surely, for which $\{Z_{t \wedge T_n}\}_{t \geq 0}$

is a uniformly integrable martingale, for each $n \in \mathbb{N}$, i.e.

$$\sup_{t \geq 0} \int_{T_n > 0} \|Z_{t \wedge T_n}\| \, d\mathbb{P} < \infty.$$

The class of local martingales includes time-changed Brownian motions and Itô diffusion processes. For more on Itô diffusion, see also [31].

With the decomposition of a semi-martingale in the form $Y_t = Y_0 + F_t + Z_t$, for an FV process $\{F_t\}_{t \geq 0}$ and a local martingale process $\{Z_t\}_{t \geq 0}$ with $F_0 = Z_0 = 0$, we recall from [28, Lemma 18.7] 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 $\{F_t\}_{t \geq 0}$, 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

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

We also have

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

In the more general case of semi-martingale, the process $\{Y_t\}_{t \geq 0}$ has finite quadratic variation [32, Section 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 [32, Theorem II.22]),

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

Since we are considering a bounded interval $I = [0, T]$, the local martingale $Z_t = Z_{t \wedge T}$, $0 \leq t \leq T$, where $t \wedge T = \min\{t, T\}$, is a square-integrable martingale, and the Doob Maximal inequality (see e.g. [32, Theorem I.20]) yields

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

which is finite by the definition of martingale. Since $F_0 = 0$, the finite variation part is bounded by

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

which implies, in particular,

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

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

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

and

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

For the quadratic variation part of a local martingale $\{Z_t\}_{t\in I}$, the Burkholder-David-Gundy inequalities [32, Theorem IV.48] relate the moments of the maximum of the local martingale with moments of its quadratic variation,

$$\frac{1}{c_p}\mathbb{E}\left[\|[Z, Z]_t\|^{p/2}\right] \leq \mathbb{E}\left[\max_{0\leq s\leq t}\|Z_t\|^p\right] \leq C_p\mathbb{E}\left[\|[Z, Z]_t\|^{p/2}\right],$$

for $p \geq 1$ and suitable constants $c_p, C_p > 0$. Thus, we obtain the bound

$$\mathbb{E}\left[\|[Z, Z]_t\|\right] \leq 4c_2\mathbb{E}\left[\|Z_T\|^2\right] < \infty,$$

for any $t \in I$, and then

$$\mathbb{E}\left[\|[Y, Y]_t\|\right] \leq 4c_2\mathbb{E}\left[\|Z_T\|^2\right] + \mathbb{E}\left[\|V(\{F_t\}_t; I)\|^2\right]. \tag{4.6}$$

This, together with (4.2), also yields

$$\mathbb{E}\left[\sum_{s\in J_Y}\|\Delta Y_t\|^2\right] \leq 4c_2\mathbb{E}\left[\|Z_T\|^2\right] + \mathbb{E}\left[\|V(\{F_t\}_t; I)\|^2\right]. \tag{4.7}$$

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 [32, Theorems II.31 and II.33]):

$$f(Y_t) - f(Y_0) = \int_{0^+}^t Df(Y_{s-})\,\mathrm{d}Y_s + \sum_{0<s\leq t}\left(f(Y_s) - f(Y_{s-}) - Df(Y_{s-})\Delta Y_s\right). \tag{4.8}$$

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 (4.10) 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 $\{(t, Y_t)\}_{t \geq 0}$. The first component is continuous hence has no jumps and the corresponding summation term vanishes. This leads to the following change of variables formula for FV processes:

$$f(t, Y_t) - f(0, Y_0) = \int_0^t \partial_s f(s, Y_{s-}) \, \mathrm{d}s + \int_{0+}^t D_y f(s, Y_{s-}) \, \mathrm{d}Y_s$$
$$+ \sum_{0 < s \leq t} (f(s, Y_s) - f(s, Y_{s-}) - D_y f(s, Y_{s-}) \Delta Y_s). \quad (4.9)$$

In the more general case of a semi-martingale noise $\{Y_t\}_{t \geq 0}$, the change of variables formula (4.9) takes the form (see [32, Theorems II.32 and II.33])

$$f(t, Y_t) - f(0, Y_0) = \int_0^t \partial_s f(s, Y_{s-}) \, \mathrm{d}s + \int_{0+}^t D_y f(s, Y_{s-}) \, \mathrm{d}Y_s$$
$$+ \sum_{0 < s \leq t} (f(s, Y_s) - f(s, Y_{s-}) - D_y f(s, Y_{s-}) \Delta Y_s) \quad (4.10)$$
$$+ \frac{1}{2} \int_0^t D_{yy} f(s, Y_{s-}) \, \mathrm{d}[Y, Y]_s^c,$$

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

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

Since $[Y, Y]_t^c$ is of finite variation, the last integral is a Lebesgue-Stieltjes integral.

We actually apply the formula (4.10) 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 $\{Z_t\}_{t \geq 0}$ and an adapted matrix-valued process $\{A_t\}_{t \geq 0}$, takes the form

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

provided the right hand side is finite, and where $\odot$ denotes the Hadamard product of two matrices of same dimensions.

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*

$$X_{t_j} - X_{t_j}^N = X_0 - X_0^N + \int_0^{t_j} \left(f(s, X_s, Y_s) - f(s, X_{\tau^N(s)}, Y_s) \right) \, \mathrm{d}s$$

$$+ \int_0^{t_j} \left(f(s, X_{\tau^N(s)}, Y_s) - f(s, X_{\tau^N(s)}^N, Y_s) \right) \, \mathrm{d}s \qquad (5.1)$$

$$+ \int_0^{t_j} \left(f(s, X_{\tau^N(s)}^N, Y_s) - f(\tau^N(s), X_{\tau^N(s)}^N, Y_{\tau^N(s)}) \right) \, \mathrm{d}s,$$

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:

$$\tau^N(t) = \max_{t_j \le t}\{t_j\} = \left[\frac{t}{\Delta t_N} \right] \Delta t_N = \left[\frac{tN}{T} \right] \frac{T}{N}. \qquad (5.2)$$

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

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

Adding and subtracting appropriate terms yield

$$X_{t_j} - X_{t_j}^N = X_{t_{j-1}} - X_{t_{j-1}}^N$$

$$+ \int_{t_{j-1}}^{t_j} \left(f(s, X_s, Y_s) - f(s, X_{t_{j-1}}, Y_s) \right) \, \mathrm{d}s$$

$$+ \int_{t_{j-1}}^{t_j} \left(f(s, X_{t_{j-1}}, Y_s) - f(s, X_{t_{j-1}}^N, Y_s) \right) \, \mathrm{d}s \qquad (5.3)$$

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

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

$$X_{t_j} - X_{t_j}^N = X_0 - X_0^N + \sum_{i=1}^{j} \left(\int_{t_{i-1}}^{t_i} (f(s, X_s, Y_s) - f(s, X_{t_i}, Y_s)) \, \mathrm{d}s \right.$$

$$+ \int_{t_{i-1}}^{t_i} \left(f(s, X_{t_{i-1}}, Y_s) - f(s, X_{t_{i-1}}^N, Y_s) \right) \, \mathrm{d}s$$

$$+ \left. \int_{t_{i-1}}^{t_i} \left(f(s, X_{t_{i-1}}^N, Y_s) - f(t_{i-1}, X_{t_{i-1}}^N, Y_{t_{i-1}}) \right) \, \mathrm{d}s \right).$$

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

6. BASIC ESTIMATES FOR THE GLOBAL ERROR

Here we derive an estimate which is the basis for the convergence estimate. For that, we use the following classical discrete version of the Grownwall Lemma, which can be deduce from e.g. [36, Lemma 1].

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*

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

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

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

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

Proposition 6.1 (Basic strong estimate). *Under the Standing Hypothesis 2.1, suppose further that (2.6) holds. Then, for every $j = 0, \dots, N$,*

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

Proof. We estimate 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,

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

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

$$\left\| \int_0^{t_j} \left(f(s, X_{\tau^N(s)}, Y_s) - f(s, X_{\tau^N(s)}^N, Y_s) \right) \, \mathrm{d}s \right\| \leq c_L \int_0^{t_j} \|X_{\tau^N(s)} - X_{\tau^N(s)}^N\| \, \mathrm{d}s$$

$$\leq c_L \sum_{i=0}^{j-1} \|X_{t_i} - X_{t_i}^N\| \Delta t.$$

With these two estimates, we bound (5.1) as

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

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

$$\|X_{t_j} - X_{t_j}^N\| \le \left(\|X_0 - X_0^N\| + c_L \int_0^{t_j} \|X_s - X_{\tau^N(s)}\| \, \mathrm{d}s \right.$$
$$\left. + \left\| \int_0^{t_j} \left(f(s, X_{\tau^N(s)}^N, Y_s) - f(\tau^N(s), X_{\tau^N(s)}^N, Y_{\tau^N(s)}) \right) \, \mathrm{d}s \right\| \right) e^{c_L t_j}, \quad (6.4)$$

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). $\qquad\square$

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 and assuming conditions (2.4), (2.5) and (2.6), we find that, for $j = 1, \ldots, N$,*

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

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

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

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

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

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

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

Thanks to (2.4), (2.5) and (2.6), Lemma 2.2 applies and inequality (2.8) holds, so that we obtain

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

This simplifies to (6.5). $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

7. Strong convergence for semi-martingale noises

We now prove the desired strong order 1 convergence.

Theorem 7.1. *Under the Standing Hypothesis 2.1, suppose further that (2.4), (2.5) and (2.6) hold; that X_0^N is a first order approximation of X_0, in the sense that*

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

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, $\{Z_t\}_{t \in I}$ is a local martingale, and $F_0 = Z_0 = 0$, and assume they satisfy

$$\mathbb{E}\left[\|| Y_0 \|| \right] < \infty, \tag{7.5}$$

$$\mathbb{E}\left[V(\{F_t\}_t; I) \right] < \infty. \tag{7.6}$$

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

$$\max_{j=0,\ldots,N} \mathbb{E}\left[\left\|X_{t_j} - X_{t_j}^N\right\|\right] \leq c\Delta t_N, \qquad \forall N \in \mathbb{N}, \tag{7.7}$$

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 (4.10), the integral in this last term reads

$$\int_0^{t_j} \left(f(s, X_{\tau^N(s)}^N, Y_s) - f(\tau^N(s), X_{\tau^N(s)}^N, Y_{\tau^N(s)}) \right) \, \mathrm{d}s$$

$$= \int_0^{t_j} \int_{\tau^N(s)}^s \partial_\xi f(\xi, X_{\tau^N(s)}^N, Y_{\xi^-}) \, \mathrm{d}\xi \, \mathrm{d}s$$

$$+ \int_0^{t_j} \int_{\tau^N(s)+}^s D_y f(\xi, X_{\tau^N(s)}^N, Y_{\xi^-}) \, \mathrm{d}Y_\xi \, \mathrm{d}s$$

$$+ \int_0^{t_j} \sum_{\tau^N(s) < \xi \leq s} \left(f(\xi, X_{\tau^N(s)}^N, Y_\xi) - f(\xi, X_{\tau^N(s)}^N, Y_{\xi^-}) \right. \tag{7.8}$$

$$\left. - D_y f(\xi, X_{\tau^N(s)}^N, Y_{\xi^-}) \Delta Y_\xi \right) \, \mathrm{d}s$$

$$+ \frac{1}{2} \int_0^{t_j} \int_{\tau^N(s)}^s D_{yy} f(\xi, X_{\tau^N(s)}^N, Y_{\xi^-}) \, \mathrm{d}[Y,Y]_\xi^c \, \mathrm{d}s.$$

The expectation of the first term in (7.8) is estimated using (7.2), so that

$$\mathbb{E}\left[\left\| \int_0^{t_j} \int_{\tau^N(s)}^s \partial_\xi f(\xi, X_{\tau^N(s)}^N, Y_{\xi^-}) \, \mathrm{d}\xi \, \mathrm{d}s \right\|\right]$$

$$\leq \mathbb{E}\left[\int_0^{t_j} \int_{\tau^N(s)}^s \left(c_1 + c_2 \|X_{\tau^N(s)}^N\| + c_3 \|Y_{\xi^-}\| \right) \, \mathrm{d}\xi \, \mathrm{d}s \right] \tag{7.9}$$

$$\leq t_j \Delta t_N \left(c_1 + c_2 \sup_{0 \leq s \leq t_j} \mathbb{E}\left[\|X_{\tau^N(s)}^N\| \right] + c_3 \sup_{0 \leq s \leq t_j} \mathbb{E}\left[\|Y_s\| \right] \right).$$

For the second term in (7.8), we use the decomposition $Y_t = Y_0 + F_t + Z_t$. With that, the second term can be written as

$$\int_0^{t_j} \int_{\tau^N(s)+}^s D_y f(\xi, X_{\tau^N(s)}^N, Y_{\xi^-}) \, \mathrm{d}Y_\xi \, \mathrm{d}s = \int_0^{t_j} \int_{\tau^N(s)+}^s D_y f(\xi, X_{\tau^N(s)}^N, Y_{\xi^-}) \, \mathrm{d}F_\xi \, \mathrm{d}s$$

$$+ \int_0^{t_j} \int_{\tau^N(s)+}^s D_y f(\xi, X_{\tau^N(s)}^N, Y_{\xi^-}) \, \mathrm{d}Z_\xi \, \mathrm{d}s, \tag{7.10}$$

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 first integral on the right hand side of (7.10) is estimated using (7.3), so that

$$\left\| \int_0^{t_j} \int_{\tau^N(s)+}^s D_y f(\xi, X_{\tau^N(s)}^N, Y_{\xi-}) \, dF_\xi \, ds \right\| \leq \int_0^{t_j} \int_{\tau^N(s)}^s c_4 \, \|dF_\xi\| \, ds$$

$$\leq c_4 \int_0^{t_j} \int_\xi^{\tau^N(\xi)+\Delta t_N} ds \, \|dF_\xi\| \leq c_4 \Delta t_N \int_0^{t_j} \|dF_\xi\| \leq c_4 \Delta t_N V(\{F_t\}_t; [0, t_j]).$$

Taking the expectation and using (7.6), we obtain

$$\mathbb{E}\left[\left\| \int_0^{t_j} \int_{\tau^N(s)+}^s D_y f(\xi, X_{\tau^N(s)}^N, Y_{\xi-}) \, dF_\xi \, ds \right\| \right] \leq c_4 \Delta t_N \mathbb{E}\left[V(\{F_t\}_t; [0, t_j]) \right]. \quad (7.11)$$

For the second integral on the right hand side of (7.10), we once again switch the order of integration, then use the Lyapunov inequality, the Itô isometry for local martingales, and then again the bound (7.3), so that

$$\mathbb{E}\left[\left\| \int_0^{t_j} \int_{\tau^N(s)+}^s D_y f(\xi, X_{\tau^N(s)}^N, Y_{\xi-}) \, dZ_\xi \, ds \right\| \right]$$

$$= \mathbb{E}\left[\left\| \int_0^{t_j} \int_\xi^{\tau^N(\xi)+\Delta t_N} D_y f(\xi, X_{\tau^N(s)}^N, Y_{\xi-}) \, ds \, dZ_\xi \right\| \right]$$

$$\leq \mathbb{E}\left[\left\| \int_0^{t_j} \int_\xi^{\tau^N(\xi)+\Delta t_N} D_y f(\xi, X_{\tau^N(\xi)}^N, Y_{\xi-}) \, ds \, dZ_\xi \right\|^2 \right]^{1/2}$$

$$\leq c_4 \Delta t_N \mathbb{E}\left[\sum_{l_1, l_2=1}^k \int_0^{t_j} d[Z^{(l_1)}, Z^{(l_2)}]_\xi \right]^{1/2}$$

$$= c_4 \Delta t_N \mathbb{E}\left[\sum_{l_1, l_2=1}^k [Z^{(l_1)}, Z^{(l_2)}]_{t_j} - [Z^{(l_1)}, Z^{(l_2)}]_0 \right]^{1/2},$$

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 $X_{\tau^N(s)}^N = X_{\tau^N(\xi)}^N$. Now, using that $[Z, Z]_0 = Z_0^2 = 0$, we obtain

$$\mathbb{E}\left[\left\| \int_0^{t_j} \int_{\tau^N(s)+}^s D_y f(\xi, X_{\tau^N(s)}^N, Y_{\xi-}) \, dZ_\xi \, ds \right\| \right] \leq c_4 \sqrt{k} \Delta t_N \mathbb{E}\left[\|[Z, Z]_{t_j}\| \right]^{1/2}. \quad (7.12)$$

Now we move to the third term in (7.8). We use the bound (7.4) on the second derivative of $D_{yy} f$ to write

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

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,

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

$$\leq \frac{c_7}{2} \int_0^{t_j} \sum_{\tau^N(s) < \xi \leq s} \|\Delta Y_\xi\|^2 \, ds$$

$$\leq \frac{c_7}{2} \sum_{0 < \xi \leq t_j} \int_\xi^{\tau^N(\xi) + \Delta t_N} \|\Delta Y_\xi\|^2 \, ds$$

$$\leq \frac{c_7}{2} \Delta t_N \sum_{0 < \xi \leq t_j} \|\Delta Y_\xi\|^2.$$

Taking the expectation,

$$\mathbb{E}\left[\left\| \int_0^{t_j} \sum_{\tau^N(s) < \xi \leq s} \left(f(\xi, X^N_{\tau^N(s)}, Y_\xi) - f(\xi, X^N_{\tau^N(s)}, Y_{\xi-}) \right. \right. \right.$$

$$\left. \left. \left. - D_y f(\xi, X^N_{\tau^N(s)}, Y_{\xi-}) \Delta Y_\xi \right) ds \right\| \right] \leq \frac{c_7}{2} \Delta t_N \mathbb{E}\left[\left\| \sum_{0 < \xi \leq t_j} \|\Delta Y_\xi\|^2 \right\| \right]. \quad (7.13)$$

The final term in (7.8) involves a Lebesgue-Stieltjes integral and is handled in a similar way, so that

$$\left\| \frac{1}{2} \int_0^{t_j} \int_{\tau^N(s)}^s D_{yy} f(\xi, X^N_{\tau^N(s)}, Y_{\xi-}) \, d[Y,Y]_\xi^c \, ds \right\|$$

$$\leq \frac{1}{2} \int_0^{t_j} \int_\xi^{\tau^N(\xi) + \Delta t_N} \left\| D_{yy} f(\xi, X^N_{\tau^N(s)}, Y_{\xi-}) \right\| \, ds \, d[Y,Y]_\xi^c$$

$$\leq \frac{c_7}{2} \Delta t_N \int_0^{t_j} d[Y,Y]_\xi^c \leq \frac{c_7}{2} \Delta t_N [Y,Y]_{t_j}^c,$$

where we used that $[Y,Y]_0^c = 0$. Taking the expectation,

$$\mathbb{E}\left[\left\| \frac{1}{2} \int_0^{t_j} \int_{\tau^N(s)}^s D_{yy} f(\xi, X^N_{\tau^N(s)}, Y_{\xi-}) \, d[Y,Y]_\xi^c \, ds \right\| \right] \leq \frac{c_7}{2} \Delta t_N \mathbb{E}\left[[Y,Y]_{t_j}^c \right]. \quad (7.14)$$

Putting the estimates (7.9), (7.10), (7.11), (7.12), (7.13), and (7.14) together and using that $0 \leq [Y,Y]_t^c \leq [Y,Y]_t$ are increasing, we obtain, from (7.8),

$$\mathbb{E}\left[\left\| \int_0^{t_j} \left(f(s, X^N_{\tau^N(s)}, Y_s) - f(\tau^N(s), X^N_{\tau^N(s)}, Y_{\tau^N(s)}) \right) ds \right\| \right] \leq C_T \Delta t_N, \quad (7.15)$$

where

$$C_T = c_1 T + c_2 T \sup_{0 \leq s \leq T} \mathbb{E}\left[\|X^N_{\tau^N(s)}\|\right] + c_3 T \sup_{0 \leq s \leq T} \mathbb{E}\left[\|Y_s\|\right] + c_4 \mathbb{E}\left[V(\{F_t\}_t; I)\right]$$

$$+ c_4 \mathbb{E}\left[[Z,Z]_T\right]^{1/2} + \frac{c_7}{2}\mathbb{E}\left[\sum_{0 < s \leq T} \|\Delta Y_s\|^2\right] + \frac{c_7}{2}\mathbb{E}\left[[Y,Y]_T\right].$$

Thanks to (2.4), (2.5), (2.6), (3.3), (7.5), (7.6), (4.4), (4.3), (4.6), and (4.7), the constant C_T above is finite and independent of N. Plugging (7.15) into (6.3) and applying (6.5) and (7.1) yield the desired strong order 1 estimate (7.7). □

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. Supplementary Material

In this supplementary material, we describe in details the numerical examples illustrating the main work. In all the examples driven by semi-martingale noises, we numerically verify the strong order 1 convergence of the Euler method. In the example with fractional Brownian motion, we verify the strong convergence of order $\min\{H + 1/2, 1\}$, where H is the Hurst parameter. The order of convergence is displayed with a 95% confidence interval that perfectly matches the theory developed in the paper. They also show that the theoretical estimate is sharp.

The code for all the examples is written in the Julia Programming Language [6] and is available in the github repository https://github.com/rmsrosa/rode_convergence_euler/ [23]. The Julia language itself is available for download on the website https://julialang.org/.

One of the distinguishing powers of the Julia language is to work both as a high-level, interactive, scripting language (with its dynamically typed and just-in-time compilation system), and as a fast, high-performance language (with its ahead-of-time and aggressive compilation capabilities). It allows for both easy prototyping and the production of efficient end code. As such, it is growing quickly as a language suitable for scientific computing and computationally-demanding real-world applications.

Our code has been tested in the currently most recent Julia version 1.11 as well as in previous versions down to version 1.8. The code is set up in a way that any change to the repository triggers an automatic rerun of all the documentation code and all the unit tests in a github server, to make sure it is always up to date. The

result of the github documentation run is displayed online, on the website `https://rmsrosa.github.io/rode_conv_em/`.

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 $i = 1,\ldots,I$, 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.

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 $n_i \in \mathbb{N}$, $i = 1,\ldots,I$, so that finer meshes are refinements of coarser meshes.
(v) A number $N_{\mathrm{tgt}} = 2^{n_{\mathrm{tgt}}}$ for a finer resolution to compute a target solution path, typically $N_{\mathrm{tgt}} = \max_i\{N_i^2\}$, 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, $\{Y_{t_j}\}_{j=0,\ldots,N_{\mathrm{tgt}}}$, on the finest grid $t_j = j\Delta t_{N_{\mathrm{tgt}}}$, $j = 0,\ldots,N_{\mathrm{tgt}}$, 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 $\{X_{t_j}\}_{j=0,\ldots,N_{\mathrm{tgt}}}$, 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,

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

(vi) We fit $(\Delta t_{N_i}, \epsilon^{N_i})$ to the power law $C\Delta t^p_{N_i}$, 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 $(A^t A)\mathbf{v} = A^t \ln(\boldsymbol{\epsilon})$, 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,

$$s^{N_i}_{t_j} = \frac{\sigma^{N_i}_{t_j}}{\sqrt{M}},$$

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

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

and compute the 95% confidence interval $[\epsilon_{\min}, \epsilon_{\max}]$ for the strong error with

$$\epsilon^{N_i}_{\min} = \max_{j=0,\ldots,N_i} (\epsilon^{N_i}_{t_j} - 2\sigma^{N_i}_{t_j}), \quad \epsilon^{N_i}_{\max} = \max_{j=0,\ldots,N_i} (\epsilon^{N_i}_{t_j} + 2\sigma^{N_i}_{t_j}).$$

(viii) Finally, from the normal equations above, we compute the 95% confidence interval $[p_{\min}, p_{\max}]$ by computing the minimum and maximum values of p in the image, by the linear map $\mathbf{e} \mapsto (\ln(C), p) = (A^t A)^{-1} A^t \ln(\mathbf{e})$, of the polytope formed by all combinations of the indices in $(\epsilon^{N_i}_{\min})_i$ and $(\epsilon^{N_i}_{\max})_i$, which is the image of the set of 95% confidence intervals for the errors obtained with the Monte-Carlo approximation of the strong error.

8.2. **Numerical setup.** Julia has a feature-rich `DifferentialEquations.jl` ecosystem of packages for solving differential equations [33], including random and stochastic differential equations, as well as delay equations, differential-algebraic equations, jump diffusions, partial differential equations, neural differential equations, and so on. It also has packages to seemlessly compose such equations in optimization problems, Bayesian parameter estimation, global sensitivity analysis, uncertainty quantification, and domain specific applications.

Although all the source code for `DifferentialEquations.jl` is publicly available, it involves a quite large ecosystem of packages, with an intricate interplay between them. Hence, for the numerical results presented here, we chose to implement our own routines, with a minimum set of methods necessary for the convergence estimates. This is done mostly for the sake of transparency, so that the reviewing process for

checking the accuracy of the implementation becomes easier. All the source code for the numerical simulations presented below are in the github repository [23].

8.2.1. *Simulating a Random ODE.* For the simulation of an approximation of a pathwise solution of a random ODE

$$\begin{cases} \dfrac{\mathrm{d}X_t}{\mathrm{d}t} = f(t, X_t, Y_t), & 0 \leq t \leq T, \\ X_t|_{t=0} = X_0, \end{cases} \tag{8.2}$$

we define the following variables

- We fix the initial time, usually `t0=0.0`, and set the final time `tf` as the desired time T for the time interval $[0, T]$.
- The initial condition `x0` as a scalar or vector, depending on whether it is a scalar equation or a system of equations.
- A noise sample `yt` as an `AbstractVector`, indexed by a time index `j`, where each `yt[j]` represents the state of the noise sample at the j-th time in the time mesh. The time mesh is computed from the initial and final times and from the length `n` of `yt`, starting at time `t0`, at index `j=1`, and advancing at equal time steps `dt = ( tf - t0 ) / (n - 1)`, up to time `tf`, at index `j=n`. The value `yt[j]` is either a scalar or a vector, depending on the dimension of the noise.
- A function representing $f(t, X_t, Y_t)$ which can either be in the out-of-place form `f(t, x, y, p)`, in the case of a scalar equation, or in the in-place form `f!(dx, t, x, y, p)`, in the case of a system of equations, where `dx` is a vector to store the value of the numerical representation of $f(t, X_t, Y_t)$. In either case, `p` is used to store extra parameters in the equation.
- A set of parameters `p = params` to be passed on to either `f(t, x, y, p)` or `f!(dx, t, x, y, p)`, when solving the equation.
- An `AbstractVector`, denoted here as `xt`, with the same length as `yt` and with each `xt[j]` of the same type as `x0`, pre-allocated in memory to hold the solution sample path.

Then, one computes the Euler approximation via

```
solve!(xt, t0, tf, x0, f, yt, params, ::RandomEuler).
```

This fills up `xt` with the sample solution.

8.2.2. *Calculating the order of convergence.* For the implementation of the calculation of the order of convergence as described in Section 8.1, we also define the following variables.

- A random seed `rng` as an `AbstractRNG`, from the `Random` standard Julia library, so the computed samples are reproducible. In the applications, we use the Julia implementation of the Xoshiro256++ pseudorandom number generator [8].

- The initial distribution law `x0law` as a univariate or multivariate distribution from the `Distributions.jl` package [5], depending on whether X_t is a scalar or vector valued.
- A `noise` defining a univariate or multivariate process. There are several implemented noises, with the corresponding parameter options. The implemented univariate processes are
    - `WienerProcess(...)`
    - `OrnsteinUhlenbeckProcess(...)`
    - `HomogeneousLinearItoProcess(...)`
    - `GeometricBrownianMotionProcess(...)`
    - `CompoundPoissonProcess(...)`
    - `PoissonStepProcess(...)`
    - `ExponentialHawkesProcess(...)`
    - `TransportProcess(...)`
    - `FractionalBrownianMotionProcess(...)`.

  Then there is the option to create a multivariate process
    - `ProductProcess(noises...)`

  with any combination of the univariate process as components. For a thorough explanation of the noises, see the github repository [23].
- A `target` method, which is either `RandomEuler()` to solve the equation via Euler approximation on a finer mesh or a custom method to yield an exact sample of the solution, when a closed-form solution for the exact process $\{X_t\}_t$ is available.
- The method `method = RandomEuler()` for solving the approximations at each resolution N_i via Euler method. The method `RandomEuler()` is just an instance of a `RandomEuler` type, which indicates the solver to use the routine for computing the Euler approximation with the given parameters and stores cached vectors for the efficient computation in the case of systems of equations.
- A target resolution `ntg` representing N_{tgt}.
- A vector `ns` containing the desired resolutions $(N_i)_{i=1,\ldots,I}$ for the coarsers meshes.
- An integer `m` representing the number M of Monte Carlo simulations.
- When approximating a PDE, a vector `ks` of steps used to reduce the spatial resolution may be provided. It is a vector of the same length as `ns`, to accompany each time mesh resolution N_i. If the target approximation $X_{t_j}^{N_{\mathrm{tgt}}}$ has dimension/spatial resolution `d = length(xt[i])`, then each $X_{t_j}^{N_i}$ has dimension/resolution `div(d-1,ks[i])+1`. If `ks` is not provided, it defaults to a vector of the same length as `ns` and all components equal to `1` for no change in resolution.

- With all these variables set up, the "convergence suite" is defined by

```
suite = ConvergenceSuite(t0, tf, x0law, f, noise,
                         params, target, method, ntgt, ns, m, ks).
```

Then, we compute the convergence order via

```
result = solve(rng, suite).
```

The returned `result` is of a type `ConvergenceResult`, which stores all the necessary info about the errors at each resolution, their confidence intervals, the order of convergence and its confidence interval, and so on. The error table can be accessed via `generate_error_table(results)`. The order of convergence and the bounds for the 95% confidence interval are stored in `result.p`, `result.pmin`, `result.pmax`. A plot of the convergence errors, the confidence intervals, and the fitted error estimate $C\Delta t^p$ can be visualized via `plot(results)`, which contains a plotting recipe based on the `Plots.jl` package [11].

8.3. **Numerical examples.** We now describe in more details the examples used to illustrate the order of convergence. We start with some linear equations with all sorts of noises. Then we illustrate the $\min\{H + 1/2, 1\}$ order of convergence in the case of a fractional Brownian motion (fBm) noise with Hurst parameter $0 < H < 1$. Next we move to a number of linear and nonlinear models, as mentioned in the Section 1.

8.3.1. *Homogeneous linear equation with Wiener noise.* We start with one of the simplest Random ODEs, that of a linear homogenous equation with a Wiener process as the coefficient:

$$\begin{cases} \dfrac{\mathrm{d}X_t}{\mathrm{d}t} = W_t X_t, & 0 \le t \le T, \\ X_t|_{t=0} = X_0. \end{cases} \tag{8.3}$$

Since the noise is simply a Wiener process, the corresponding RODE can be turned into an SDE with an additive noise. In this 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. Moreover, it is known that the Euler-Maruyama method for an SDE with additive noise is of strong order 1 [19]. Nevertheless, we illustrate the strong convergence directly for the Euler method for this RODE equation, for the sake of completeness.

Equation (8.3) has the explicit solution

$$X_t = e^{\int_0^t W_s \, \mathrm{d}s} X_0. \tag{8.4}$$

When we compute an approximate solution via Euler's method, however, we only draw the realizations $\{W_{t_i}\}_{i=0}^N$ of a sample path, on the mesh points. We cannot compute the exact integral $\int_0^{t_j} W_s \, \mathrm{d}s$ just from these values, and, in fact, an exact

solution is not uniquely defined from them. We can, however, find its exact distribution and use that to draw feasible exact solutions and use them to estimate the error.

The exact distribution of $\int_0^\tau W_s \, \mathrm{d}s$ given the step $\Delta W = W_\tau - W_0 = W_\tau$ is computed in [18, Section 14.2] as

$$\int_0^\tau W_s \, \mathrm{d}s \sim \frac{\tau}{2}\Delta W + \sqrt{\frac{\tau^3}{12}}\mathcal{N}(0,1) = \frac{\tau}{2}\Delta W + \mathcal{N}\left(0, \frac{\tau^3}{12}\right). \tag{8.5}$$

As for the distribution of the integral over a mesh interval $[t_i, t_{i+1}]$ when given the endpoints W_{t_i} and $W_{t_{i+1}}$, we use that $s \mapsto W_{t_i+s} - W_{t_i}$ is a standard Wiener process to find, from (8.5), that

$$\begin{aligned}
\int_{t_i}^{t_{i+1}} W_s \, \mathrm{d}s &= W_{t_i}(t_{i+1} - t_i) + \int_{t_i}^{t_{i+1}} (W_s - W_{t_i}) \, \mathrm{d}s \\
&= W_{t_i}(t_{i+1} - t_i) + \int_0^{t_{i+1}-t_i} (W_{t_i+s} - W_{t_i}) \, \mathrm{d}s \\
&= W_{t_i}(t_{i+1} - t_i) + \frac{(t_{i+1} - t_i)}{2}(W_{t_{i+1}} - W_{t_i}) + Z_i \\
&= \frac{(W_{t_{i+1}} + W_{t_i})}{2}(t_{i+1} - t_i) + Z_i,
\end{aligned}$$

where

$$Z_i \sim \mathcal{N}\left(0, \frac{(t_{i+1} - t_i)^3}{12}\right) = \frac{\sqrt{(t_{i+1} - t_i)^3}}{\sqrt{12}}\mathcal{N}(0,1). \tag{8.6}$$

Thus, breaking down the sum over the mesh intervals,

$$\int_0^{t_j} W_s \, \mathrm{d}s = \sum_{i=0}^{j-1} \int_{t_i}^{t_{i+1}} W_s \, \mathrm{d}s = \sum_{i=0}^{j-1} \left(\frac{(W_{t_{i+1}} + W_{t_i})}{2}(t_{i+1} - t_i) + Z_i\right), \tag{8.7}$$

where $Z_0, \ldots, Z_{N-1}$ are independent random variables with distribution (8.6).

Notice the first term in (8.7) is the trapezoidal rule, with the second one being a correction term.

For a normal variable $N \sim \mathcal{N}(\mu, \sigma)$, the expectation of the random variable e^N is $\mathbb{E}[e^N] = e^{\mu + \sigma^2/2}$. Hence,

$$\mathbb{E}[e^{Z_i}] = e^{((t_{i+1} - t_i)^3)/24}. \tag{8.8}$$

This is the contribution of this random variable to the mean of the exact solution. But we actually draw directly Z_i and compute $e^{\sum_i Z_i}$.

Hence, once an Euler approximation of (8.3) is computed, along with realizations $\{W_{t_i}\}_{i=0}^N$ of a sample path of the noise, we consider an exact pathwise solution given by

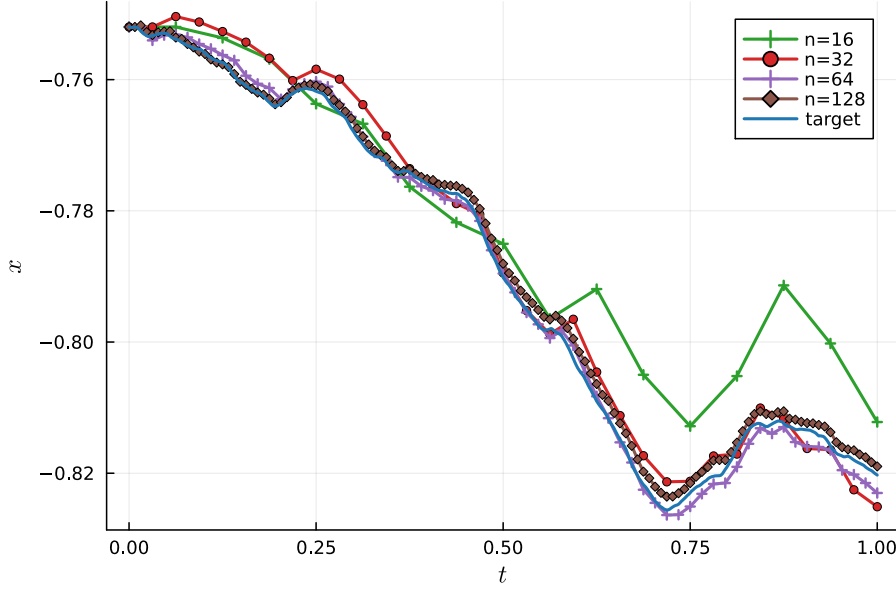$$X_{t_j} = X_0 e^{\sum_{i=0}^{j-1}\left(\frac{1}{2}\left(W_{t_i} + W_{t_{i+1}}\right)(t_{i+1} - t_i) + Z_i\right)}, \tag{8.9}$$

FIGURE 7. Euler approximation of a sample solution of $\mathrm{d}X_t/\mathrm{d}t = W_t X_t$ with $X_0 \sim \mathcal{N}(0,1)$, with a few mesh resolutions and with the target resolution.

for realizations Z_i drawn from a normal distributions given by (8.6). Figure 7 shows an approximate solution and a few sample paths of exact solutions associated with the given realizations of the noise on the mesh points.

The function $f(t, x, y) = yx$ associated with equation (8.3) does not satisfy the uniform Lipschitz condition (2.6). Nevertheless, thanks to the form of the exact solution (8.4) the form of the approximation (8.9) and estimates such as (8.8), the proof of Theorem 7.1 can be adapted to show strong order 1 convergence in this case, as well.

Table 1 shows the estimated strong error obtained in the numerical experiments, based on a target solution computed with the exact distribution. Figure 8 illustrates the order of convergence.

8.3.2. *Non-homogeneous linear system of RODEs with different types of noises.* Next, 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 second, 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.

N	dt	error	std err
16	0.0625	0.0402	0.00602
32	0.0312	0.0214	0.00351
64	0.0156	0.0102	0.00176
128	0.00781	0.00519	0.000931
256	0.00391	0.00253	0.000431
512	0.00195	0.00123	0.000201
1024	0.000977	0.000621	9.97e-5
2048	0.000488	0.00031	4.93e-5
4096	0.000244	0.000166	2.57e-5
8192	0.000122	8.22e-5	1.33e-5
16384	6.1e-5	3.9e-5	6.94e-6

TABLE 1. Mesh points (N), time steps (dt), strong error (error), and standard error (std err) of the Euler method for $\mathrm{d}X_t/\mathrm{d}t = W_t X_t$ for each mesh resolution N, with initial condition $X_0 \sim \mathcal{N}(0,1)$ and a standard Wiener process noise $\{W_t\}_t$, on the time interval $I = [0.0, 1.0]$, based on $M = 200$ sample paths for each fixed time step, with the target solution calculated with 65536 points. The order of strong convergence is estimated to be $p = 1.001$, with the 95% confidence interval $[0.936, 1.0664]$.

The equations take the form

$$\begin{cases} \dfrac{\mathrm{d}X_t}{\mathrm{d}t} = -\left\| Y_t \right\|^2 X_t + Y_t, & 0 \le t \le T, \\ X_t|_{t=0} = X_0, \end{cases} \tag{8.10}$$

where $\{X_t\}_t$ is either a scalar or a vector-valued process, and $\{Y_t\}_t$ is, correspondingly, either a scalar or a vector-valued noise process with the same dimension as X_t. In the case of the system, each coordinate of $\{Y_t\}_t$ 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) $\{H_t\}_t$ 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 $H_0 = 0.2$;
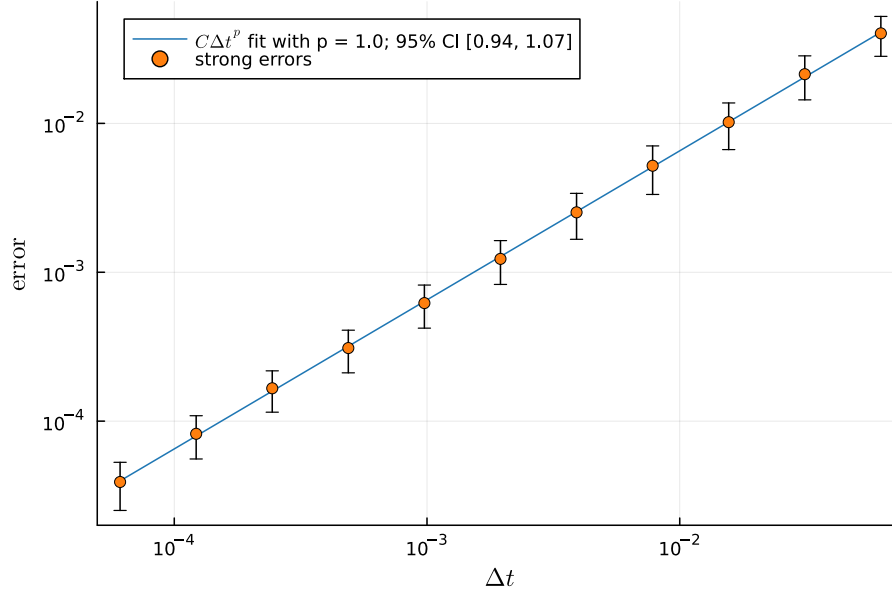
FIGURE 8. Estimated order of convergence of the strong error of the Euler method for $\mathrm{d}X_t/\mathrm{d}t = W_t X_t$, based on Table 1.

(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 2 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 9 illustrates the obtained order of convergence, while Figure 10 illustrates some sample paths of all the noises used in this system.

Table 3 and Figure 11 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.

N	dt	error	std err
64	0.0156	0.209	0.0198
128	0.00781	0.102	0.00973
256	0.00391	0.0498	0.00473
512	0.00195	0.0251	0.00239

TABLE 2. 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 $M = 92$ sample paths for each fixed time step, with the target solution calculated with 262144 points. The order of strong convergence is estimated to be $p = 1.021$, with the 95% confidence interval $[0.9108, 1.1308]$.
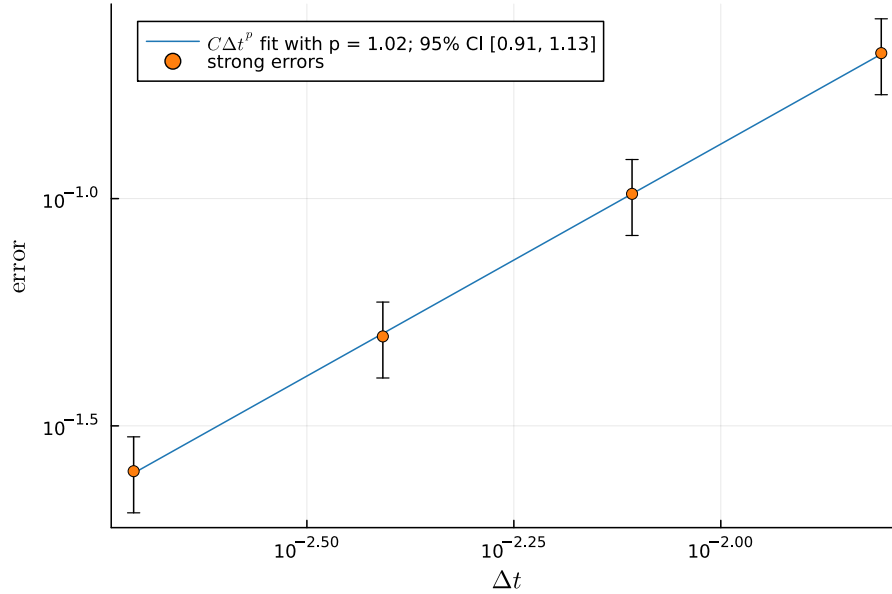

FIGURE 9. Order of convergence of the strong error of the Euler method for $\mathrm{d}X_t/\mathrm{d}t = -\|Y_t\|^2 X_t + Y_t$, based on Table 2.

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 this 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 [19].
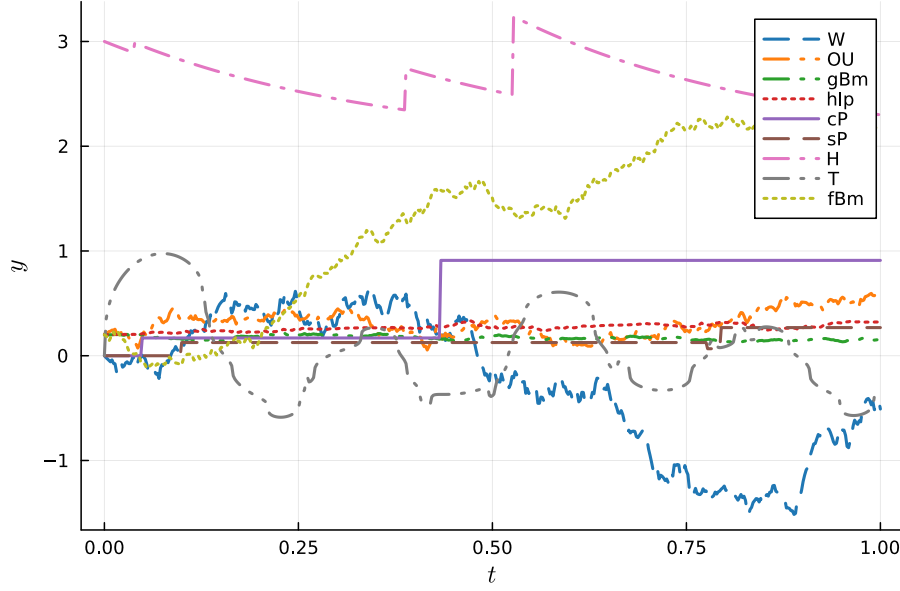
FIGURE 10. Sample paths of all the noises used in the linear system (8.10), mixing all different types of implemented noises.

noise	p	$p_{\min}$	$p_{\max}$
all noises combined (all)	1.02082	0.910835	1.13082
Wiener (W)	1.00851	0.893154	1.11785
OrnsteinUhlenbeck (OU)	0.957151	0.854775	1.06022
GeometricBrownianMotion (gBm)	0.925382	0.770407	1.08201
HomogeneousLinearIto (hlp)	1.00657	0.878322	1.13411
CompoundPoisson (cP)	0.975898	0.735864	1.22554
PoissonStep (sP)	0.969316	0.851299	1.08621
ExponentialHawkes (H)	1.0213	0.92944	1.11311
Transport (T)	1.06509	0.983371	1.14687
FractionalBrownianMotion (fBm)	1.01897	0.887446	1.14729

TABLE 3. Type of noise, order p of strong convergence, and 95% confidence interval $p_{\min}$ and $p_{\max}$, for the linear equation (8.10) with various types of noise.

The geometric Brownian motion is also expected to yield strong order 1 convergence thanks to the results in [39] and the fact that it is an Itô diffusion process. For the remaining noises, however, previous works would estimate the order of convergence to be below the order 1 attained here.
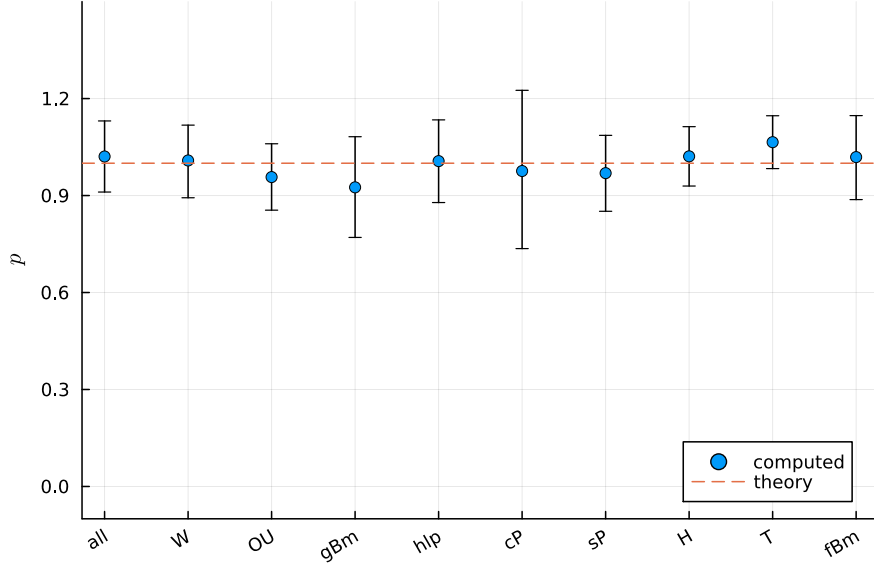
FIGURE 11. Order of convergence for the equation (8.10) as it varies among different types of noises, starting with all the noises combined in a system of equations (all) and continuing with the scalar cases, each with a different type of noise, according to the table Table 3

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

8.3.3. *Fractional Brownian motion noise.* Now, we consider again a linear equation, this time of the form

$$\begin{cases} \dfrac{\mathrm{d}X_t}{\mathrm{d}t} = -X_t + B_t^H, & 0 \leq t \leq T, \\ X_t|_{t=0} = X_0, \end{cases} \tag{8.11}$$

and where the noise $\{B_t^H\}_t$ is assumed to be a fractional Brownian motion (fBm) with Hurst parameter $0 < H < 1$. It turns out that, for $0 < H < 1/2$, the order of convergence is $H + 1/2$, while for $1/2 \leq H < 1$, the order of convergence is 1. In any case, this is higher than the order H estimated in previous works. The same seems to hold for a nonlinear dependency on the fBm, but the proof is more involved, depending on a fractional Itô formula (see [7, Theorem 4.2.6], [4, Theorem 4.1]and [29, Theorem 2.7.4]), based on the Wick Itô Skorohod (WIS) integral (see [7, Chapter 4]). A corresponding WIS isometry is also needed (see e.g. [7, Theorem 4.5.6]), involving Malliavin calculus and fractional derivatives. For these reasons, we leave

the nonlinear case to a subsequent work and focus on this simple linear example, which suffices to illustrate the peculiarity of the dependence on H of the order of convergence. For this linear equation, the proof of convergence is done rigorously below, with the framework developed in the main article.

We need to estimate the last term in (6.3) of Proposition 6.1, involving the steps of the term $f(t, x, y) = -x + y$, which in this case reduce to

$$f(s, X^N_{\tau^N(s)}, Y_s) - f(\tau^N(s), X^N_{\tau^N(s)}, Y_{\tau^N(s)}) = B^H_s - B^H_{\tau^N(s)}, \tag{8.12}$$

for $0 \le s \le T$. There are several ways to represent an fBm (see e.g. [7, 29]). We use the formula (see [26, eq. (2.1)] or [7, eq. (1.1)])

$$B^H_t = \frac{1}{\Gamma(H + 1/2)} \left(\int_{-\infty}^0 \left((t - s)^{H-1/2} - (-s)^{H-1/2} \right) \, dW_s \right.$$
$$\left. + \int_0^t (t - s)^{H-1/2} \, dW_s \right), \tag{8.13}$$

where $\Gamma(\cdot)$ is the well-known Gamma function. For the step, (8.13) means that

$$B^H_s - B^H_{\tau^N(s)} = \frac{1}{\Gamma(H + 1/2)} \left(\int_{-\infty}^{\tau^N(s)} \left((s - \xi)^{H-1/2} - (\tau^N(s) - \xi)^{H-1/2} \right) \, dW_\xi \right.$$
$$\left. + \int_{\tau^N(s)}^s (s - \xi)^{H-1/2} \, dW_\xi \right). \tag{8.14}$$

Then, using the stochastic Fubini Theorem to exchange the order of integration (see, again, [32, Section IV.6]),

$$\int_0^{t_j} \left(f(s, X^N_{\tau^N(s)}, Y_s) - f(\tau^N(s), X^N_{\tau^N(s)}, Y_{\tau^N(s)}) \right) \, ds$$

$$= \frac{1}{\Gamma(H + 1/2)} \int_0^{t_j} \int_{-\infty}^{\tau^N(s)} \left((s - \xi)^{H-1/2} - (\tau^N(s) - \xi)^{H-1/2} \right) \, dW_\xi \, ds$$

$$+ \frac{1}{\Gamma(H + 1/2)} \int_0^{t_j} \int_{\tau^N(s)}^s (s - \xi)^{H-1/2} \, dW_\xi \, ds$$

$$= \frac{1}{\Gamma(H + 1/2)} \int_{-\infty}^0 \int_0^{t_j} \left((s - \xi)^{H-1/2} - (\tau^N(s) - \xi)^{H-1/2} \right) \, ds \, dW_\xi$$

$$+ \frac{1}{\Gamma(H + 1/2)} \int_0^{t_j} \int_{\tau^N(\xi) + \Delta t_N}^{t_j} \left((s - \xi)^{H-1/2} - (\tau^N(s) - \xi)^{H-1/2} \right) \, ds \, dW_\xi$$

$$+ \frac{1}{\Gamma(H + 1/2)} \int_0^{t_j} \int_\xi^{\tau^N(\xi) + \Delta t_N} (s - \xi)^{H-1/2} \, ds \, dW_\xi. \tag{8.15}$$

For the first term, notice $\sigma \mapsto 1/(\sigma - \xi)^{H-1/2}$ is continuously differentiable on the interval $\sigma > \xi$, so that

$$(s - \xi)^{H-1/2} - (\tau^N(s) - \xi)^{H-1/2} = -(H - 1/2) \int_{\tau^N(s)}^{s} (\sigma - \xi)^{H-3/2} \, d\sigma.$$

Thus,

$$\int_0^{t_j} \left((s - \xi)^{H-1/2} - (\tau^N(s) - \xi)^{H-1/2} \right) \, ds$$

$$= (H - 1/2) \int_0^{t_j} \int_{\tau^N(s)}^{s} (\sigma - \xi)^{H-3/2} \, d\sigma \, ds.$$

Exchanging the order of integration yields

$$\int_0^{t_j} \left((s - \xi)^{H-1/2} - (\tau^N(s) - \xi)^{H-1/2} \right) \, ds$$

$$= (H - 1/2) \int_0^{t_j} \int_\sigma^{\tau^N(\sigma) + \Delta t_N} (\sigma - \xi)^{H-3/2} \, ds \, d\sigma$$

$$= (H - 1/2) \int_0^{t_j} \left(\tau^N(\sigma) + \Delta t_N - \sigma \right) (\sigma - \xi)^{H-3/2} \, d\sigma.$$

Hence,

$$\left| \int_0^{t_j} \left((s - \xi)^{H-1/2} - (\tau^N(s) - \xi)^{H-1/2} \right) \, ds \right|$$

$$\leq (1/2 - H) \int_0^{t_j} \Delta t_N (\sigma - \xi)^{H-3/2} \, d\sigma.$$

Now, using the Lyapunov inequality and the Itô isometry, and using the same trick as above,

$$\mathbb{E} \left[\left| \int_{-\infty}^0 \int_0^{t_j} \left((s - \xi)^{H-1/2} - (\tau^N(s) - \xi)^{H-1/2} \right) \, ds \, dW_\xi \right| \right]$$

$$\leq \left(\int_{-\infty}^0 \left(\int_0^{t_j} \left((s - \xi)^{H-1/2} - (\tau^N(s) - \xi)^{H-1/2} \right) \, ds \right)^2 \, d\xi \right)^{1/2}$$

$$\leq \Delta t_N \left(\int_{-\infty}^0 \left((1/2 - H) \int_0^{t_j} (\sigma - \xi)^{H-3/2} \, d\sigma \right)^2 \, d\xi \right)^{1/2}$$

$$\leq (1/2 - H) \Delta t_N \left(\int_{-\infty}^0 \left(\int_0^{T} (\sigma - \xi)^{H-3/2} \, d\sigma \right)^2 \, d\xi \right)^{1/2}.$$

Therefore,

$$\frac{1}{\Gamma(H+1/2)}\Delta t_N \mathbb{E}\left[\left|\int_{-\infty}^{0}\int_{0}^{t_j}\left((s-\xi)^{H-1/2}-(\tau^N(s)-\xi)^{H-1/2}\right)\,\mathrm{d}s\,\mathrm{d}W_\xi\right|\right]$$
$$\le c_H^{(1)}\Delta t_N, \quad (8.16)$$

for a suitable constant $c_H^{(1)}$. We see this term is of order 1 in Δt_N.

The second term is similar,

$$\int_{\tau^N(\xi)+\Delta t_N}^{t_j}\left((s-\xi)^{H-1/2}-(\tau^N(s)-\xi)^{H-1/2}\right)\,\mathrm{d}s$$

$$= (H-1/2)\int_{\tau^N(\xi)+\Delta t_N}^{t_j}\int_{\tau^N(s)}^{s}(\sigma-\xi)^{H-3/2}\,\mathrm{d}\sigma\,\mathrm{d}s$$

$$= (H-1/2)\int_{\tau^N(\xi)+\Delta t_N}^{t_j}\int_{\sigma}^{\tau^N(\sigma)+\Delta t_N}(\sigma-\xi)^{H-3/2}\,\mathrm{d}s\,\mathrm{d}\sigma$$

$$= (H-1/2)\int_{\tau^N(\xi)+\Delta t_N}^{t_j}\left(\tau^N(\sigma)+\Delta t_N-\sigma\right)(\sigma-\xi)^{H-3/2}\,\mathrm{d}\sigma.$$

Thus,

$$\left|\int_{\tau^N(\xi)+\Delta t_N}^{t_j}\left((s-\xi)^{H-1/2}-(\tau^N(s)-\xi)^{H-1/2}\right)\,\mathrm{d}s\right|$$

$$\le (1/2-H)\Delta t_N \int_{\tau^N(\xi)+\Delta t_N}^{t_j}(\sigma-\xi)^{H-3/2}\,\mathrm{d}\sigma.$$

Hence,

$$\mathbb{E}\left[\left|\int_{0}^{t_j}\int_{\tau^N(\xi)+\Delta t_N}^{t_j}\left((s-\xi)^{H-1/2}-(\tau^N(s)-\xi)^{H-1/2}\right)\,\mathrm{d}s\,\mathrm{d}W_\xi\right|\right]$$

$$\le \left(\int_{0}^{t_j}\left(\int_{\tau^N(\xi)+\Delta t_N}^{t_j}\left((s-\xi)^{H-1/2}-(\tau^N(s)-\xi)^{H-1/2}\right)\,\mathrm{d}s\right)^2\,\mathrm{d}\xi\right)^{1/2}$$

$$\le \Delta t_N(1/2-H)\left(\int_{0}^{t_j}\left(\int_{\tau^N(\xi)+\Delta t_N}^{T}(\sigma-\xi)^{H-3/2}\,\mathrm{d}\sigma\right)^2\,\mathrm{d}\xi\right)^{1/2}.$$

Therefore,

$$\frac{1}{\Gamma(H+1/2)}\mathbb{E}\left[\left|\int_{0}^{t_j}\int_{\tau^N(\xi)+\Delta t_N}^{t_j}\left((s-\xi)^{H-1/2}-(\tau^N(s)-\xi)^{H-1/2}\right)\,\mathrm{d}s\,\mathrm{d}W_\xi\right|\right]$$
$$\le c_H^{(2)}\Delta t_N, \quad (8.17)$$

for a possibly different constant $c_H^{(2)}$. This term is also of order 1.

For the last term, we have

$$0 \leq \int_{\xi}^{\tau^N(\xi)+\Delta t_N} (s-\xi)^{H-1/2}\, \mathrm{d}s = \frac{1}{H+1/2}(\tau^N(\xi)+\Delta t_N - \xi)^{H+1/2}$$

$$\leq \frac{1}{H+1/2}\Delta t_N^{H+1/2}.$$

so that, using the Lyapunov inequality and the Itô isometry

$$\mathbb{E}\left[\left\|\int_0^{t_j}\int_{\xi}^{\tau^N(\xi)+\Delta t_N}(s-\xi)^{H-1/2}\,\mathrm{d}s\,\mathrm{d}W_{\xi}\right\|\right]$$

$$\leq \left(\int_0^{t_j}\left(\int_{\xi}^{\tau^N(\xi)+\Delta t_N}(s-\xi)^{H-1/2}\,\mathrm{d}s\right)^2\,\mathrm{d}\xi\right)^{1/2}$$

$$\leq \left(\int_0^{t_j}\Delta t_N^{2H+1}\,\mathrm{d}\xi\right)^{1/2} \leq t_j^{1/2}\Delta t_N^{H+1/2}.$$

Therefore,

$$\frac{1}{\Gamma(H+1/2)}\mathbb{E}\left[\left\|\int_0^{t_j}\int_{\xi}^{\tau^N(\xi)+\Delta t_N}(s-\xi)^{H-1/2}\,\mathrm{d}s\,\mathrm{d}W_{\xi}\right\|\right] \leq c_H^{(3)}\Delta t_N^{H+1/2}, \qquad (8.18)$$

for a third constant $c_H^{(3)}$.

Putting the three estimates (8.16), (8.17), (8.18) in (8.15) we find that

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

$$\leq c_H^{(4)}\Delta t_N + c_H^{(3)}\Delta t_N^{H+1/2}, \quad (8.19)$$

where $c_H^{(4)} = c_H^{(1)} + c_H^{(2)}$. Using this estimate in Proposition 6.1 shows that the Euler method is of order $H+1/2$, when $0 < H < 1/2$, and is of order 1, when $1/2 \leq H < 1$, having in mind that $H = 1/2$ corresponds to the classical Wiener process.

In summary, we have proved the following result.

Theorem 8.1. *Consider the equation (8.11) where $\{B_t^H\}_t$ is a fractional Brownian motion (fBm) with Hurst parameter $0 < H < 1$. Suppose the initial condition X_0 satisfies*

$$\mathbb{E}[\|X_0\|^2] < \infty. \qquad (8.20)$$

Then, the Euler scheme for this initial value problem is of strong order $H+1/2$, for $0 < H < 1/2$, and is of order 1, for $1/2 \leq H < 1$. More precisely,

$$\max_{j=0,\dots,N}\mathbb{E}\left[\left\|X_{t_j} - X_{t_j}^N\right\|\right] \leq c_1\Delta t_N + c_2\Delta t_N^{H+1/2}, \qquad \forall N \in \mathbb{N}, \qquad (8.21)$$

for suitable constants $c_1, c_2 \geq 0$.

H	p	$p_{\min}$	$p_{\max}$
0.1	0.640103	0.582705	0.697128
0.2	0.742881	0.681001	0.804453
0.3	0.794755	0.735716	0.853437
0.4	0.921899	0.865631	0.978194
0.5	1.00777	0.94886	1.06662
0.7	1.00637	0.945425	1.06735
0.9	1.00559	0.94014	1.07104

TABLE 4. Hurst parameter H, order p of strong convergence, and 95% confidence interval $p_{\min}$ and $p_{\max}$, for a number of Hurst values.

As for illustrating numerically the order of strong convergence, although the above linear equation has the explicit solution

$$X_t = e^{-t} X_0 + \int_0^t e^{-(t-s)} B_s^H \, \mathrm{d}s, \tag{8.22}$$

computing a distributionally exact solution of this form is a delicate process. Thus we check the convergence numerically by comparing the approximations with another Euler approximation on a much finer mesh.

The fractional Browninan motion noise term is generated with the $\mathcal{O}(N \log N)$ fast Fourier transform (FFT) method of Davies and Harte, as presented in [13] (see also [18, Section 14.4]). Table 4 shows the obtained convergence estimates, for a number of Hurst parameters, which is illustrated in Figure 12, matching the theoretical estimate of $p = \min\{H + 1/2, 1\}$.

8.3.4. *Population dynamics with harvest.* Here, we consider a population dynamics problem modeled by a logistic equation with random coefficients, loosely inspired by [18, Section 15.2], with an extra term representing harvest:

$$\frac{\mathrm{d}X_t}{\mathrm{d}t} = \Lambda_t X_t (1 - \frac{X_t}{r}) - \alpha H_t \frac{X_t}{r + X_t}, \tag{8.23}$$

where $r, \alpha > 0$ are constants and $\{\Lambda_t\}_{t \geq 0}$ and $\{H_t\}_{t \geq 0}$ are stochastic processes. The first term is a logistic growth, with carrying capacity r and random specific growth Λ_t. The second term is the harvest term, with α being an overall modulation factor, the term $X_t/(r + X_t)$ accounting for the scarcity of the population when $X_t \ll r$, and with $0 \leq H_t \leq 1$ being a random harvest factor.

More specifically, $\{\Lambda_t\}_{t \geq 0}$ is given by

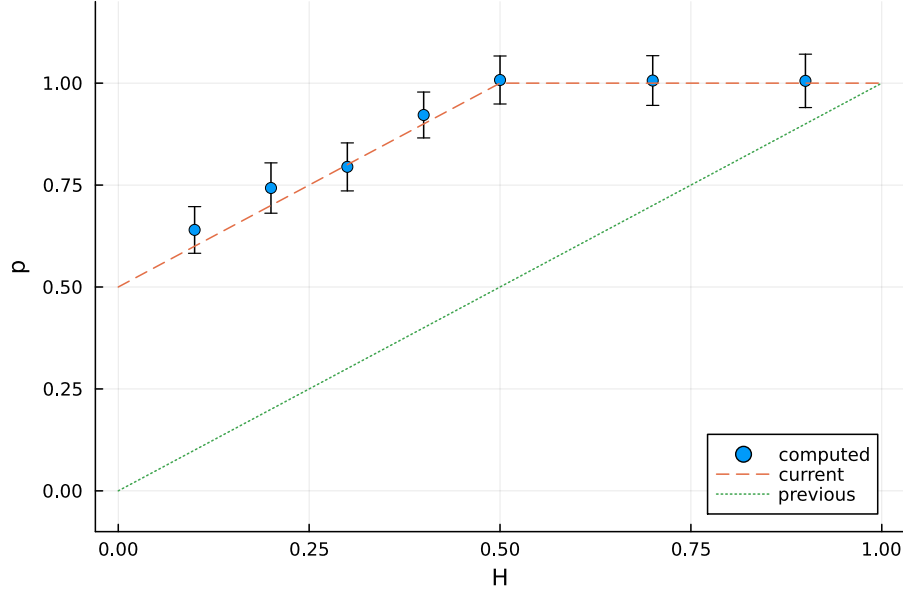$$\Lambda_t = \gamma(1 + \varepsilon \sin(G_t)),$$

FIGURE 12. 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).

where $\gamma > 0$, $0 < \varepsilon < 1$, and $\{G_t\}_{t \geq 0}$ is a geometric Brownian motion (gBm) process. A Wiener process is a natural choice, but we choose a gBm process instead since it is a multiplicative noise.

The harvest term $\{H_t\}_{t \geq 0}$ is a "Poisson step" process of the form

$$H_t = S_{N_t},$$

where $\{N_t\}_{t \geq 0}$ is a Poisson point-process with rate λ, $S_0 = 0$, and the S_i, for $i = 1, 2, \ldots$, are independent and identically distributed random variables with non-negative values within the interval $[0, 1]$, independent also of the Poisson counter $\{N_t\}_{t \geq 0}$.

We suppose the initial condition is nonnegative and bounded almost surely, i.e.

$$0 \leq X_0 \leq R,$$

for some $R > r$.

The noise process $\{\Lambda_t\}_{t \geq 0}$ itself satisfies

$$0 < \gamma - \varepsilon \leq \Lambda_t \leq \gamma + \varepsilon < 2\gamma, \qquad \forall t \geq 0.$$

Define $f : \mathbb{R} \times \mathbb{R} \times \mathbb{R}^2 \to \mathbb{R}$ by

$$f(t, x, y) = \begin{cases} \gamma(1 + \varepsilon \sin(y_1))x(r - x) - \alpha \dfrac{xy_2}{r + x}, & x \geq 0, \\ 0, & x < 0. \end{cases}$$

The equation (8.23) becomes

$$\frac{\mathrm{d}X_t}{\mathrm{d}t} = f(t, X_t, Y_t),$$

where $\{Y_t\}_{t \geq 0}$ is the vector-valued process given in coordinates by $Y_t = (G_t, H_t)$.

Notice that $f(t, x, y) = 0$, for $x < 0$, for arbitrary $y = (y_1, y_2)$, while $f(t, x, y) < 0$, for $x \geq r$, $y_2 \geq 0$, and for arbitrary y_1. Since the noise $y_2 = H_t$ is always nonnegative, we see that the interval $0 \leq x \leq R$ is positively invariant and attracts the orbits with a nonnegative initial condition. Thus, the pathwise solutions of the initial-value problem under consideration are almost surely bounded as well.

The function $f = f(t, x, y)$ is continuously differentiable infinitely many times and with bounded derivatives within the positively invariant region. Hence, within the region of interest, all the conditions of Theorem 7.1 hold and the Euler method is of strong order 1.

We simulate numerically the solutions of the above problem, with $\gamma = 0.8$, $\varepsilon = 0.3$, $r = 1.0$, and $\alpha = \gamma r^2$. The geometric Brownian motion process $\{G_t\}_{t \geq 0}$ is taken with drift coefficient $\mu = 1.0$, diffusion coefficient $\sigma = 0.8$, and initial condition $y_0 = 1.0$. The Poisson process $\{N_t\}_{t \geq 0}$ is taken with rate $\lambda = 15.0$. And the step process $\{H_t\}_{t \geq 0}$ is taken with steps following a Beta distribution with shape parameters $\alpha = 5.0$ and $\beta = 7.0$. The initial condition X_0 is taken to be a Beta distribution with shape parameters $\alpha = 7.0$ and $\beta = 5.0$, hence we can take $R = 1$.

Notice that we can write

$$\frac{\mathrm{d}X_t}{\mathrm{d}t} = \frac{X_t}{r + X_t}\left(r\Lambda_t - \alpha H_t - \frac{\Lambda_t}{r}X_t^2\right).$$

Hence, when $\alpha H_t \geq r\Lambda_t$, the population decays for arbitrary $X_t > 0$, leading to an extinction of the population. The parameters chosen above keep the population from extinction but may often get close to the critical values.

Table 5 shows the estimated strong error obtained for each mesh resolution, while Figure 13 illustrates the order of convergence, estimated to be close enough to the theoretical value of strong order 1. Finally, Figure 14 shows an approximation sequence of a sample path.

8.3.5. *A toggle-switch model for gene expression.* Here, we consider the toggle-switch model in [3, Section 7.8], originated from [38]; see also [35].

Toogle switches in gene expression consist of genes that mutually repress each other and exhibit two stable steady states of ON and OFF. It is a regulatory mechanism which is active during cell differentiation and is believed to act as a memory device, able to choose and maintain cell fate decisions.

N	dt	error	std err
16	0.0625	0.00529	0.000367
32	0.0312	0.00262	0.000172
64	0.0156	0.00126	7.86e-5
128	0.00781	0.000626	4.35e-5
256	0.00391	0.00029	2.2e-5
512	0.00195	0.00017	1.18e-5

TABLE 5. Mesh points (N), time steps (dt), strong error (error), and standard error (std err) of the Euler method for population dynamics for each mesh resolution N, with initial condition $X_0 \sim \text{Beta}(7.0, 5.0)$ and gBm and step process noises, on the time interval $I = [0.0, 1.0]$, based on $M = 100$ sample paths for each fixed time step, with the target solution calculated with 262144 points. The order of strong convergence is estimated to be $p = 1.009$, with the 95% confidence interval $[0.9603, 1.0571]$.
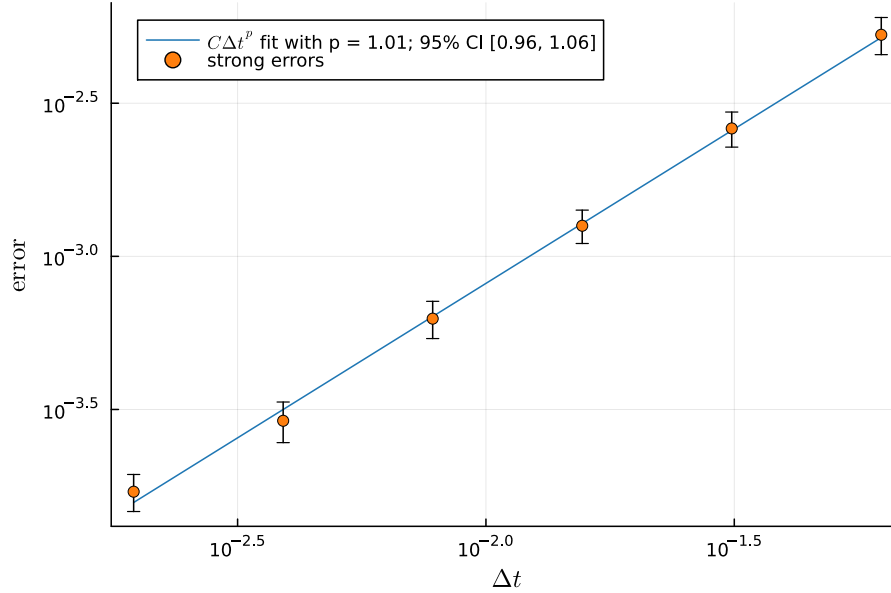

FIGURE 13. Order of convergence of the strong error of the Euler method for the population dynamics equation (8.23), based on Table 5.

We consider the following simple model as discussed in [3, Section 7.8], of two interacting genes, X and Y, with the concentration of their corresponding protein products at each time t denoted by X_t and Y_t. These are stochastic processes defined
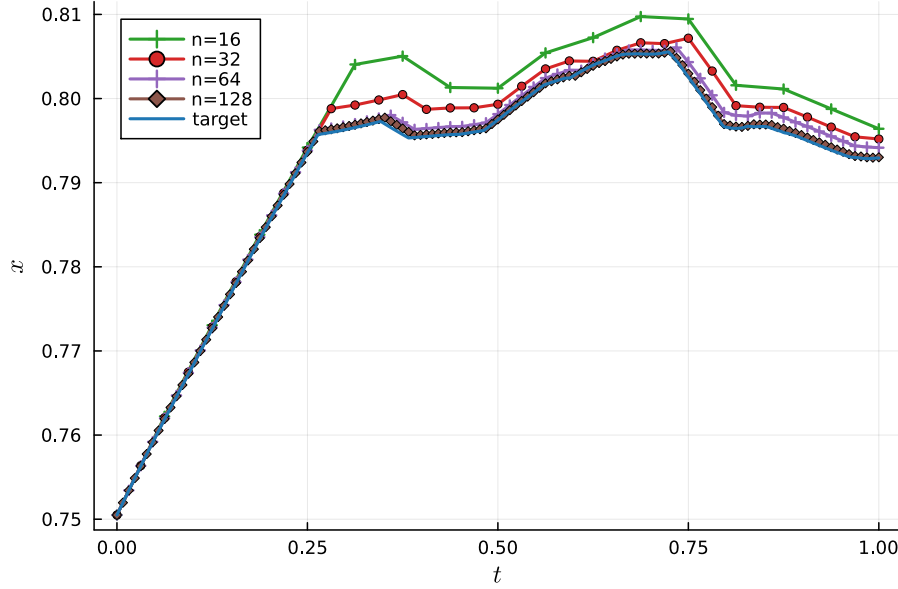
FIGURE 14. An approximation sequence of a sample path solution of the population dynamics equation (8.23).

by the system of equations

$$
\begin{cases}
\dfrac{\mathrm{d}X_t}{\mathrm{d}t} = \left(A_t + \dfrac{X_t^4}{a^4 + X_t^4}\right)\left(\dfrac{b^4}{b^4 + Y_t^4}\right) - \mu X_t, \\[3mm]
\dfrac{\mathrm{d}Y_t}{\mathrm{d}t} = \left(B_t + \dfrac{Y_t^4}{c^4 + Y_t^4}\right)\left(\dfrac{d^4}{d^4 + X_t^4}\right) - \nu Y_t,
\end{cases}
\tag{8.24}
$$

on $t \geq 0$, with initial conditions X_0 and Y_0, where $\{A_t\}_{t\geq 0}$ and $\{B_t\}_{t\geq 0}$ are given stochastic process representing the external activation on each gene; a and c determine the auto-activation thresholds; b and d determine the thresholds for mutual repression; and μ and ν are protein decay rates. In this model, the external activations A_t is taken to be a compound Poisson process, while B_t is non-autonomous homogeneous linear Itô process of the form $\mathrm{d}B_t = (\mu_1 + \mu_2 \sin(\vartheta t))B_t\,\mathrm{d}t + \sigma\sin(\vartheta t)B_t\,\mathrm{d}W_t$.

In the simulations below, we use parameters similar to those in [3, Section 7.8]. We fix $a = c = 0.25$; $b = d = 0.4$; and $\mu = \nu = 0.75$. The initial conditions are set to $X_0 = Y_0 = 4.0$. The external activation $\{A_t\}$ is a compound Poisson process with Poisson rate $\lambda = 5.0$ and jumps uniformly distributed within $[0.0, 0.5]$, while $\{B_t\}_t$ is taken with $\mu_1 = 0.7$, $\mu_2 = 0.3$, $\sigma = 0.3$, $\vartheta = 3\pi$, and initial condition $B_0 = 0.2$.

We do not have an explicit solution for the equation so we use as target for the convergence an approximate solution via Euler method at a much higher resolution.

Table 6 shows the estimated strong error obtained with the Monte Carlo method, while Figure 15 illustrates the order of convergence. Figure 16 shows a sample solution, while Figure 17 illustrates the two components (A_t, B_t) of a sample noise.

N	dt	error	std err
32	0.156	0.645	0.0344
64	0.0781	0.31	0.0158
128	0.0391	0.15	0.00774
256	0.0195	0.0733	0.0039
512	0.00977	0.0362	0.00195

TABLE 6. Mesh points (N), time steps (dt), strong error (error), and standard error (std err) of the Euler method for a toggle-switch model of gene regulation for each mesh resolution N, with initial condition $X_0 = 4.0, Y_0 = 4.0$ and coupled compound Poisson process and geometric Brownian motion noises, on the time interval $I = [0.0, 5.0]$, based on $M = 100$ sample paths for each fixed time step, with the target solution calculated with 262144 points. The order of strong convergence is estimated to be $p = 1.039$, with the 95% confidence interval $[1.0103, 1.0681]$.
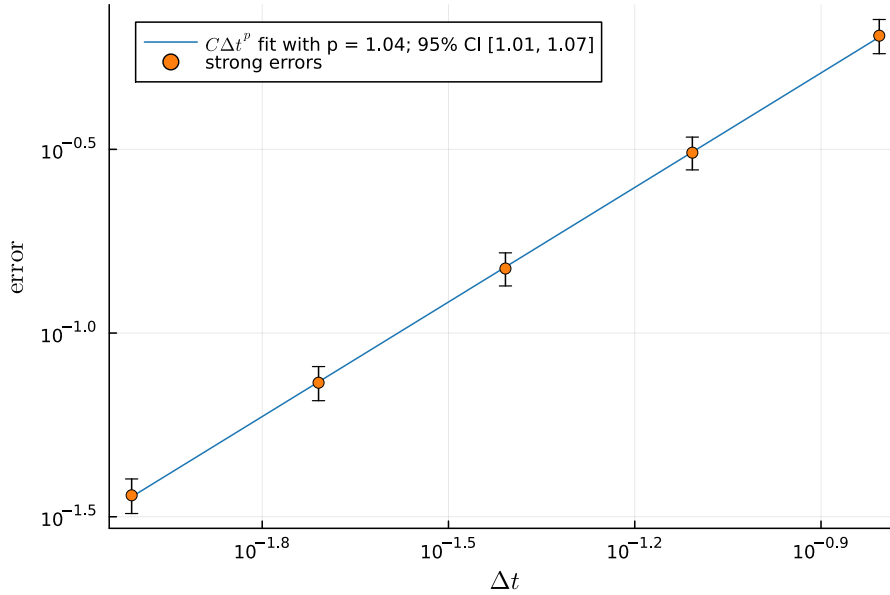

FIGURE 15. Order of convergence of the strong error of the Euler method for the toggle-switch model (8.24), based on Table 6.

8.3.6. *Mechanical structure under random Earthquake-like seismic disturbances.* Now we consider a simple mechanical structure model driven by a random disturbance in the form of a transport process, simulating seismic ground-motion excitations,
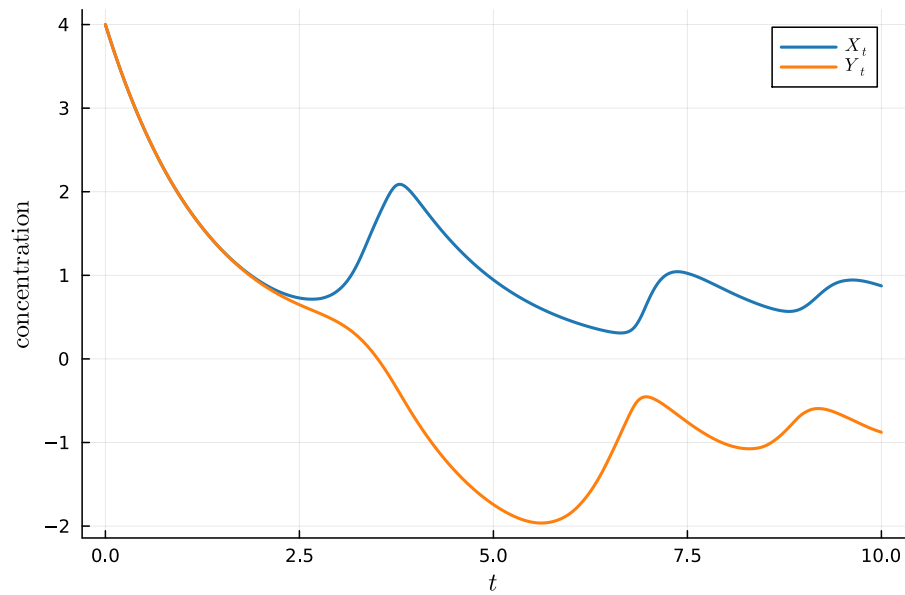
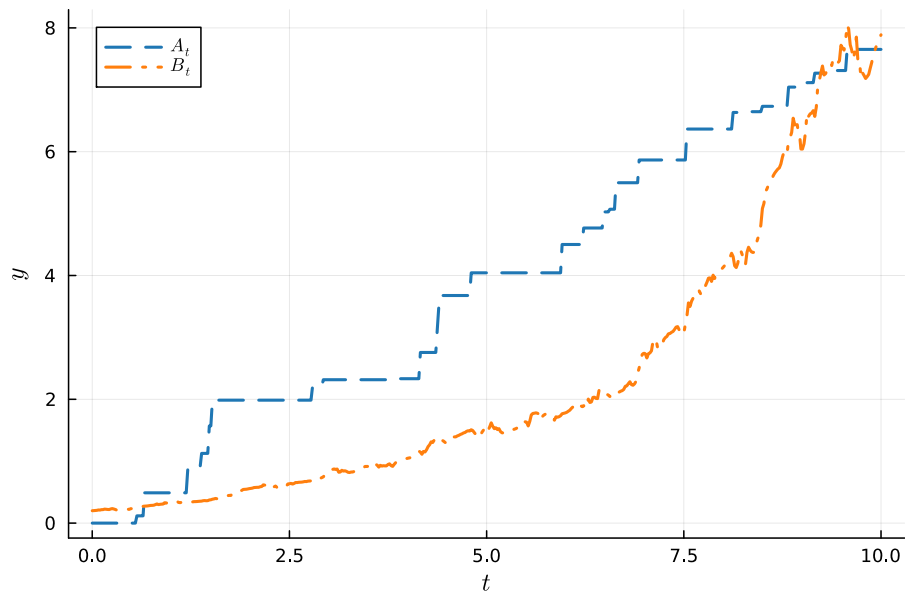FIGURE 16. Evolution of a sample solution of the toggle-switch model (8.24).

FIGURE 17. Noise sample paths for the toggle-switch model (8.24).

inspired by the model in [9] (see also [30, Chapter 18] and [20] with this and other models).

There are a number of models for earthquake-type forcing, such as the ubiquitous Kanai-Tajimi and the Clough-Penzien models, where the noise has a characteristic spectral density, determined by the mechanical properties of the ground layer. The ideia, from [22] and [37], is that the spectrum of the noise at bedrock is characterized by a constant pattern, while at the ground surface it is modified by the vibration property of the ground layer. This interaction between the bedrock and the ground layer is modeled as a stochastic oscillator driven by a zero-mean Gaussian white noise, and whose solution leads to a noise with a characteristic power spectrum.

We follow, however, the Bogdanoff-Goldberg-Bernard model, which takes the form of a transport process noise and models qualitatively well the main shock and subsequent aftershocks. We chose the later not only because of its realistic modeling but also so we can illustrate the improved convergence for such type of transport noise, complementing the other examples. The noise is described in more details shortly. Let us first introduce the model for the vibrations of the mechanical structure.

A single-storey building is considered, with its ground floor centered at position M_t and its ceiling at position $M_t + X_t$. The random process X_t refers to the motion relative to the ground. The ground motion M_t affects the motion of the relative displacement X_t as an excitation force proportional to the ground acceleration $\ddot{M}_t$. The damping and elastic forces are in effect within the structure. In this framework, the equation of motion for the relative displacement X_t of the ceiling of the single-storey building takes the form

$$\ddot{X}_t + 2\zeta_0\omega_0\dot{X}_t + \omega_0^2 X_t = -\ddot{M}_t. \tag{8.25}$$

where ζ_0 and ω_0 are damping and elastic model parameters depending on the structure of the building.

For the numerical simulations, the second-order equation is written as a system of first-order equations,

$$\begin{cases} \dot{X}_t = V_t, \\ \dot{V}_t = -\omega_0^2 X_t - 2\zeta_0\omega_0 X_t - Y_t, \end{cases}$$

where $\{V_t\}_t$ is the random velocity of the celing relative to the ground and where $\{Y_t\}_t$ is the stochastic noise excitation term given as the ground acceleration, $Y_t = \ddot{M}_t$, generated by an Earthquake and its aftershocks, or any other type of ground motion.

The structure is originally at rest, so we have the conditions

$$X_0 = 0, \quad V_0 = \dot{X}_0 = 0.$$

In the Bogdanoff-Goldberg-Bernard model [9], the excitation $\ddot{M}_t$ is made of a composition of oscillating signals with random frequencies, modulated by a linear attack rate followed by an exponential decay. This can be written, more precisely, as

$$\sum_{j=1}^{n} a_j t e^{-\delta_j t} \cos(\omega_j t + \theta_j).$$

In order to simulate the start of the first shock-wave and the subsequent aftershocks, we modify this model sligthly to be a combination of such terms but at different incidence times. We also remove the attack rate from the excitation to obtain a rougher instantaneous, discontinuous excitation. This jump-like excitation is connected with a square power attack rate for the displacement itself. Finally, for simulation purposes, we model the displacement M_t instead of modeling directly the excitation $\ddot{M}_t$, but in such a way that the ground-motion excitation follows essentially the proposed signal.

Thus, with this framework in mind, we model the ground displacement as a transport process composed of a series of time-translations of a square-power "attack" front, with an exponentially decaying tail and an oscillating background wave:

$$M_t = \sum_{i=1}^{k} \gamma_i (t - \tau_i)_+^2 e^{-\delta_i (t - \tau_i)} \cos(\omega_i (t - \tau_i)), \tag{8.26}$$

where $k \in \mathbb{N}$ is given, $(t - \tau_i)_+ = \max\{0, t - \tau_i\}$ is the positive part of the function, and the parameters γ_i, τ_i, δ_i, and ω_i are all random variables, with τ_i being exponentially distributed, and γ_i, δ_i, and ω_i being uniformly distributed, each with different support values, and all of them independent of each other.

The excitation itself becomes

$$
\begin{aligned}
\ddot{M}(t) = & 2 \sum_{i=1}^{k} \gamma_i H(t - \tau_i) e^{-\delta_i (t - \tau_i)} \cos(\omega_i (t - \tau_i)) \\
& + \sum_{i=1}^{k} \gamma_i (\delta_i^2 - \omega_i^2)(t - \tau_i)_+^2 e^{-\delta_i (t - \tau_i)} \cos(\omega_i (t - \tau_i)) \\
& - 2 \sum_{i=1}^{k} \gamma_i (\delta_i + \omega_i)(t - \tau_i)_+ e^{-\delta_i (t - \tau_i)} \cos(\omega_i (t - \tau_i)) \\
& + \delta_i \sum_{i=1}^{k} \omega_i \gamma_i (t - \tau_i)_+^2 e^{-\delta_i (t - \tau_i)} \sin(\omega_i (t - \tau_i)),
\end{aligned}
$$

where $H = H(s)$ is the Heaviside function, where, for definiteness, we set $H(s) = 1$, for $s \geq 1$, and $H(s) = 0$, for $s < 0$.

More specifically, for the numerical simulations, we use $\zeta_0 = 0.6$ and $\omega_0 = 15$ as the structural parameters. For the transport process, we set $k = 8$ and define the random parameters as the following exponential and uniform distributions: $\tau_i \sim \text{Exponential}(0.5)$, $\gamma_i \sim \text{Uniform}(16.0, 32.0)$, $\delta_i \sim \text{Uniform}(12.0, 16.0)$, and $\omega_i \sim \text{Uniform}(16\pi, 32\pi)$.

Table 7 shows the estimated strong error obtained with the Monte-Carlo method, while Figure 18 illustrates the order of convergence. Figure 19 shows a sample ground motion and the corresponding ground acceleration and its envelope.

N	dt	error	std err
64	0.0312	9.27	0.721
128	0.0156	3.66	0.308
256	0.00781	1.75	0.142
512	0.00391	0.838	0.0672

TABLE 7. Mesh points (N), time steps (dt), strong error (error), and standard error (std err) of the Euler method for mechanical structure model under ground-shaking random excitations for each mesh resolution N, with initial condition $X_0 = \mathbf{0}$ and transport process noise, on the time interval $I = [0.0, 2.0]$, based on $M = 100$ sample paths for each fixed time step, with the target solution calculated with 262144 points. The order of strong convergence is estimated to be $p = 1.147$, with the 95% confidence interval $[1.0595, 1.2341]$.
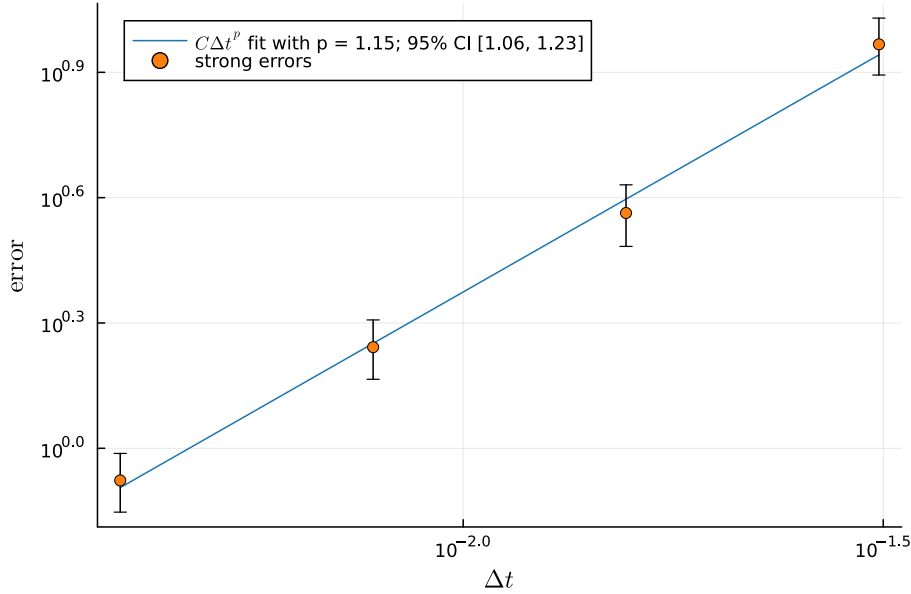

FIGURE 18. Order of convergence of the strong error of the Euler method for the mechanical structure model (8.25), based on Table 7.

8.3.7. *An actuarial risk model.* A classical model for the surplus U_t at time t of an insurance company is the Cramér-Lundberg model (see e.g. [16]) given by

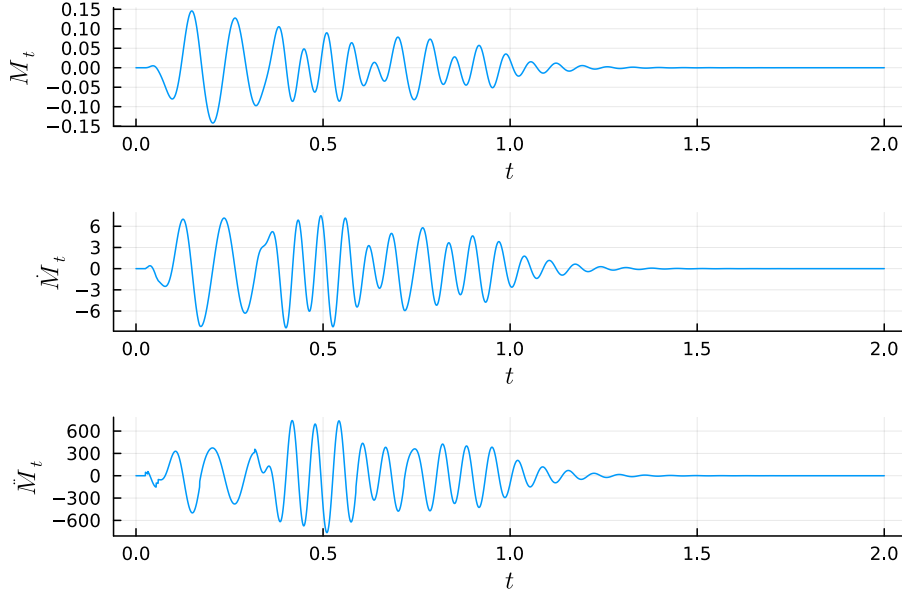$$U_t = U_0 + \gamma t - \sum_{i=1}^{N_t} C_i,$$

FIGURE 19. Sample ground motion, ground velocity, and ground acceleration for the mechanical structure model (8.25).

where U_0 is the initial capital, γ is a constant premium rate received from the insurees, C_i is a random variable representing the value of the i-th claim paid to a given insuree, and N_t is the number of claims up to time t. The process $\{N_t\}_t$ is modeled as a Poisson counter, so that the accumulated claims form a compound Poisson process. It is also common to use inhomogeneous Poisson processes and Hawkes self-exciting process, or combinations of such processes for the incidence of the claim, but the classical model uses a homogeneous Poisson counter.

The model above, however, does not take into account the variability of the premium rate received by the company, nor the investiment of the accumulated reserves, among other things. Several diffusion type models have been proposed to account for these and other factors. We consider here a simple model, with a randomly perturbed premium and with variable rentability.

More precisely, we start by rewriting the above expression as the following jump (or impulse) differential equation

$$\mathrm{d}U_t = \gamma \, \mathrm{d}t - \mathrm{d}C_t,$$

where

$$C_t = \sum_{i=1}^{N_t} C_i.$$

The addition of an interest rate r leads to

$$\mathrm{d}U_t = rU_t\mathrm{d}t + \gamma \, \mathrm{d}t - \mathrm{d}C_t.$$

Assuming a premium rate perturbed by a white noise and assuming the interest rate as a process $\{R_t\}_t$, we find

$$\mathrm{d}U_t = R_t U_t \, \mathrm{d}t + \gamma \, \mathrm{d}t + \varepsilon \, \mathrm{d}W_t - \mathrm{d}C_t,$$

so the equation becomes

$$\mathrm{d}U_t = (\gamma + R_t U_t) \, \mathrm{d}t + \varepsilon \, \mathrm{d}W_t - \mathrm{d}C_t.$$

Since we can compute exactly the accumulated claims C_t, we subtract it from U_t to get rid of the jump term. We also subtract an Ornstein-Uhlenbeck process, in the classical way to transform an SDE into a RODE. So, defining

$$X_t = U_t - C_t - O_t,$$

where $\{O_t\}_t$ is given by

$$\mathrm{d}O_t = -\nu O_t \, \mathrm{d}t + \varepsilon \, \mathrm{d}W_t,$$

we obtain

$$\mathrm{d}X_t = (\gamma + R_t U_t) \, \mathrm{d}t + \nu O_t \, \mathrm{d}t = (\gamma + R_t(X_t + C_t + O_t)) \, \mathrm{d}t + \nu O_t \, \mathrm{d}t.$$

This leads us to the linear random ordinary differential equation

$$\frac{\mathrm{d}X_t}{\mathrm{d}t} = R_t X_t + R_t(C_t + O_t) + \nu O_t + \gamma. \tag{8.27}$$

This equation has the explicit solution

$$X_t = X_0 e^{\int_0^t R_s \, \mathrm{d}s} + \int_0^t e^{\int_s^t R_\tau \, \mathrm{d}\tau}(R_s(C_s + O_s) + \nu O_s + \gamma) \, \mathrm{d}s.$$

As for the interest rate process $\{R_t\}_t$, there is a vast literature with models for it, see e.g. [10, Chapter 3] in particular Table 3.1. Here, we consider the Dothan model ([10, Section 3.2.2] of the aforementioned reference), which consists simply of a geometric Brownian motion process

$$\mathrm{d}R_t = \mu R_t \, \mathrm{d}t + \sigma R_t \, \mathrm{d}t,$$

with $R_t = r_0$, where μ, σ, r_0 are positive constants. This has an explicit solution

$$R_t = r_0 e^{(\mu - \sigma^2/2)t + \sigma W_t},$$

so that the equation (8.27) for $\{X_t\}_t$ is a genuine random ODE.

Once the solution of $\{X_t\}_t$ is obtained, we find an explicit formula for the surplus $U_t = X_t + C_t + O_t$, namely

$$U_t = C_t + O_t + X_0 e^{\int_0^t R_s \, \mathrm{d}s} + \int_0^t e^{\int_s^t R_\tau \, \mathrm{d}\tau}(R_s(C_s + O_s) + \nu O_s + \gamma) \, \mathrm{d}s,$$

with $\{R_t\}_t$ as above.

For the numerical simulations, we use $O_0 = 0.0$, $\nu = 5.0$ and $\varepsilon = 0.8$, for the Ornstein-Uhlenbeck process $\{O_t\}_t$; $\lambda = 8.0$ and $C_i \sim \text{Uniform}(0, 0.2)$, for the compound Poisson process $\{C_t\}$; $R_0 = 0.2$, $\mu = 0.02$ and $\sigma = 0.4$, for the interest rate process $\{R_t\}_t$; and we take $X_0 = 1.0$, so that $U_0 = X_0 + O_0 + R_0 = 1.2$.

N	dt	error	std err
64	0.0469	0.224	0.0164
128	0.0234	0.112	0.00847
256	0.0117	0.0559	0.00397
512	0.00586	0.0295	0.00199

TABLE 8. Mesh points (N), time steps (dt), strong error (error), and standard error (std err) of the Euler method for a risk model for each mesh resolution N, with initial condition $X_0 = 1.0$ and coupled Ornstein-Uhlenbeck, geometric Brownian motion, and compound Poisson processes, on the time interval $I = [0.0, 3.0]$, based on $M = 400$ sample paths for each fixed time step, with the target solution calculated with 262144 points. The order of strong convergence is estimated to be $p = 0.976$, with the 95% confidence interval $[0.8938, 1.0584]$.

Although we have an explicit solution, it is not in closed form, so the target approximation for computing the error is obtained by solving the equation via Euler method in the much finer mesh. Table 8 shows the estimated strong error obtained with the Monte-Carlo method, while Figure 20 illustrates the order of convergence. Figure 21 shows a sample path of the noise, which is composed of three processes, while Figure 22 shows a sample path of the surplus.

8.3.8. *A random Fisher-KPP nonlinear PDE driven by boundary noise.* Finally, we simulate a Fisher-KPP equation with random boundary conditions, as inspired by the works of [34] and [15]. The first work addresses the Fisher-KPP equation with a random reaction coefficient, while the second work considers more general reaction-diffusion equations but driven by random boundary conditions.

The intent here is to illustrate the strong order 1 convergence rate on a discretization of a nonlinear partial differential equation. We use the method of lines (MOL), with finite differences in space, to approximate the random partial differential equation (PDE) by a system of random ODEs.

The deterministic Fisher-KPP equation is a nonlinear parabolic equation of reaction-diffusion type, with origins in [14] and [25]. It models inhomogeneous population growth and many other phenomena displaying wave propagation, such as combustion front wave propagation, physiollogy, crystallography pattern formation, and so on.

We consider the Fisher-KPP equation driven by Neumann boundary conditions, with a random influx on the left end point and no flux on the right end point. The equation takes the form

$$\frac{\partial u}{\partial t} = \mu \frac{\partial^2 u}{\partial x^2} + \lambda u \left(1 - \frac{u}{u_m} \right), \quad (t, x) \in (0, \infty) \times (0, 1), \qquad (8.28)$$
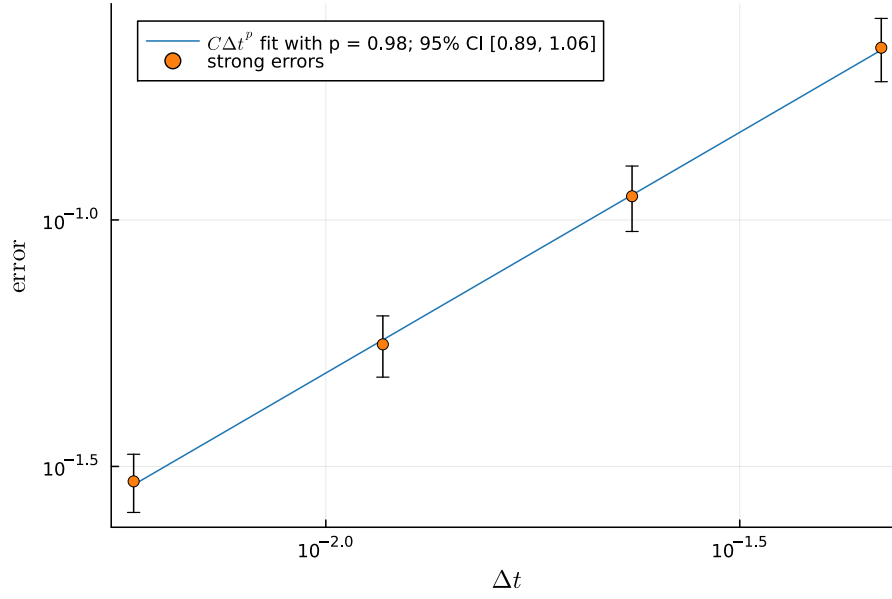
FIGURE 20. Order of convergence of the strong error of the Euler method for the actuarial risk model (8.27), based on Table 8.
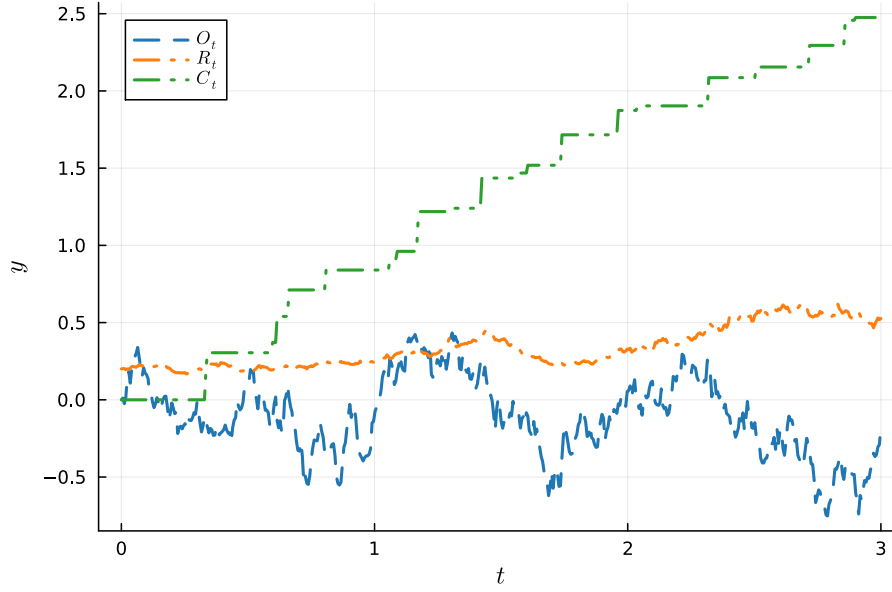

FIGURE 21. Sample noises for the risk model (8.27).

endowed with the boundary conditions

$$\frac{\partial u}{\partial x}(t,0) = -Y_t, \quad \frac{\partial u}{\partial x}(t,1) = 0, \tag{8.29}$$
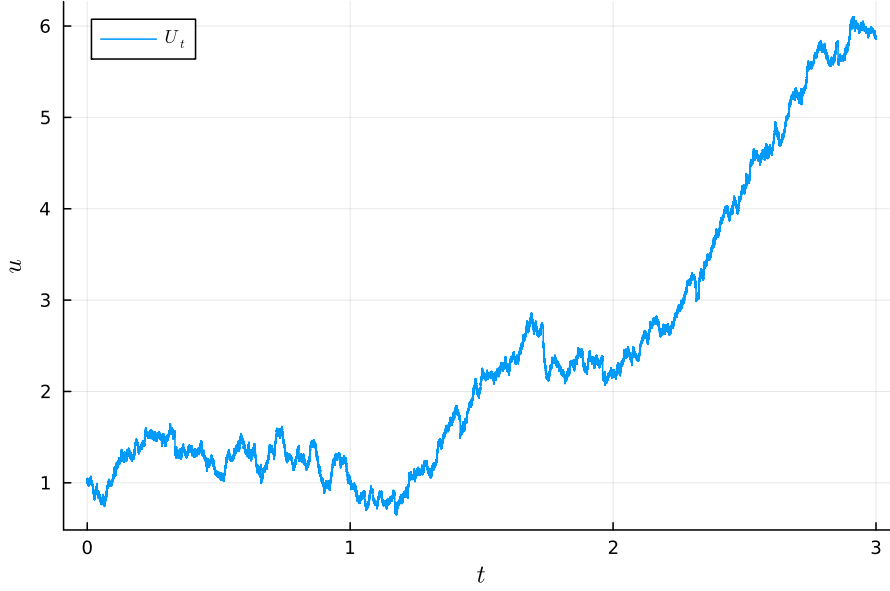
FIGURE 22. Sample surplus solution for the risk model (8.27).

and a given initial condition

$$u(0, x) = u_0(x).$$

The unknown $u(t, x)$ represents the population density at time t and point x; D is a diffusivity coefficient; λ is a reaction, or proliferation, coefficient; and u_m is a carrying capacity density coefficient.

The random process $\{Y_t\}_t$, which drives the flux on the left boundary point, is taken to be a colored noise modulated by an exponentially decaying Hawkes process, representing random wave trains of incoming populations. More precisely, $Y_t = H_t O_t$, where $\{H_t\}_t$ is a Hawkes process with initial rate $\lambda_0 = 5.0$, base rate $a = 1.4$, exponential decay rate $\delta = 8.0$, and jump law following an exponential distribution with scale $\theta = 1/0.4$; and $\{O_t\}_t$ is an Ornstein-Uhlenbeck process with $O_0 = 0.0$, time-scale $\tau = 0.005$, drift $\nu = 1/\tau$ and diffusion $\sigma = \zeta/\tau$, where $\zeta = \sqrt{\tau}$. Since both $\{H_t\}_t$ and $\{O_t\}_t$ are semi-martingales and the class of semi-martingales is an algebra [32, Corollary II.3], the modulated process $\{Y_t\}_t$ is also a semi-martingale.

This equation displays traveling wave solutions with a minimum wave speed of $2\sqrt{\lambda\mu}$. We choose $\lambda = 10$ and $\mu = 0.009$, so the limit traveling speed is about 0.6. The carrying capacity is set to $u_m = 1.0$.

The initial condition is taken to be zero, $u_0(x) = 0$, so all the population originates from the left boundary influx.

The mass within the region $0 \le x \le 1$ satisfies

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_0^1 u(t, x) \, \mathrm{d}x = \mu \int_0^1 u_{xx}(t, x) \, \mathrm{d}x + \lambda \int_0^1 u(t, x) \left(1 - \frac{u(t, x)}{u_m}\right) \, \mathrm{d}x.$$

N	dt	error	std err
32	0.0625	132.0	21.2
128	0.0156	31.5	5.03
512	0.00391	6.01	0.959

TABLE 9. Mesh points (N), time steps (dt), strong error (error), and standard error (std err) of the Euler method for the Fisher-KPP equation for each mesh resolution N, with initial condition $X_0 = 0$ and Hawkes-modulated Ornstein-Uhlenbeck colored noise, on the time interval $I = [0.0, 2.0]$, based on $M = 40$ sample paths for each fixed time step, with the target solution calculated with 262144 points. The order of strong convergence is estimated to be $p = 1.116$, with the 95% confidence interval $[0.9995, 1.2318]$.

Using the boundary conditions, we find that

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_0^1 u(t,x) \, \mathrm{d}x = \mu Y_t + \frac{\lambda}{u_m} \int_0^1 u(t,x) \, (u_m - u(t,x)) \, \mathrm{d}x,$$

which is nonnegative, provided $0 \leq u \leq u_m$ and $Y_t \geq 0$.

The equation involves a nonlinear term which is not globally Lipschitz continuous, but, similiarly to the population dynamics model considered in Section 8.3.4, the region $0 \leq u(t,x) \leq u_m$ is invariant, so that the nonlinear term can be modified outside this region in order to satisfy the required uniform global estimates without affecting the dynamics within this region. The initial condition is chosen to be within this region almost surely. The procedure is the same as that done in Section 8.3.4, so the details are omited.

For the time-mesh parameters, we set $N_{\mathrm{tgt}} = 2^{18}$ and $N_i = 2^5, 2^7, 2^9$. The spatial discretization is done with finite differences, with the number of spatial points depending on the time mesh, for stability and convergence reasons. Indeed, the Von Neumann stability analysis requires that $2\mu\Delta t/\Delta_x^2 \leq 1$. With that in mind, for each $N_i = 2^5, 2^7, 2^9$, we take the $K_i + 1$ spatial points $0 = x_0 < \ldots x_{K_i}$, with $K_i = 2^3$, 2^4, and 2^5, respectively, while for the target solution, we use $K_{\mathrm{tgt}} = 2^9$.

Table 9 shows the estimated strong error obtained with the Monte-Carlo method, while Figure 23 illustrates the order of convergence.

REFERENCES

[1] Charalambos D. Aliprantis and Kim C. Border. *Infinite Dimensional Analysis: a Hitchhiker's Guide.* Springer, Berlin, Heidelberg, 2006.

[2] Y. Asai and P. E. Kloeden. Numerical schemes for random odes with affine noise. *Numer. Algorithms*, 72(1):155–171, 2016.
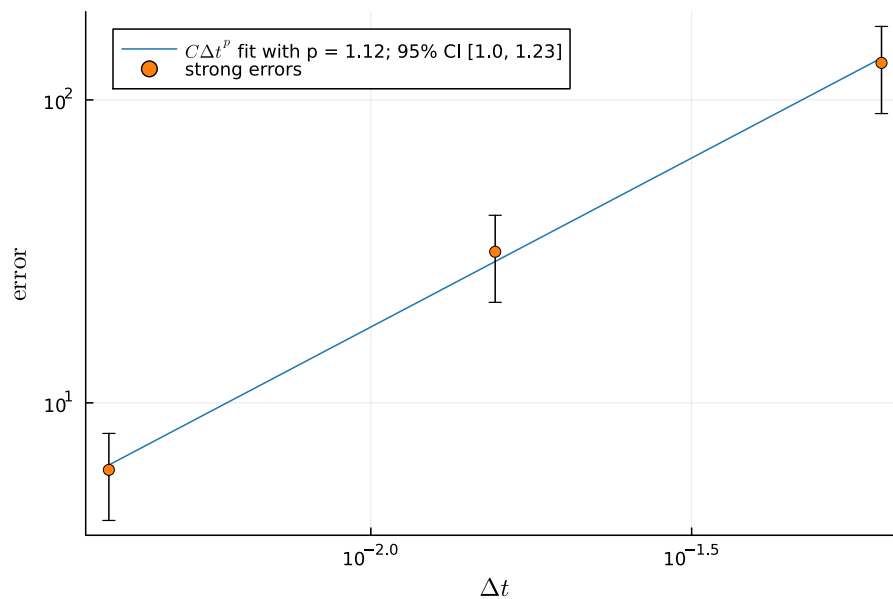
FIGURE 23. Order of convergence of the strong error of the Euler
method for the Fisher-KPP model (8.28)-(8.29), based on Table 9.

[3] Yusuke Asai. *Numerical methods for random ordinary differential equations and their applications in biology and medicine.* doctoral thesis, Johann Wolfgang Goethe University Franfurt am Main, 2016.

[4] Christian Bender. An ItÃ´ formula for generalized functionals of a fractional Brownian motion with arbitrary Hurst parameter. *Stochastic Process. Appl.*, 104(1):81–106, 2003.

[5] Mathieu BesanÃ§on, Theodore Papamarkou, David Anthoff, Alex Arslan, Simon Byrne, Dahua Lin, and John Pearson. Distributions.jl: Definition and modeling of probability distributions in the JuliaStats ecosystem. *Journal of Statistical Software*, 98(16):1–30, 2021.

[6] Jeff Bezanson, Alan Edelman, Stefan Karpinski, and Viral B. Shah. Julia: A fresh approach to numerical computing. *SIAM Rev.*, 59(1):65–98, 2017.

[7] Francesca Biagini, Yaozhong Hu, Bernt Øksendal, and Tusheng Zhang. *Stochastic Calculus for Fractional Brownian Motion and Applications.* Probability and Its Applications. Springer-Verlag, London, 2008.

[8] David Blackman and Sebastiano Vigna. Scrambled linear pseudorandom number generators. *ACM Trans. Math. Software*, 47(4), September 2021.

[9] John L. Bogdanoff, John E. Goldberg, and M. C. Bernard. Response of a simple structure to a random earthquake-type disturbance. *Bulletin of the Seismological Society of America*, 51(2):293–310, 04 1961.

[10] Damiano Brigo and Fabio Mercurio. *Interest-Rate Models: Theory and Practice.* Springer Finance. Springer-Verlag, Berlin, Heidelberg, 2001.

[11] Simon Christ, Daniel Schwabeneder, Christopher Rackauckas, Michael Krabbe Borregaard, and Thomas Breloff. Plots.jl – a user extendable plotting API for the Julia programming language. *Journal of Open Research Software*, 2023.

[12] A. Coddington and N. Levinson. *Theory of Ordinary Differential Equations.* International series in pure and applied mathematics. R.E. Krieger, 1984.

[13] A.B. Dieker and M. Mandjes. On spectral simulation offFractional Brownian motion. *Probab. Engrg. Inform. Sci.*, 17(3):417–434, 2003.

[14] R. A. Fisher Sc. D., F. R. S. The wave of advance of advantageous genes. *Annals Eugen.*, 7(4):355–369, 1937.

[15] Mark I. Freidlin and Alexander D. Wentzell. Reaction-diffusion equations with randomly perturbed boundary conditions. *Ann. Probab.*, 20(2):963 – 986, 1992.

[16] Hans U. Gerber and Elias S.W. Shiu. On the time value of ruin. *N. Am. Actuar. J.*, 2(1):48–72, 1998.

[17] L. Grüne and P. E. Kloeden. Higher order numerical schemes for affinely controlled nonlinear systems. *Numer. Math.*, 89(4):669–690, 2001.

[18] Xiaoying Han and Peter E. Kloeden. *Random Ordinary Differential Equations and Their Numerical Solution.* Springer Publishing Company, Incorporated, 1st edition, 2017.

[19] Desmond Higham and Peter Kloeden. *An Introduction to the Numerical Simulation of Stochastic Differential Equations.* Society for Industrial and Applied Mathematics, Philadelphia, PA, 2021.

[20] George W. Housner and Paul C. Jennings. Generation of artificial earthquakes. *Journal of the Engineering Mechanics Division*, 90(1):113–150, 1964.

[21] Arnulf Jentzen and Peter E. Kloeden. *Taylor Approximations for Stochastic Partial Differential Equations.* Society for Industrial and Applied Mathematics, 2011.

[22] K. Kanai. Semi-empirical formula for the seismic characteristics of the ground. *Bull. Earthq. Res. Inst.*, 35:309–325, 1957.

[23] 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, 2024.

[24] Peter E Kloeden and Arnulf Jentzen. Pathwise convergent higher order numerical schemes for random ordinary differential equations. *Proc. Roy. Soc. Edinburgh Sect. A.*, 463:2229–2944, 2007.

[25] 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(6):1–25, 1937.

[26] Benoit B. Mandelbrot and John W. Van Ness. Fractional Brownian motions, fractional noises and applications. *SIAM Rev.*, 10(4):422–437, 1968.

[27] Mehmet Merdan, Zafer Bekiryazici, Tulay Kesemen, and Tahir Khaniyev. Comparison of stochastic and random models for bacterial resistance. *Adv. Difference Equ.*, 2017(1):133, 2017.

[28] Michel Métivier. *Semimartingales.* De Gruyter, Berlin, New York, 1982.

[29] Y. S. Mishura. *Stochastic calculus for fractional Brownian motion and related processes.* Lecture Notes in Mathematics 1929. Springer-Verlag, Berlin, Heidelberg, 2011.

[30] Tobias Neckel and Florian Rupp. *Random Differential Equations in Scientific Computing.* De Gruyter Open Poland, Versita Ltd, 78 York Street, London W1H 1DP, Great Britain., 2013.

[31] Bernt Øksendal. *Stochastic Differential Equations : An Introduction with Applications.* Springer-Verlag, Berlin, Heidelberg, 2003.

[32] Philip E. Protter. *Stochastic Integration and Differential Equations.* Springer-Verlag, Berlin, Heidelberg, 2 edition, 2005.

[33] Christopher Rackauckas and Qing Nie. DifferentialEquations.jl – A performant and feature-rich ecosystem for solving differential equations in Julia. *The Journal of Open Research Software*, 5(1), 2017.

[34] Rachidi B. Salako and Wenxian Shen. Long time behavior of random and nonautonomous Fisher–KPP equations: Part I—stability of equilibria and spreading speeds. *J. of Dynam. Differential Equations*, 33(2):1035–1070, 2021.

[35] Michael Strasser, Fabian J. Theis, and Carsten Marr. Stability and multiattractor dynamics of a toggle switch based on a two-stage model of stochastic gene expression. *Biophysical Journal*, 102(1):19–29, 2012.

[36] Shohei Sugiyama. Stability problems on difference and functional-differential equations. *Proceedings of the Japan Academy*, 45(7):526–529, 1969.

[37] H Tajimi. A statistical method of determining the maximum response of a building structure during an earthquake. *Proceedings of the 2nd World Conference on Earthquake Engineering, Tokyo, Japan, 1960*, pages 781–797, 1960.

[38] Berta Verd, Anton Crombach, and Johannes Jaeger. Classification of transient behaviours in a time-dependent toggle switch model. *BMC Systems Biology*, 8(1):43, 2014.

[39] Peng Wang, Yanzhao Cao, Xiaoying Han, and Peter Kloeden. Mean-square convergence of numerical methods for random ordinary differential equations. *Numer. Algorithms*, 87(1):299–333, 2021.

(Peter E. Kloeden) MATHEMATICS DEPARTMENT, UNIVERSITY OF TÜBINGEN, GERMANY

(Ricardo M. S. Rosa) INSTITUTO DE MATEMÁTICA, UNIVERSIDADE FEDERAL DO RIO DE JANEIRO, BRAZIL

Email address, P. E. Kloeden: `kloeden@math.uni-frankfurt.de`

Email address, R. M. S. Rosa: `rrosa@im.ufrj.br`