.

IMPROVED ERROR ESTIMATES FOR THE ORDER OF 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 approximating the solutions of a random ordinary differential equation (RODE) $\mathrm{d}X_t/\mathrm{d}t = f(t, X_t, Y_t)$ 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 increases to 1 in some special cases, such as that of an Itô diffusion noise, in which case the Euler for RODE can be seen as a particular case of the Milstein method for an associated augmented system of stochastic differential equations. 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. This applies not only to yield a direct proof of order 1 convergence for Itô process noises, but also to point-process noises, transport-type processes with sample paths of bounded variation, time-changed Brownian motion, and any other noise that is a semi-martingale. The result follows from estimating the global error as an iterated integral over both large and small mesh scales, and then by switching the order of integration to move the critical regularity to the large scale. The work is complemented with numerical simulations illustrating the strong order 1 convergence in those cases, and with an example with fractional Brownian motion noise with Hurst parameter $0 < H < 1/2$ for which the order of convergence is $H + 1/2$, hence lower than the attained order 1 in the semi-martingale examples above, but still higher than the order H of convergence expected from previous works.

1. Introduction

Consider the following initial value problem for a **random ordinary differential equation (RODE)**:

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

Date: March 30, 2024.

2020 *Mathematics Subject Classification.* 60H35 (Primary), 65C30, 34F05 (Secondary).

Key words and phrases. random ordinary differential equations, Euler method, strong convergence, Itô process, point process, fractional Brownian motion.

The second author was partly supported by the Laboratório de Matemática Aplicada, Instituto de Matemática, Universidade Federal do Rio de Janeiro (LabMA/IM/UFRJ) and the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Brasil, grant 001.

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 and independent of each other. The function $f = f(t, X_t, Y_t)$ 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.

The Euler method for solving this initial value problem consists in approximating the solution on a uniform time mesh $t_j = j\Delta t_N$, $j = 0, \dots, N$, with fixed time step $\Delta t_N = T/N$, for a given $N \in \mathbb{N}$. In such a mesh, the Euler scheme takes the form

$$\begin{cases} X^N_{t_j} = X^N_{t_{j-1}} + \Delta t_N f(t_{j-1}, X^N_{t_{j-1}}, Y_{t_{j-1}}), \quad j = 1, \dots, N, \\ X^N_0 = X_0. \end{cases} \tag{1.2}$$

Notice $t_j = j\Delta t_N = jT/N$ also depends on N, but we do not make this dependency explicit, for the sake of notational simplicity.

We are mostly interested in the order of *strong* convergence, i.e. the approximation $\{X^N_{t_j}\}_j$ is said to converge to $\{X_t\}_t$ with *strong order* $\theta > 0$ when there exists a constant $C \geq 0$ such that

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

where $\mathbb{E}[\cdot]$ indicates the expectation of a random variable on Ω, and $\|\cdot\|$ is the norm in the appropriate phase space. The strong convergence is considered in the case of an arbitrary semi-martingale noise. For the particular case of a finite-variation noise (FV process), we also consider the *pathwise convergence of order* θ, where

$$\max_{j=0,\dots,N} \left\|X_{t_j} - X^N_{t_j}\right\| \leq C\Delta t_N^\theta, \qquad \forall N \in \mathbb{N}, \tag{1.4}$$

almost surely, for a random variable $C = C(\omega)$ on Ω.

There are other important notions of convergence, such as weak convergence, mean-square convergence, and p-th mean convergence (see e.g. [19, 20, 22]), but here we focus only on the strong and pathwise convergence.

Under certain regularity conditions on f, it is proved in [39, Theorem 3] that, when the noise $\{Y_t\}_{t\in I}$ has θ-Hölder continuous sample paths, the Euler scheme converges to the exact solution in the mean square sense with order θ with respect to the time step. This implies the strong convergence (1.3) with the same order θ. For pathwise convergence, see e.g. [18, 25, 22, 3, 19]. In the case of fractional Brownian motion noise with Hurst parameter H, it is proved in [39, Theorem 2] that the mean square convergence is of order H, for any $0 < H \leq 1$.

In some special cases, the order of convergence of the Euler method may be higher than the associated Hölder exponent of the noise, such as for Itô diffusion noises, in which case the RODE can be seen as part of a system of stochastic differential equations (SDE), with the Euler method for the RODE part of the system being part of a special case of the Milstein method for the SDE, which is known to be of strong

order one (see [39, Section 5], in particular [39, Section 5.2, Example 12 and Remark 13]).

Our aim is to show that, in many more classical examples, it is possible to exploit further conditions that yield in fact a higher strong order of convergence, with the sample noise paths still being Hölder continuous or even discontinuous. This is the case, for instance, when the noise is a point process, a transport process, an Itô process or a time-changed Brownian motion. We, show, in fact, that this is true for any noise that is a semi-martingale with bounded square variation.

We also consider a particular equation with fractional Brownian motion noise with Hurst parameter H, for which the sample paths are H-Hölder continuous, but the strong convergence is of order 1, when $1/2 \leq H < 1$, and of order $H + 1/2$, when $0 < H < 1/2$.

The global condition on f is a natural assumption when looking for strong convergence. Pathwise convergence, on the other hand, usually requires less stringent conditions (see e.g. [23, 22]), as seen in the FV process case.

The first main idea 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, namely

$$\begin{aligned} X_{t_j} - X^N_{t_j} = X_0 - X^N_0 &+ \int_0^{t_j} \left(f(s, X_s, Y_s) - f(s, X_{\tau^N(s)}, Y_s)\right) \, \mathrm{d}s \\ &+ \int_0^{t_j} \left(f(s, X_{\tau^N(s)}, Y_s) - f(s, X^N_{\tau^N(s)}, Y_s)\right) \, \mathrm{d}s \\ &+ \int_0^{t_j} \left(f(s, X^N_{\tau^N(s)}, Y_s) - f(\tau^N(s), X^N_{\tau^N(s)}, Y_{\tau^N(s)})\right) \, \mathrm{d}s, \end{aligned} \tag{1.5}$$

for $j = 1, \ldots, N$, where $\tau^N(t) = \max_{t_j \leq t} t_j$ is a piecewise constant function jumping to the mesh points t_j (see (5.2)).

The first term vanishes due to the initial condition $X^N_0 = X_0$. The second term only depends on the solution and can be estimated with natural regularity conditions on the term $f = f(t, x, y)$. The third term is handled solely with the typical required condition on $f = f(t, x, y)$ of being globally Lipschitz continuous with respect to x. With those, we obtain the following basic bound for the global error (see Lemma 6.2)

$$\begin{aligned} \|X_{t_j} - X^N_{t_j}\| \leq \Bigg(&\|X_0 - X^N_0\| + \ell_x \int_0^{t_j} \|X_s - X_{\tau^N(s)}\| \, \mathrm{d}s \\ &\left\| \int_0^{t_j} \left(f(s, X^N_{\tau^N(s)}, Y_s) - f(\tau^N(s), X^N_{\tau^N(s)}, Y_{\tau^N(s)})\right) \, \mathrm{d}s \right\| \Bigg) e^{\ell_x t_j}. \end{aligned} \tag{1.6}$$

The only problematic, noise-sensitive term is the last one. The classical analysis is to use an assumed θ-Hölder regularity of the noise sample paths and estimate the

local error as

$$\mathbb{E}\left[\left\|f(s,X^N_{\tau^N(s)},Y_s)-f(\tau^N(s),X^N_{\tau^N(s)},Y_{\tau^N(s)})\right\|\right]\leq C\Delta t_N^\theta.$$

Instead, we look at the whole noise error

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

and assume that the steps of the process given by $F_t=f(t,X^N_{\tau^N(t)},Y_t)$ 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

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

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

$$\begin{aligned}&\mathbb{E}\left[\left\|\int_0^{t_j}\left(f(s,X^N_{\tau^N(s)},Y_s)-f(\tau^N(s),X^N_{\tau^N(s)},Y_{\tau^N(s)})\right)\,\mathrm{d}s\right\|\right]\\&\qquad\leq\Delta t_N\mathbb{E}\left[\left\|\int_0^{t_j}\mathrm{d}F_\xi\right\|\right].\end{aligned}$$

We consider two cases, one exclusive for FV processes (whose paths are almost surely of finite variation and càdlàg, i.e. right continuous with left limits), for which we prove the pathwise convergence of order 1, under less stringent conditions on $f=f(t,x,y)$, and a more general case of semi-martingale noises (which can be characterized as a sum of an FV process and a local martingale), for which we prove the strong convergence of order 1. These two cases are treated, respectively, in Section 7 and Section 8, with the main results given in Theorem 7.1 and Theorem 8.1.

We complement this work with a number of explicit examples and their numerical implementation, illustrating the strong order 1 convergence in the cases above. We include, for instance, a system of linear equations with all sorts of noises, encompassing noises with sample paths with bounded variation and Itô process noises. We also

include an example with a fractional Brownian motion noise (fBm), for which the order of convergence drops to $H + 1/2$, when the Hurst parameter is in the range $0 < H < 1/2$. We do not present a general proof of this order of convergence in the case of fBm noise, but we prove it here in a particular linear equation. In this example, we essentially have (see (9.11) and (9.13))

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

In this case, disregarding the higher order term,

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

which, upon taking the expectation of the absolute value, yields a strong convergence of order $H + 1/2$.

Many other examples are included and also presented in more details in the github repository [34], such as a logistic model of population dynamics with random coefficients, loosely inspired by [19, Section 15.2], where the specific growth is the sine of a geometric Brownian motion process and with an extra point-process random term representing harvest; a toggle-switch model of gene expression (similar to [2, Section 7.8], originated from [38], see also [36]) driven by a combination of a compound Poisson point process and an Itô process, illustrating, again, the two main types of noises considered here; a mechanical structure model driven by a random disturbance simulating seismic ground-motion excitations in the form of a transport process, inspired by the Bogdanoff-Goldberg-Bernard model in [7] (see also [30, Chapter 18], [21], and [24], with this and other models, such as the ubiquotous Kanai-Tajimi and Clough-Penzien colored-noise models); an actuarial risk model for the surplus of an insurance company, inspired by [14] and [8]; and a Fisher-KPP partial differential equation with random boundary conditions, as inspired by the works of [35] and [13] (see also [12] and [26]), and where the noise is a colored noise modulated by a decaying self-exciting Hawkes point process.

2. Pathwise solutions

For the notion and main results on pathwise solutions for RODEs, we refer the reader to [19, 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 $\{Y_t\}_{t\in I}$, where $I = [0,T]$, $T > 0$, and $d, k \in \mathbb{N}$. We make the following standing assumptions:*

(i) *The mapping $x \mapsto f(t,x,Y_t)$ is almost surely globally Lipschitz continuous on x, for every $t \in I$, with*

$$\|f(t,x_1,Y_t) - f(t,x_2,Y_t)\| \leq L_t\|x_1 - x_2\|, \quad \forall t \in I,\ \forall x_1, x_2 \in \mathbb{R}^d, \tag{2.1}$$

for a non-negative stochastic process $\{L_t\}_{t\in I}$ with Lebesgue integrable sample paths $t \mapsto L_t(\omega)$ on I.

(ii) *The mapping $t \mapsto f(t,x,Y_t(\omega))$ is Lebesgue measurable on I, for every $x \in \mathbb{R}^d$ and almost every sample path $t \mapsto Y_t(\omega)$;*

(iii) *The bound $\|f(t,0,Y_t)\| \leq M_t$ holds for all $t \in I$, where $\{M_t\}_{t\in I}$ is a non-negative stochastic process with Lebesgue integrable sample paths $t \mapsto M_t(\omega)$ on I.*

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 $t \mapsto X_t(\omega)$, in the Lebesgue sense, for the realizations $X_0 = X_0(\omega)$ of the initial condition and $t \mapsto Y_t(\omega)$ of the noise process (see [10, Theorem 1.1]). Moreover, the mapping $(t,\omega) \mapsto X_t(\omega)$ is (jointly) measurable (see [19, Section 2.1.2] and also [1, Lemma 4.51]) and, hence, give rise to a well-defined stochastic process $\{X_t\}_{t\in I}$.

Each pathwise solution of (2.2) is absolutely continuous, and, from assumptions (i) and (iii), the solutions satisfy almost surely

$$\|X_t\| \leq \|X_0\| + \int_0^t (M_t + L_t\|X_t\|)\, \mathrm{d}s.$$

Using Grownwall's inequality and the assumptions that $\{M_t\}_{t\in I}$ and $\{L_t\}_{t\in I}$ are integrable almost surely, it follows that

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

More generally, since $t \mapsto X_t$ is absolutely continuous almost surely, we take the derivative of $\|X_t\|^\beta$ for an arbitrary power $\beta \geq 1$ to find that

$$\|X_t\|^\beta = \|X_0\|^\beta + q \int_0^t \|X_s\|^{\beta-2} X_s \cdot f(t, X_s, Y_s)\, \mathrm{d}s.$$

Using again (i) and (iii), we obtain

$$\begin{aligned}\|X_t\|^\beta &\leq \|X_0\|^\beta + \beta \int_0^t \|X_s\|^{\beta-1} \left(M_s + L_s \|X_s\|\right)\, \mathrm{d}s \\ &\leq \|X_0\|^\beta + \beta \int_0^t \left(\frac{1}{\beta} M_s^\beta + \frac{\beta-1}{\beta} \|X_s\|^\beta + L_s \|X_s\|^\beta\right)\, \mathrm{d}s.\end{aligned}$$

By the Grownwall Lemma, the solutions satisfy almost surely

$$\|X_t\|^\beta \leq \left(\|X_0\|^\beta + \int_0^t M_s^\beta\, \mathrm{d}s\right) e^{\int_0^t (\beta - 1 + \beta L_s)\, \mathrm{d}s}. \tag{2.4}$$

This is a useful pathwise estimate. For a strong norm estimate, though, we need to assume more, as illustrated by the following result.

Lemma 2.1. *Under the Standing Hypothesis 2.1, suppose further that*

$$\mathbb{E}[\|X_0\|] < \infty \tag{2.5}$$

and

$$\int_0^T \mathbb{E}[M_s]\, \mathrm{d}s < \infty, \tag{2.6}$$

and that $\{L_t\}_{t \in I}$ is essentially bounded, uniformly on I, i.e. there exists a constant $\ell_x > 0$ such that

$$L_t \leq \ell_x, \quad t \in I, \tag{2.7}$$

almost surely. Then

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

If, moreover,

$$\mathbb{E}[\|X_0\|^\beta] < \infty, \quad \int_0^T \mathbb{E}[M_s^\beta]\, \mathrm{d}s < \infty, \tag{2.9}$$

for a given $\beta \geq 1$, then,

$$\mathbb{E}[\|X_t\|^\beta] \leq \left(\mathbb{E}[\|X_0\|^\beta] + \int_0^t \mathbb{E}[M_s^\beta]\, \mathrm{d}s\right) e^{((\beta-1) + \beta \ell_x) t}, \quad t \in I. \tag{2.10}$$

Proof. Thanks to (2.3) and (2.4), the result is straightfoward. □

Remark 2.1. When $f = f(t, x, y)$ is continuous on all three variables, as well as uniformly globally Lipschitz continuous in x, and the sample paths of $\{Y_t\}_{t\geq 0}$ are almost surely continuous except on a set of Lebesgue measure zero, then the integrand in (2.2) is almost surely continuous in t and the integral becomes a Riemann integral. In this case, the integral form (2.2) of the pathwise solutions of (1.1) holds in the Riemann sense.

Remark 2.2. In special *dissipative* cases, depending on the structure of the equation, we might not need the second condition (2.6) and only require $\mathbb{E}[\|X_0\|] < \infty$. More generally, 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. We leave these cases to be handled in the applications, such as the population dynamics example in Section 9.4.

3. Pathwise approximation

Similarly, under the Standing Hypothesis 2.1, we may obtain estimates for the Euler approximation (1.2). Notice that

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

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

$$\|X^N_{t_j}\| \leq e^{L_{t_{j-1}} \Delta t_N} \|X^N_{t_{j-1}}\| + \Delta t_N M_{t_{j-1}}.$$

Iterating this relation we obtain the following bound analogous to (2.3).

$$\|X^N_{t_j}\| \leq \left(\|X^N_0\| + \Delta t_N \sum_{i=0}^{j-1} M_{t_i} \right) e^{\Delta t_N \sum_{i=0}^{j-1} L_{t_i}}, \quad \forall j = 0, \ldots, N. \tag{3.1}$$

4. Semi-martingale noises

For the theory of semi-martingales, we refer the reader to [32]. Thanks to the Bichteler-Dellacherie Theorem [32, Theorem III.47], a *semi-martingale* is a càdlàg process $\{Y_t\}_{t\in I}$ characterized by being decomposable as a sum $Y_t = F_t + Z_t$ of an FV process $\{F_t\}_{t\in I}$ and a local martingale process $\{Z_t\}_{t\in I}$.

A *càdlàg process* is a stochastic process $\{Y_t\}_{t\in I}$ whose sample paths are almost surely càdlàg paths. 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). Thus, $\{Y_t\}_{t\in I}$ is càglàg process when for almost every $\omega \in \Omega$ and for all $t \in I$, we have that $\lim_{s\to t^+} Y_s(\omega) = Y_t(\omega)$ (right continuous) and the limit $\lim_{s\to t^-} Y_s(\omega)$ exists (left limit).

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

$$V(\{Y_t\}_{t\in I}; I) = \sup_{t_0<\ldots<t_n\in I, n\in\mathbb{N}} \sum_{i=1}^{n} \|Y_{t_i} - Y_{t_{i-1}}\| < \infty.$$

For a càdlàg process such as semi-martingales, we define the left-limit process

$$Y_{t^-} = \lim_{t\to t^-} Y_s \tag{4.1}$$

and the jump process

$$\Delta Y_t = Y_t - Y_{t^-}. \tag{4.2}$$

Almost surely, the sample paths of the jump process are zero except possibly at a countable number of points, denoted J, i.e. for almost every $\omega \in \Omega$, there exists a countable set $J(\omega) \subset I$ for which $\Delta Y_t(\omega)$ is zero on each connected component of $I \setminus J(\omega)$.

In the case of finite variation, we also have

$$\sum_{s\in J} \|\Delta Y_t\| \le V(\{Y_t(\omega)\}_{t\in I}; I) < \infty. \tag{4.3}$$

In the more general case of semi-martingale, then $\{Y_t\}_{t\in I}$ has finite quadratic variation [32, Section II.6] and we find that

$$\sum_{s\in J} \|\Delta Y_t\|^2 < \infty.$$

In any case, 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.

A fundamental tool for integration and differentiation is the chain rule. In the case of an FV process, given a continuously differentiable function $f : \mathbb{R}^k \to \mathbb{R}^k$, the following change of variables formula holds (see [32, Theorems II.31 and II.33]):

$$f(Y_t) - f(Y_0) = \int_{0^+}^{t} Df(Y_{s^-})\, \mathrm{d}Y_s + \sum_{0<s\le t} \left(f(Y_s) - f(Y_{s^-}) - Df(Y_{s^-})\Delta Y_s\right). \tag{4.4}$$

The summation on the right hand side is an at most countable summation since the summand vanishes for $s \notin J$. The summation can be written as an integral with respect to a jump measure, 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 $f(y) = (f_i(y_1, \ldots, y_k))_{i=1,\ldots,k}$ and $Y_t = ((Y_t)_i)_{i=1,\ldots,k}$, and

$$Df(y)\Delta Y_t = (\nabla f_i(y)\Delta Y_t)_{i=1,\ldots,k} = \left(\sum_{j=1,\ldots,k} \partial y_j f_i(y)(Y_t)_j\right)_{i=1,\ldots,k}.$$

When $f = f(t, y)$ depends also on t, one can apply the above formula to the extended process $\{(t, Y_t)\}_{t\in I}$. The first component is continuous hence has no jumps

and the corresponding summation term vanishes. This leads to the following change of variables formula for FV processes:

$$f(t,Y_t)-f(0,Y_0)=\int_0^t \partial_s f(s,Y_{s^-})\,\mathrm{d}s+\int_{0^+}^t D_y f(s,Y_{s^-})\,\mathrm{d}Y_s + \sum_{0<s\le t}\left(f(s,Y_s)-f(s,Y_{s^-})-D_y f(s,Y_{s^-})\Delta Y_s\right). \quad (4.5)$$

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

The class of FV processes include the compound Poisson process, inhomogeneous Poisson process, Hawkes self-exciting process, and so on. 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 càdlág function of finite variation, for every sample value $y_0=Y_0(\omega)$.

In the more general case of a semi-martingale noise $\{Y_t\}_{t\in I}$, the change of variables formula (4.5) takes the form (see [32, Theorems II.32 and II.33])

$$f(t,Y_t)-f(0,Y_0)=\int_0^t \partial_s f(s,Y_{s^-})\,\mathrm{d}s+\int_{0^+}^t D_y f(s,Y_{s^-})\,\mathrm{d}Y_s + \sum_{0<s\le t}\left(f(s,Y_s)-f(s,Y_{s^-})-D_y f(s,Y_{s^-})\Delta Y_s\right) + \frac{1}{2}\int_0^t D_{yy} f(s,Y_{s^-})\,\mathrm{d}[Y,Y]^c_t, \quad (4.6)$$

where $\{[Y,Y]^c_t\}_{t\in I}$ is the continuous part of the quadratic variation process $\{[Y,Y]_t\}_{t\in I}$ defined by

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

The quadratic variation is a cáglág, increasing, and adapted process [32, Theorem II.22]. The continuous part is given by

$$[Y,Y]^c_t=[Y,Y]_t-\sum_{0\le s\le t}(\Delta X_s)^2.$$

Since $[Y,Y]^c_t$ is continuous and increasing, the corresponding integral is a Lebesgue-Stieltjes integral.

The class of local martingales include Itô diffusion processes and time-changed Brownian motion, for example. For Itô diffusion, see also [31].

5. Integral formula for the global pathwise error

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

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

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

for $j = 1, \ldots, N$, where τ^N is the piecewise constant jump function along the time mesh:

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

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

The exact pathwise solution satisfies

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

The Euler step is given by $X^N_{t_j} = X^N_{t_{j-1}} + \Delta t_N f(t_{j-1}, X^N_{t_{j-1}}, Y_{t_{j-1}})$. Subtracting, we obtain

$$X_{t_j} - X^N_{t_j} = X_{t_{j-1}} - X^N_{t_{j-1}} + \int_{t_{j-1}}^{t_j} \left(f(s, X_s, Y_s) - f(t, X^N_t, Y_t)\right)\, \mathrm{d}s.$$

Adding and subtracting appropriate terms yield

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

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

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

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

6. Basic estimate for the global pathwise error

Here we derive an estimate that is the basis for the specific estimates for each type of noise. For that, we use the following discrete version of the Grownwall Lemma, which is a particular case of the result found in [16] (see also [9]). Its proof follows from [16, Lemma V.2.4] by taking $n = j$, $a_n = e_j$, $b_n = 0$, $c_n = b$, and $\lambda = a$.

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*

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

for every j, with $e_0 = 0$, and where $a_i, b \geq 0$. Then,

$$e_j \leq b e^{\sum_{i=0}^{j-1} a_i}, \qquad \forall j. \tag{6.2}$$

With that, we are ready to start proving our basic estimate for the global pathwise error.

Lemma 6.2. *Under the Standing Hypothesis 2.1, the global error (5.1) is estimated as*

$$\begin{aligned} \|X_{t_j} - X_{t_j}^N\| \leq \Bigg(& \|X_0 - X_0^N\| + \int_0^{t_j} L_s \|X_s - X_{\tau^N(s)}\| \, \mathrm{d}s \\ & \left\| \int_0^{t_j} \left(f(s, X_{\tau^N(s)}^N, Y_s) - f(\tau^N(s), X_{\tau^N(s)}^N, Y_{\tau^N(s)}) \right) \, \mathrm{d}s \right\| \Bigg) e^{\int_0^{t_j} L_s \, \mathrm{d}s}, \end{aligned} \tag{6.3}$$

for $j = 1, \ldots, N$, where τ^N is given by (5.2).

Proof. We estimate the first two integrals in (5.1). For the first one, we use (2.1), so that

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

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

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

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

$$\| f(s, X_t, Y_s) - f(s, X_t^N, Y_s) \| \leq L_s \| X_t - X_t^N \|,$$

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

$$\begin{aligned} \left\| \int_0^{t_j} \left(f(s, X_{\tau^N(s)}, Y_s) - f(s, X^N_{\tau^N(s)}, Y_s) \right) \, \mathrm{d}s \right\| &\leq \int_0^{t_j} L_s \| X_{\tau^N(s)} - X^N_{\tau^N(s)} \| \, \mathrm{d}s \\ &\leq \sum_{i=0}^{j-1} \left(\int_{t_i}^{t_{i+1}} L_s \, \mathrm{d}s \right) \| X_{t_i} - X_{t_i}^N \|. \end{aligned}$$

With these two estimates, we bound (5.1) as

$$\begin{aligned} \| X_{t_j} - X_{t_j}^N \| \leq \| X_0 - X_0^N \| &+ \int_0^{t_j} L_s \| X_s - X_{\tau^N(s)} \| \, \mathrm{d}s \\ &+ \sum_{i=0}^{j-1} \left(\int_{t_i}^{t_{i+1}} L_s \, \mathrm{d}s \right) \| X_{t_i} - X_{t_i}^N \| \\ &+ \left\| \int_0^{t_j} \left(f(s, X^N_{\tau^N(s)}, Y_s) - f(\tau^N(s), X^N_{\tau^N(s)}, Y_{\tau^N(s)}) \right) \, \mathrm{d}s \right\|. \end{aligned}$$

This can be cast in the form of (6.1). Then, using the discrete Gronwall Lemma 6.1, we obtain (6.3). □

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

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

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

Lemma 6.3. *Under the Standing Hypothesis 2.1 and assuming moreover that $t \mapsto L_t$ is bounded almost surely, we find that*

$$\int_0^{t_j} L_s \left\| X_s - X_{\tau^N(s)} \right\| \, \mathrm{d}s \leq \Delta t_N \left(\sup_{0 \leq t \leq T} L_t \right) \int_0^{t_j} (M_s + L_s \| X_s \|) \, \mathrm{d}s. \tag{6.4}$$

Proof. By assumptions (i) and (iii) of the Standing Hypothesis 2.1, we have almost surely that

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

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

$$\begin{aligned}\int_0^{t_j} L_s \left\|X_s - X_{\tau^N(s)}\right\|\,\mathrm{d}s &\leq \int_0^{t_j} L_s \int_{\tau^N(s)}^{s} (M_\xi + \ell_x\|X_\xi\|)\,\mathrm{d}\xi\,\mathrm{d}s \\ &= \int_0^{t_j} (M_\xi + L_\xi\|X_\xi\|) \int_{\xi}^{\tau^N(\xi)+\Delta t_N} L_s\,\mathrm{d}s\,\mathrm{d}\xi.\end{aligned}$$

Using that $t \mapsto L_t$ is bounded almost surely, we find that

$$\int_0^{t_j} L_s \left\|X_s - X_{\tau^N(s)}\right\|\,\mathrm{d}s \leq \int_0^{t_j} (\tau^N(\xi) + \Delta t_N - \xi) \left(\sup_{0\leq t\leq T} L_t\right) (M_\xi + L_s\|X_\xi\|)\,\mathrm{d}\xi$$

Using that $\tau^N(\xi) \leq \xi$ and that the remaining terms are nonnegative, we have $\tau^N(\xi) + \Delta t_N - \xi \leq \Delta t_N$, and we obtain exactly (6.4). □

Combining the two previous results we obtain the following:

Proposition 6.1. *Under the Standing Hypothesis 2.1, suppose further that (2.5), (2.6) and (2.7) hold and that, for some constant $C_0 \geq 0$,*

$$\mathbb{E}[\|X_0 - X_0^N\|] \leq C_0\Delta t_N, \qquad N \in \mathbb{N}. \tag{6.5}$$

Then, for every $j = 0, \ldots, N$,

$$\begin{aligned}\mathbb{E}\left[\|X_{t_j} - X_{t_j}^N\|\right] \leq \Bigg(&C_0\Delta t_N + \Delta t_N \ell_x \left(\mathbb{E}[\|X_0\|] + \int_0^{t_j} \mathbb{E}[M_\xi]\,\mathrm{d}\xi\right) e^{\ell_x t_j} \\ &\mathbb{E}\left[\left\|\int_0^{t_j} \left(f(s, X^N_{\tau^N(s)}, Y_s) - f(\tau^N(s), X^N_{\tau^N(s)}, Y_{\tau^N(s)})\right)\,\mathrm{d}s\right\|\right]\Bigg) e^{\ell_x t_j}.\end{aligned} \tag{6.6}$$

Proof. Estimate (6.6) is obtained by taking the expectation of (6.3) in Lemma 6.2, using the assumption (2.7) to bound the exponential term, and properly estimating the first two terms inside the parentheses. The first term is handled with the assumption (6.5). We need some more work to take care of the second term.

Under the Standing Hypothesis 2.1, together with (2.7), estimate Lemma 6.3 applies and inequality (6.4) holds. Using (2.7), that inequality yields

$$\int_0^{t_j} L_s\mathbb{E}[\|X_s - X_{\tau^N(s)}\|]\,\mathrm{d}s \leq \Delta t_N \ell_x \int_0^{t_j} (\mathbb{E}[M_s] + \ell_x\mathbb{E}[\|X_s\|])\,\mathrm{d}s.$$

Thanks to (2.5) and (2.6), inequality (2.8) holds, and we obtain

$$\begin{aligned}
&\int_0^{t_j} L_s \mathbb{E}[\|X_s - X_{\tau^N(s)}\|]\,\mathrm{d}s \\
&\quad\le \Delta t_N \ell_x \int_0^{t_j} \left(\mathbb{E}[M_s] + \ell_x \left(\mathbb{E}[\|X_0\|] + \int_0^s \mathbb{E}[M_\xi]\,\mathrm{d}\xi\right) e^{\ell_x s}\right)\,\mathrm{d}s \\
&\quad\le \Delta t_N \ell_x \left(\int_0^{t_j} \mathbb{E}[M_s]\,\mathrm{d}s + \ell_x \int_0^{t_j} \left(\mathbb{E}[\|X_0\|] + \int_0^{t_j} \mathbb{E}[M_\xi]\,\mathrm{d}\xi\right) e^{\ell_x s}\,\mathrm{d}s\right) \\
&\quad= \Delta t_N \ell_x \left(\int_0^{t_j} \mathbb{E}[M_s]\,\mathrm{d}s + \left(\mathbb{E}[\|X_0\|] + \int_0^{t_j} \mathbb{E}[M_\xi]\,\mathrm{d}\xi\right)\left(e^{\ell_x t_j} - 1\right)\right).
\end{aligned}$$

Thus,

$$\int_0^{t_j} L_s \mathbb{E}[\|X_s - X_{\tau^N(s)}\|]\,\mathrm{d}s \le \Delta t_N \ell_x \left(\mathbb{E}[\|X_0\|] + \int_0^{t_j} \mathbb{E}[M_\xi]\,\mathrm{d}\xi\right) e^{\ell_x t_j}. \tag{6.7}$$

Now we look at Lemma 6.2 and put the pieces together. Taking the expectation of the global error formula (6.3) and using (2.7) to bound the exponential term, we arrive at

$$\begin{aligned}
\mathbb{E}\left[\|X_{t_j} - X^N_{t_j}\|\right] \le \Bigg(&\mathbb{E}\left[\|X_0 - X^N_0\|\right] + \int_0^{t_j} L_s \mathbb{E}\left[\|X_s - X_{\tau^N(s)}\|\right]\,\mathrm{d}s \\
&\mathbb{E}\left[\left\|\int_0^{t_j} \left(f(s, X^N_{\tau^N(s)}, Y_s) - f(\tau^N(s), X^N_{\tau^N(s)}, Y_{\tau^N(s)})\right)\,\mathrm{d}s\right\|\right]\Bigg) e^{\ell_x t_j}.
\end{aligned}$$

Using now estimate (6.7) and condition (6.5), we find (6.6), completing the proof. □

7. Pathwise convergence for FV noises

Here, the noise $\{Y_t\}_{t\in I}$ is assumed to be an FV process, i.e. almost surely with sample paths that are càdlàg and with finite variation.

Theorem 7.1. *Under the Standing Hypothesis 2.1, suppose also that*

$$t \mapsto L_t \text{ is bounded almost surely;} \tag{7.1}$$

that

$$\|X_0 - X^N_0\| \le C_0 \Delta t, \tag{7.2}$$

for a random variable C_0 which is finite almost surely; and that $f = f(t,x,y)$ is uniformly globally Lipschitz continuous in x and is continuously differentiable in (t,y), with differentials $\partial_t f$ and $D_y f$ with at most linear growth in x and y, i.e.

$$\|\partial_t f(t,x,y)\| \le c_1 + c_2\|x\| + c_3\|y\|, \quad \|D_y f(t,x,y)\| \le c_4 + c_5\|x\| + c_6\|y\|, \tag{7.3}$$

in $(t,x,y) \in I \times \mathbb{R}^d \times \mathbb{R}^k$, for suitable constants $c_1, c_2, c_3, c_4, c_5, c_6 \ge 0$. Assume, further, that $\{Y_t\}_{t\in I}$ is an FV process. Then, the Euler scheme is pathwise convergent

of order 1, i.e.

$$\max_{j=0,\dots,N} \left\| X_{t_j}(\omega) - X^N_{t_j}(\omega) \right\| \le C(\omega)\Delta t_N, \qquad \forall N \in \mathbb{N}, \tag{7.4}$$

for a suitable random variable $C = C(\omega)$ which is finite almost surely on Ω.

Proof. In view of Lemma 6.2, we need to estimate the last term of (6.3), namely

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

Thanks to (4.5), we have

$$\begin{aligned} f(s,X^N_{\tau^N(s)}, Y_s) - f(\tau^N(s), X^N_{\tau^N(s)}, Y_{\tau^N(s)}) &= \int_{\tau^N(s)}^s D_\xi f(\xi, X^N_{\tau^N(s)}, Y_{\xi^-}) \, \mathrm{d}\xi \\ &+ \int_{\tau^N(s)^+}^s D_y f(\xi, X^N_{\tau^N(s)}, Y_{\xi^-}) \, \mathrm{d}Y_\xi \\ &+ \sum_{\tau^N(s)<\xi\le s} \left(f(\xi, X^N_{\tau^N(s)}, Y_\xi) - f(\xi, X^N_{\tau^N(s)}, Y_{\xi^-}) - D_y f(\xi, X^N_{\tau^N(s)}, Y_{\xi^-}) \Delta Y_\xi \right). \end{aligned}$$

For the first term, we have, using (7.3),

$$\begin{aligned} &\left\| \int_0^{t_j} \int_{\tau^N(s)}^s D_\xi f(\xi, X^N_{\tau^N(s)}, Y_{\xi^-}) \, \mathrm{d}\xi \, \mathrm{d}s \right\| \\ &\qquad \le \int_0^{t_j} \int_{\tau^N(s)}^s \left(c_1 + c_2 \| X^N_{\tau^N(s)} \| + c_3 \| Y_{\xi^-} \| \right) \, \mathrm{d}\xi \, \mathrm{d}s. \end{aligned}$$

Exchanging the order of integration,

$$\begin{aligned} &\left\| \int_0^{t_j} \int_{\tau^N(s)}^s D_\xi f(\xi, X^N_{\tau^N(s)}, Y_{\xi^-}) \, \mathrm{d}\xi \, \mathrm{d}s \right\| \\ &\qquad \le \int_0^{t_j} \int_{\xi}^{\tau^N(\xi)+\Delta t} \left(c_1 + c_2 \| X^N_{\tau^N(s)} \| + c_3 \| Y_{\xi^-} \| \right) \, \mathrm{d}s \, \mathrm{d}\xi. \end{aligned}$$

For s within $\xi \le s < \tau^N(\xi) + \Delta t$, we have $\tau^N(s) = \tau^N(\xi)$. Moreover, $\tau^N(\xi) + \Delta t - \xi \le \Delta t$. Hence

$$\begin{aligned} &\left\| \int_0^{t_j} \int_{\tau^N(s)}^s D_\xi f(\xi, X^N_{\tau^N(s)}, Y_{\xi^-}) \, \mathrm{d}\xi \, \mathrm{d}s \right\| \\ &\qquad \le \int_0^{t_j} \int_{\xi}^{\tau^N(\xi)+\Delta t} \left(c_1 + c_2 \| X^N_{\tau^N(\xi)} \| + c_3 \| Y_{\xi^-} \| \right) \, \mathrm{d}s \, \mathrm{d}\xi \\ &\qquad \le \Delta t \int_0^{t_j} \left(c_1 + c_2 \| X^N_{\tau^N(\xi)} \| + c_3 \| Y_{\xi^-} \| \right) \, \mathrm{d}\xi. \end{aligned}$$

The second term is similar, except that now we integrate with respect to a finite variation process, so it is a Lebesgue-Stieltjes integral:

$$\begin{aligned}
&\left\| \int_0^{t_j} \int_{\tau^N(s)^+}^{s} D_y f(\xi, X^N_{\tau^N(s)}, Y_{\xi^-}) \, \mathrm{d}Y_\xi \, \mathrm{d}s \right\| \\
&\qquad \le \int_0^{t_j} \int_{\tau^N(s)}^{s} \left(c_4 + c_5 \|X^N_{\tau^N(s)}\| + c_6 \|Y_{\xi^-}\| \right) \, \|\mathrm{d}Y_\xi\| \, \mathrm{d}s \\
&\qquad \le \int_0^{t_j} \int_{\xi}^{\tau^N(\xi)+\Delta t} \left(c_4 + c_5 \|X^N_{\tau^N(s)}\| + c_6 \|Y_{\xi^-}\| \right) \, \mathrm{d}s \, \|\mathrm{d}Y_\xi\| \\
&\qquad \le \Delta t \int_0^{t_j} \left(c_4 + c_5 \|X^N_{\tau^N(\xi)}\| + c_6 \|Y_{\xi^-}\| \right) \, \|\mathrm{d}Y_\xi\|.
\end{aligned}$$

For the third term, we have

$$\begin{aligned}
&\left\| f(\xi, X^N_{\tau^N(s)}, Y_\xi) - f(\xi, X^N_{\tau^N(s)}, Y_{\xi^-}) - D_y f(\xi, X^N_{\tau^N(s)}, Y_{\xi^-}) \Delta Y_\xi \right\| \\
&\qquad\qquad \le 2(c_4 + c_5 \|X^N_{\tau^N(s)}\| + c_6 \max\{\|Y_{\xi^-}\|, \|Y_\xi\|)\}) \|\Delta Y_\xi\|.
\end{aligned}$$

Hence,

$$\begin{aligned}
&\left\| \int_0^{t_j} \sum_{\tau^N(s)<\xi\le s} \left(f(\xi, X^N_{\tau^N(s)}, Y_\xi) - f(\xi, X^N_{\tau^N(s)}, Y_{\xi^-}) - D_y f(\xi, X^N_{\tau^N(s)}, Y_{\xi^-}) \Delta Y_\xi \right) \, \mathrm{d}s \right\| \\
&\qquad\qquad \le \int_0^{t_j} \sum_{\tau^N(s)<\xi\le s} 2(c_4 + c_5 \|X^N_{\tau^N(s)}\| + c_6 \max\{\|Y_{\xi^-}\|, \|Y_\xi\|)\} \|\Delta Y_\xi\| \, \mathrm{d}s.
\end{aligned}$$

Switching the order of the integral and the summation and using (4.3), we find

$$\begin{aligned}
&\left\| \int_0^{t_j} \sum_{\tau^N(s)<\xi\le s} \left(f(\xi, X^N_{\tau^N(s)}, Y_\xi) - f(\xi, X^N_{\tau^N(s)}, Y_{\xi^-}) - D_y f(\xi, X^N_{\tau^N(s)}, Y_{\xi^-}) \Delta Y_\xi \right) \, \mathrm{d}s \right\| \\
&\quad \le \sum_{0<\xi\le t_j} \int_{\xi}^{\tau(\xi)+\Delta t} 2 \left(c_4 + c_5 \|X^N_{\tau^N(s)}\| + c_6 \max\{\|Y_{\xi^-}\|, \|Y_\xi\|\} \right) \|\Delta Y_\xi\| \, \mathrm{d}s \\
&\quad \le 2 \sum_{0<\xi\le t_j} \left(\int_{\xi}^{\tau(\xi)+\Delta t} \left(c_4 + c_5 \|X^N_{\tau^N(s)}\| + c_6 \max\{\|Y_{\xi^-}\|, \|Y_\xi\|\} \right) \, \mathrm{d}s \right) \|\Delta Y_\xi\| \\
&\quad \le 2 \left(c_4 + c_5 \max_{i=0,\dots,j} \|X^N_i\| + c_6 \sup_{0\le s\le t_j} \|Y_s\| \right) \Delta t \sum_{0<\xi\le t_j} \|\Delta Y_\xi\| \\
&\quad \le 2 \left(c_4 + c_5 \max_{i=0,\dots,j} \|X^N_i\| + c_6 \sup_{0\le s\le t_j} \|Y_s\| \right) V(\{Y_t\}_{t\in I}; [0, t_j]) \Delta t.
\end{aligned}$$

Putting the estimates together, we find that

$$\left\| \int_0^{t_j} \left(f(s, X^N_{\tau^N(s)}, Y_s) - f(\tau^N(s), X^N_{\tau^N(s)}, Y_{\tau^N(s)}) \right) \, \mathrm{d}s \right\| \leq \tilde{C} \Delta t,$$

where $\tilde{C}$ is the random variable

$$\begin{aligned} \tilde{C} = \int_0^T \left(c_1 + c_2 \|X^N_{\tau^N(s)}\| + c_3 \|Y_{s^-}\| \right) \, \mathrm{d}s \\ + \int_0^T \left(c_4 + c_5 \|X^N_{\tau^N(s)}\| + c_6 \|Y_{s^-}\| \right) \, \|\mathrm{d}Y_s\| \\ + 2 \left(c_4 + c_5 \max_{i=0,\ldots,N} \|X^N_i\| + c_6 \sup_{0 \leq s \leq T} \|Y_s\| \right) V(\{Y_t\}_{t \in I}; [0, t_j]). \end{aligned} \tag{7.5}$$

Using this estimate in (6.3), we obtain

$$\|X_{t_j} - X^N_{t_j}\| \leq \left(\|X_0 - X^N_0\| + \int_0^{t_j} L_s \|X_s - X_{\tau^N(s)}\| \, \mathrm{d}s \; + \tilde{C} \Delta t \right) e^{\int_0^{t_j} L_s \, \mathrm{d}s}$$

From (6.4) and the hypothesis (7.2) on the initial condition we prove (7.4), with a suitable $C = C(\omega)$. **Wait, the constant $C = C(\omega)$ cannot depend on the solution. Need to use (3.1) on the estimates above.** □

8. Strong converge for semi-martingale noises

We now consider the more general case of a semi-martingale noise. In this case, we estimate the strong convergence.

Theorem 8.1. *Under the Standing Hypothesis 2.1, suppose also that*

$$t \mapsto L_t \text{ is bounded almost surely;} \tag{8.1}$$

that

$$\mathbb{E}\left[\|X_0 - X^N_0\| \right] \leq c_0 \Delta t, \tag{8.2}$$

for a constant $c_0 \geq 0$; and that $f = f(t, x, y)$ is uniformly globally Lipschitz continuous in x and is twice continuously differentiable in (t, y), with partial differentials $\partial_t f$, $D_y f$ and $D_{yy} f$ satisfying (7.3) and

$$\|D_{yy} f(t, x, y)\| \leq c_7 + c_8 \|x\| + c_9 \|y\|, \tag{8.3}$$

in $(t, x, y) \in I \times \mathbb{R}^d \times \mathbb{R}^k$, for suitable constants $c_1, c_2, c_4, c_5, c_7, c_8 \geq 0$. Assume, further, that $\{Y_t\}_{t \in I}$ is a semi-martingale. Then, the Euler scheme is strongly convergent of order 1, i.e.

$$\max_{j=0,\ldots,N} \mathbb{E}\left[\left\| X_{t_j} - X^N_{t_j} \right\| \right] \leq c \Delta t_N, \qquad \forall N \in \mathbb{N}, \tag{8.4}$$

for a suitable constant $c \geq 0$.

Proof. As in the proof of (7.1), we need to estimate the last term of (6.3). Using (4.6), the last term becomes

$$\begin{aligned}
&\int_0^{t_j} \Big(f(s, X^N_{\tau^N(s)}, Y_s) - f(\tau^N(s), X^N_{\tau^N(s)}, Y_{\tau^N(s)}) \Big)\, \mathrm{d}s \\
&\quad = \int_0^{t_j} \int_{\tau^N(s)}^{s} D_\xi f(\xi, X^N_{\tau^N(s)}, Y_{\xi^-})\, \mathrm{d}\xi\, \mathrm{d}s \\
&\qquad + \int_0^{t_j} \int_{\tau^N(s)^+}^{s} D_y f(\xi, X^N_{\tau^N(s)}, Y_{\xi^-})\, \mathrm{d}Y_\xi\, \mathrm{d}s \\
&\qquad + \int_0^{t_j} \sum_{\tau^N(s) < \xi \le s} \Big(f(\xi, X^N_{\tau^N(s)}, Y_\xi) - f(\xi, X^N_{\tau^N(s)}, Y_{\xi^-}) \\
&\qquad\qquad - D_y f(\xi, X^N_{\tau^N(s)}, Y_{\xi^-}) \Delta Y_\xi \Big)\, \mathrm{d}s \\
&\qquad + \frac{1}{2} \int_0^{t_j} \int_{\tau^N(s)}^{s} D_{yy} f(\xi, X^N_{\tau^N(s)}, Y_{\xi^-})\, \mathrm{d}[Y,Y]^c_\xi\, \mathrm{d}s.
\end{aligned} \tag{8.5}$$

The first term in (8.5) is handled just like in the proof of Theorem 7.1 and we obtain

$$\left\| \int_0^{t_j} \int_{\tau^N(s)}^{s} D_\xi f(\xi, X^N_{\tau^N(s)}, Y_{\xi^-})\, \mathrm{d}\xi\, \mathrm{d}s \right\| \le \Delta t \int_0^{t_j} \Big(c_1 + c_2 \|X^N_{\tau^N(\xi)}\| + c_3 \|Y_{\xi^-}\| \Big)\, \mathrm{d}\xi.$$

For the second term in (8.5), we use the Bichteler-Dellacherie Theorem [32, Theorem III.47], which says that $\{Y_t\}_{t\in I}$ can be decomposed as $Y_t = F_t + Z_t$, where $\{F_t\}_{t\in I}$ is an FV process and $\{Z_t\}_{t\in I}$ is a local martingale. Hence, the second term can be written as

$$\begin{aligned}
\int_0^{t_j} \int_{\tau^N(s)^+}^{s} D_y f(\xi, X^N_{\tau^N(s)}, Y_{\xi^-})\, \mathrm{d}Y_\xi\, \mathrm{d}s &= \int_0^{t_j} \int_{\tau^N(s)^+}^{s} D_y f(\xi, X^N_{\tau^N(s)}, Y_{\xi^-})\, \mathrm{d}F_\xi\, \mathrm{d}s \\
&\quad + \int_0^{t_j} \int_{\tau^N(s)^+}^{s} D_y f(\xi, X^N_{\tau^N(s)}, Y_{\xi^-})\, \mathrm{d}Z_\xi\, \mathrm{d}s,
\end{aligned} \tag{8.6}$$

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 part of (8.6) is estimated as in the proof of Theorem 7.1, so

that

$$\begin{aligned}
&\left\| \int_0^{t_j} \int_{\tau^N(s)^+}^{s} D_y f(\xi, X^N_{\tau^N(s)}, Y_{\xi^-}) \, \mathrm{d}F_\xi \, \mathrm{d}s \right\| \\
&\quad \le \int_0^{t_j} \int_{\tau^N(s)}^{s} \left(c_4 + c_5 \|X^N_{\tau^N(s)}\| + c_6 \|Y_{\xi^-}\| \right) \, \|\mathrm{d}F_\xi\| \, \mathrm{d}s \\
&\quad \le \int_0^{t_j} \int_{\xi}^{\tau^N(\xi)+\Delta t} \left(c_4 + c_5 \|X^N_{\tau^N(s)}\| + c_6 \|Y_{\xi^-}\| \right) \, \mathrm{d}s \, \|\mathrm{d}F_\xi\| \\
&\quad \le \Delta t \int_0^{t_j} \left(c_4 + c_5 \|X^N_{\tau^N(\xi)}\| + c_6 \|Y_{\xi^-}\| \right) \, \|\mathrm{d}F_\xi\| \\
&\quad \le \Delta t \left(c_4 + c_5 \max_{i=0,\dots,j} \|X^N_i\| + c_6 \sup_{0\le\xi\le t_j} \|Y_\xi\| \right) V(\{F_t\}_{t\in I}; [0, t_j]).
\end{aligned}$$

For the second term in (8.6), we use the Itô isometry for local martingales, so that

$$\begin{aligned}
&\mathbb{E}\left[\left\| \int_0^{t_j} \int_{\tau^N(s)^+}^{s} D_y f(\xi, X^N_{\tau^N(s)}, Y_{\xi^-}) \, \mathrm{d}Z_\xi \, \mathrm{d}s \right\| \right] \\
&\quad = \mathbb{E}\left[\left\| \int_0^{t_j} \int_{\xi}^{\tau^N(\xi)+\Delta t} D_y f(\xi, X^N_{\tau^N(s)}, Y_{\xi^-}) \, \mathrm{d}s \, \mathrm{d}Z_\xi \right\| \right] \\
&\quad \le \mathbb{E}\left[\left(\int_0^{t_j} \int_{\xi}^{\tau^N(\xi)+\Delta t} D_y f(\xi, X^N_{\tau^N(s)}, Y_{\xi^-}) \, \mathrm{d}s \, \mathrm{d}Z_\xi \right)^2 \right]^{1/2} \\
&\quad \le \mathbb{E}\left[\int_0^{t_j} \left(\int_{\xi}^{\tau^N(\xi)+\Delta t} \|D_y f(\xi, X^N_{\tau^N(s)}, Y_{\xi^-})\| \, \mathrm{d}s \right)^2 \mathrm{d}[Z, Z]_\xi \right]^{1/2} \\
&\quad \le \Delta t \mathbb{E}\left[\int_0^{t_j} \|D_y f(\xi, X^N_{\tau^N(\xi)}, Y_{\xi^-})\|^2 \, \mathrm{d}[Z, Z]_\xi \right]^{1/2} \\
&\quad \le \Delta t \mathbb{E}\left[\int_0^{t_j} \left(c_4 + c_5 \|X^N_{\tau^N(\xi)}\| + c_6 \|Y_{\xi^-}\| \right)^2 \mathrm{d}[Z, Z]_\xi \right].
\end{aligned}$$

The third term can be treated as in Theorem 7.1, but instead we use the bound (8.3) on the second derivative of $D_{yy} f$ to write

$$\begin{aligned}
&\left\| f(\xi, X^N_{\tau^N(s)}, Y_\xi) - f(\xi, X^N_{\tau^N(s)}, Y_{\xi^-}) - D_y f(\xi, X^N_{\tau^N(s)}, Y_{\xi^-}) \Delta Y_\xi \right\| \\
&\qquad\qquad \le \frac{1}{2} \left(c_7 + c_8 \|X^N_{\tau^N(s)}\| + c_9 \max\{\|Y_\xi\|, \|Y_{\xi^-}\|\} \right) (\Delta Y_\xi)^2 .
\end{aligned}$$

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

$$\begin{aligned}
&\left\| \int_0^{t_j} \sum_{\tau^N(s)<\xi\le s} \left(f(\xi, X^N_{\tau^N(s)}, Y_\xi) - f(\xi, X^N_{\tau^N(s)}, Y_{\xi^-}) - D_y f(\xi, X^N_{\tau^N(s)}, Y_{\xi^-}) \Delta Y_\xi \right) \, \mathrm{d}s \right\| \\
&\quad \le \frac{1}{2} \int_0^{t_j} \sum_{\tau^N(s)<\xi\le s} \left(c_7 + c_8 \|X^N_{\tau^N(\xi)}\| + c_9 \max\{\|Y_\xi\|, \|Y_{\xi^-}\|\} \right) (\Delta Y_\xi)^2 \, \mathrm{d}s \\
&\quad \le \frac{1}{2} \sum_{0<\xi\le t_j} \int_\xi^{\tau^N(\xi)+\Delta t} \left(c_7 + c_8 \|X^N_{\tau^N(\xi)}\| + c_9 \max\{\|Y_\xi\|, \|Y_{\xi^-}\|\} \right) (\Delta Y_\xi)^2 \, \mathrm{d}s \\
&\quad \le \frac{\Delta t}{2} \left(c_7 + c_8 \max_{i=0,\dots,j} \|X^N_{t_i}\| + c_9 \sup_{0\le\xi\le t_j} \|Y_\xi\| \right) \sum_{0<\xi\le t_j} (\Delta Y_\xi)^2 .
\end{aligned}$$

The final term in (8.5) involves a Lebesgue-Stieltjes integral and is handled in a similar way, so that

$$\begin{aligned}
&\left\| \frac{1}{2} \int_0^{t_j} \int_{\tau^N(s)}^s D_{yy} f(\xi, X^N_{\tau^N(s)}, Y_{\xi^-}) \, \mathrm{d}[Y,Y]^c_\xi \, \mathrm{d}s \right\| \\
&\qquad \le \frac{1}{2} \int_0^{t_j} \int_\xi^{\tau^N(\xi)+\Delta t} \left(c_7 + c_8 \|X^N_{\tau^N(\xi)}\| + c_0 \|Y_{\xi^-}\| \right) \, \mathrm{d}s \, \mathrm{d}[Y,Y]^c_\xi \\
&\qquad \le \frac{\Delta t}{2} \int_0^{t_j} \left(c_7 + c_8 \|X^N_{\tau^N(\xi)}\| + c_9 \|Y_{\xi^-}\| \right) \, \mathrm{d}[Y,Y]^c_\xi \\
&\qquad \le \frac{\Delta t}{2} \left(c_7 + c_8 \max_{i=0,\dots,j} \|X^N_{t_i}\| + c_9 \sup_{0\le\xi\le t_j} \|Y_\xi\| \right) [Y,Y]^c_{t_j}.
\end{aligned}$$

Putting these estimates together we obtain

$$\begin{aligned}
&\left\| \int_0^{t_j} \left(f(s, X^N_{\tau^N(s)}, Y_s) - f(\tau^N(s), X^N_{\tau^N(s)}, Y_{\tau^N(s)}) \right) \, \mathrm{d}s \right\| \\
&\qquad \le \Delta t \int_0^{t_j} \left(c_1 + c_2 \|X^N_{\tau^N(\xi)}\| + c_3 \|Y_{\xi^-}\| \right) \, \mathrm{d}\xi \\
&\qquad\quad + \Delta t \mathbb{E} \left[\int_0^{t_j} \left(c_4 + c_5 \|X^N_{\tau^N(\xi)}\| + c_6 \|Y_{\xi^-}\| \right)^2 \, \mathrm{d}[Z,Z]_\xi \right] \\
&\qquad\quad + \frac{\Delta t}{2} \left(c_7 + c_8 \max_{i=0,\dots,j} \|X^N_{t_i}\| + c_9 \sup_{0\le\xi\le t_j} \|Y_\xi\| \right) \sum_{0<\xi\le t_j} (\Delta Y_\xi)^2 \\
&\qquad\quad + \frac{\Delta t}{2} \left(c_7 + c_8 \max_{i=0,\dots,j} \|X^N_{t_i}\| + c_9 \sup_{0\le\xi\le t_j} \|Y_\xi\| \right) [Y,Y]^c_{t_j}.
\end{aligned} \tag{8.7}$$

□

9. Numerical examples

In this section, we illustrate the improved order of convergence with several examples that fall into one of the cases considered above. We start with some linear equations, including all sorts of noises. Then we illustrate the $H+1/2$ order of convergence in the case of a fractional Brownian motion (fBm) noise with Hurst parameter $0 < H < 1/2$. Next we move to a number of linear and nonlinear models, as mentioned in the Introduction. Details of the numerical implementations, with the full working code for reproducibility puposes is available in the github repository [34].

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

$$\epsilon^N = \max_{j=0,\dots,N} \frac{1}{M} \sum_{m=1}^{M} \left| X_{t_j}(\omega_m) - X_{t_j}^N(\omega_m) \right|. \tag{9.1}$$

Then we fit the errors ϵ^N to the power law $C\Delta t_N^p$, in order to find p, along with the 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 $n_i \in \mathbb{N}$, so that finer meshes are refinements of coarser meshes.
(v) A number N_{tgt} for a finer resolution to compute a target solution path, typically $N_{\text{tgt}} = \max_i\{N_i^2\}$, unless an exact pathwise solution is available, in which case a coarser mesh of the order of $\max_i\{N_i\}$ can be used.

And here is the method:

(i) For each sample ω_m, $m = 1,\dots,M$, we first generate a discretization of a sample path of the noise, $\{Y_{t_j}\}_{j=0,\dots,N_{\text{tgt}}}$, on the finest grid $t_j = j\Delta t_{N_{\text{tgt}}}$, $j = 0,\dots,N_{\text{tgt}}$, using an exact distribution for the noise.
(ii) Next, we use the values of the noise at the target time mesh to generate the target solution $\{X_{t_j}\}_{j=0,\dots,N_{\text{tgt}}}$, still on the fine mesh. This is constructed either using the Euler approximation itself, keeping in mind that the mesh is sufficiently fine, or an exact distributions 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 $t_j = j\Delta t_{N_i}$, $j = 0,\dots,N_i$.

(iv) We then compare each approximation $\{X_{t_j}^{N_i}\}_{j=0,\dots,N_i}$ 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,

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

(vi) We fit $(\Delta t_{N_i}, \epsilon^{N_i})$ to the power law $C\Delta t_{N_i}^p$, via linear least-square regression in log scale, so that $\ln \epsilon^{N_i} \sim \ln C + p \ln \Delta t_{N_i}$, for suitable C and p, with p giving the order of convergence. This amounts to solving the normal equation $(A^tA)\mathbf{v} = A^t \ln(\boldsymbol{\epsilon})$, 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,

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

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

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

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

$$\epsilon_{\min}^{N_i} = \max_{j=0,\dots,N_i}(\epsilon_{t_j}^{N_i} - 2\sigma_{t_j}^{N_i}), \quad \epsilon_{\max}^{N_i} = \max_{j=0,\dots,N_i}(\epsilon_{t_j}^{N_i} + 2\sigma_{t_j}^{N_i}).$$

(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 $\mathbf{e} \mapsto (\ln(C), p) = (A^tA)^{-1}A^t \ln(\mathbf{e})$, of the polytope formed by all combinations of the indices in $(\epsilon_{\min}^{N_i})_i$ and $(\epsilon_{\max}^{N_i})_i$, which is the image of the set of 95% confidence intervals for the errors obtained with the Monte-Carlo approximation of the strong error.

As for the implementation itself, all code is written in the Julia language [5]. Julia is a high-performance programming language, suitable for scientific computing and computationally-demanding applications.

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 a github repository [34].

9.1. **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:

$$\begin{cases} \dfrac{\mathrm{d}X_t}{\mathrm{d}t} = W_t X_t, & 0 \le t \le T, \\ X_t|_{t=0} = X_0. \end{cases} \tag{9.2}$$

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 [20]. Nevertheless, we illustrate the strong convergence directly for the Euler method for this RODE equation, for the sake of completeness.

Equation (9.2) has the explicit solution

$$X_t = e^{\int_0^t W_s \,\mathrm{d}s} X_0. \tag{9.3}$$

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 [19, Section 14.2] as

$$\int_0^\tau W_s \,\mathrm{d}s \sim \frac{\tau}{2}\Delta W + \sqrt{\frac{\tau^3}{12}}\mathcal{N}(0,1) = \frac{\tau}{2}\Delta W + \mathcal{N}\left(0, \frac{\tau^3}{12}\right). \tag{9.4}$$

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 (9.4), 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

$$Z_i \sim \mathcal{N}\left(0, \frac{(t_{i+1}-t_i)^3}{12}\right) = \frac{\sqrt{(t_{i+1}-t_i)^3}}{\sqrt{12}}\mathcal{N}(0,1). \tag{9.5}$$

Thus, breaking down the sum over the mesh intervals,

$$\int_0^{t_j} W_s \, \mathrm{d}s = \sum_{i=0}^{j-1} \int_{t_i}^{t_{i+1}} W_s \, \mathrm{d}s = \sum_{i=0}^{j-1} \left(\frac{(W_{t_{i+1}} + W_{t_i})}{2}(t_{i+1}-t_i) + Z_i \right), \tag{9.6}$$

where $Z_0, \ldots, Z_{N-1}$ are independent random variables with distribution (9.5).

Notice the first term in (9.6) 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,

$$\mathbb{E}[e^{Z_i}] = e^{((t_{i+1}-t_i)^3)/24}. \tag{9.7}$$

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

Hence, once an Euler approximation of (9.2) is computed, along with realizations $\{W_{t_i}\}_{i=0}^N$ of a sample path of the noise, we consider an exact solution given by

$$X_{t_j} = X_0 e^{\sum_{i=0}^{j-1}\left(\frac{1}{2}(W_{t_i}+W_{t_{i+1}})(t_{i+1}-t_i)+Z_i\right)}, \tag{9.8}$$

for realizations Z_i drawn from a normal distributions given by (9.5). Figure 1 shows an approximate solution and a few sample paths of exact solutions associated with the given realizations of the noise on the mesh points.

Table 1 shows the estimated strong error obtained from the $M = 200$ Monte Carlo simulations for each chosen time step $N_i = 2^4, \ldots, 2^{14}$, with initial condition $X_0 \sim \mathcal{N}(0,1)$, on the time interval $[0,T] = [0.0, 1.0]$. The target resolution is set to $N_{\text{tgt}} = 2^{16}$, since we have the exact distribution for it. Figure 2 illustrates the order of convergence.

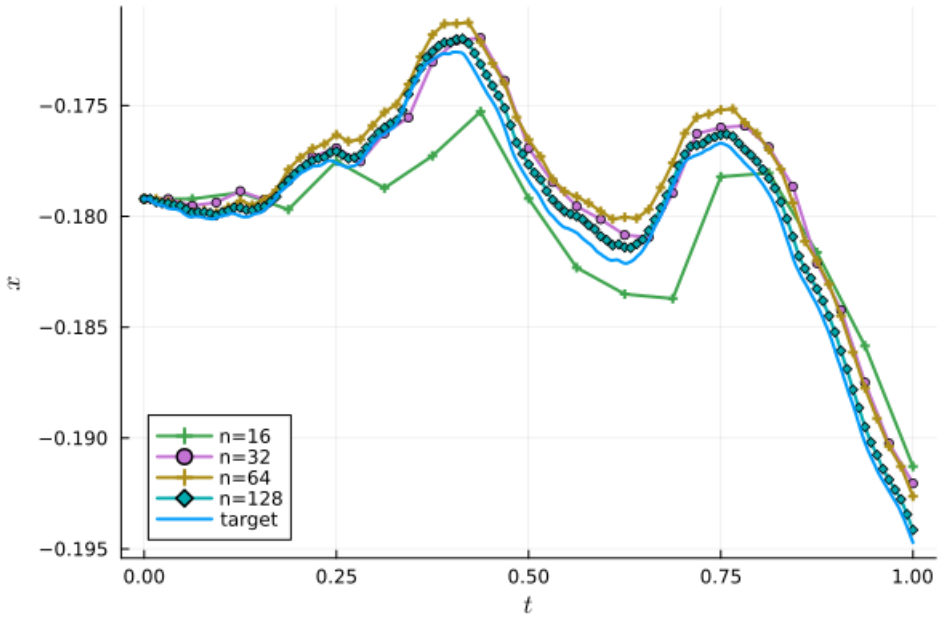

FIGURE 1. Euler approximation of a sample solution of $\mathrm{d}X_t/\mathrm{d}t = W_t X_t$ with $X_0 \sim \mathcal{N}(0,1)$, with a few mesh resolutions and with the target resolution.

9.2. **Non-homogeneous linear system of RODEs with different types of noises.** Next we consider a system of linear equations with a number of different types of noise. For most of these 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 at once, for all such noises.

The system of equations takes the form

$$\begin{cases} \dfrac{\mathrm{d}\mathbf{X}_t}{\mathrm{d}t} = -\left\|\mathbf{Y}_t\right\|^2 \mathbf{X}_t + \mathbf{Y}_t, \qquad 0 \le t \le T, \\ \mathbf{X}_t|_{t=0} = \mathbf{X}_0, \end{cases} \tag{9.9}$$

where $\{\mathbf{X}_t\}_t$ is vector valued, and $\{\mathbf{Y}_t\}_t$ is a given vector-valued noise process with the same dimension as $\mathbf{X}_t$. Each coordinate of $\{\mathbf{Y}_t\}_t$ 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 $\mathbf{Y}_t$:

(i) A standard Wiener process;
(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 $y_0 = 0.2$;

N	dt	error	std err
16	0.0625	0.0379	0.0042
32	0.0312	0.0192	0.00208
64	0.0156	0.0095	0.00103
128	0.00781	0.00478	0.000519
256	0.00391	0.0024	0.000269
512	0.00195	0.00123	0.000143
1024	0.000977	0.000633	7.43e-5
2048	0.000488	0.000324	3.74e-5
4096	0.000244	0.000154	1.65e-5
8192	0.000122	7.71e-5	8.35e-6
16384	6.1e-5	3.76e-5	4.06e-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 $M = 500$ sample paths for each fixed time step, with the target solution calculated with 65536 points. The order of strong convergence is estimated to be $p = 0.993$, with the 95% confidence interval $[0.9497, 1.0362]$.

(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 $H_0 = 0.2$;
(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 $M = 80$ Monte Carlo simulations for each chosen time step $N_i = 2^6, \ldots, 2^9$, on the time interval $[0, T] = [0.0, 1.0]$. All coordinates of the initial condition are normally distributed with mean zero and variance 1. The target resolution is set to $N_{\mathrm{tgt}} = 2^{18}$.

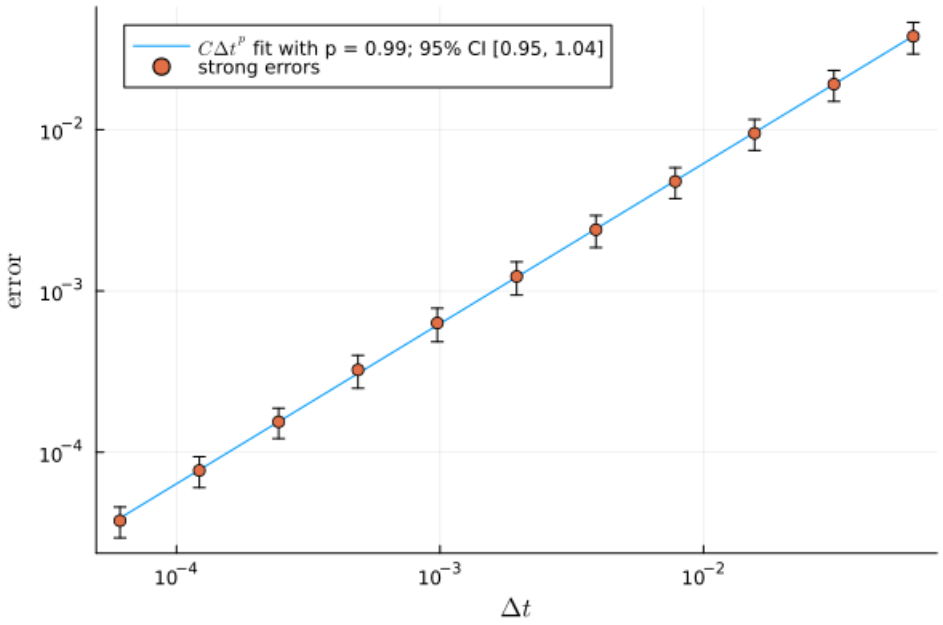

FIGURE 2. Order of convergence $p = 0.974$ of the strong error of the Euler method for $\mathrm{d}X_t/\mathrm{d}t = W_t X_t$, based on Table 1.

N	dt	error	std err
64	0.0156	0.196	0.0201
128	0.00781	0.0951	0.00972
256	0.00391	0.0473	0.00483
512	0.00195	0.0237	0.00242

TABLE 2. Mesh points (N), time steps (dt), strong error (error), and standard error (std err) of the Euler method for $\mathrm{d}\mathbf{X}_t/\mathrm{d}t = -\|\mathbf{Y}_t\|^2\mathbf{X}_t + \mathbf{Y}_t$ for each mesh resolution N, with initial condition $\mathbf{X}_0 \sim \mathcal{N}(\mathbf{0}, \mathrm{I})$ and vector-valued noise $\{\mathbf{Y}_t\}_t$ with all the implemented noises, on the time interval $I = [0.0, 1.0]$, based on $M = 80$ sample paths for each fixed time step, with the target solution calculated with 262144 points. The order of strong convergence is estimated to be $p = 1.017$, with the 95% confidence interval $[0.8982, 1.1349]$.

Figure 3 illustrates the obtained order of convergence, while Figure 4 illustrates some sample paths of all the noises used in this system.

The strong order 1 convergence is not a suprise 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

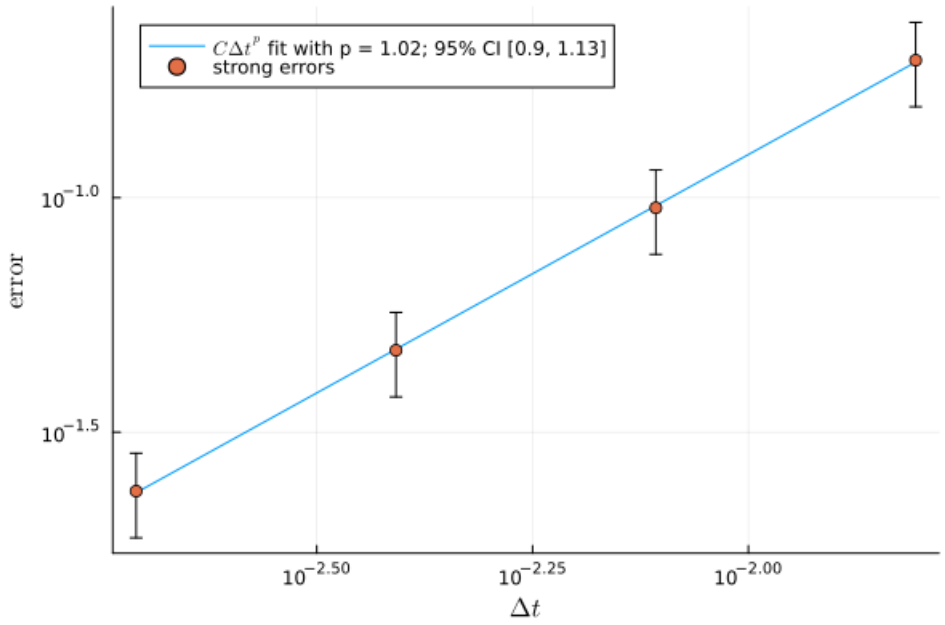

FIGURE 3. Order of convergence $p = 1.012$ of the strong error of the Euler method for $\mathrm{d}\mathbf{X}_t/\mathrm{d}t = -\|\mathbf{Y}_t\|^2 \mathbf{X}_t + \mathbf{Y}_t$, based on Table 2.

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 [20]. The geometric Brownian motion is also expected to yield strong order 1 convergence thanks to the results in [39]. For the remaining noises, however, previous works would estimate the order of convergence to be below the order 1 attained here.

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

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

$$\begin{cases} \dfrac{\mathrm{d}X_t}{\mathrm{d}t} = -X_t + B_t^H, & 0 \le t \le T, \\ X_t|_{t=0} = X_0, \end{cases} \tag{9.10}$$

except this time 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 $H + 1/2$. 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 [6, Theorem 4.2.6], [4, Theorem 4.1]and [29, Theorem 2.7.4]), based on the Wick

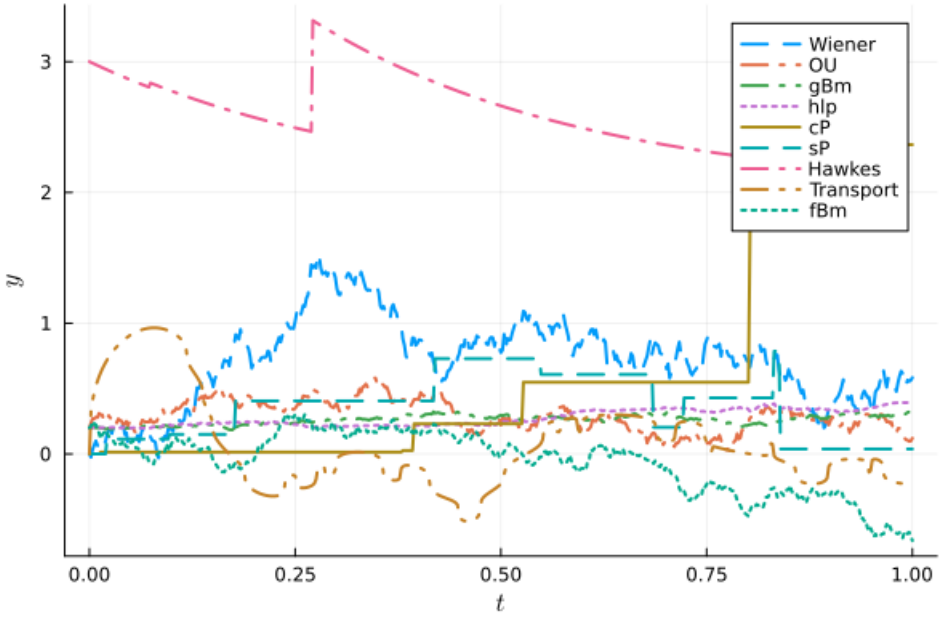

FIGURE 4. Sample paths of all the noises used in the linear system (9.9), mixing all different types of implemented noises.

Itô Skorohod (WIS) integral (see [6, Chapter 4]). A corresponding WIS isometry is also needed (see e.g. [6, 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 first sections.

We need to estimate the last term in (6.6) of Proposition 6.1, involving the steps of the term $f(t,x,y) = -x + y$, which in this case reduce to

$$f(s, X^N_{\tau^N(s)}, Y_s) - f(\tau^N(s), X^N_{\tau^N(s)}, Y_{\tau^N(s)}) = B^H_s - B^H_{\tau^N(s)}, \tag{9.11}$$

for $0 \le s \le T$. There are several ways to represent an fBm (see e.g. [6, 29]). We use the formula (see [28, eq. (2.1)] or [6, eq. (1.1)])

$$B^H_t = \frac{1}{\Gamma(H+1/2)} \left(\int_{-\infty}^0 \left((t-s)^{H-1/2} - (-s)^{H-1/2} \right) \, \mathrm{d}W_s \right. \\ \left. + \int_0^t (t-s)^{H-1/2} \, \mathrm{d}W_s \right), \tag{9.12}$$

where $\Gamma(\cdot)$ is the well-known Gamma function. For the step, (9.12) means that

$$B_s^H - B_{\tau^N(s)}^H = \frac{1}{\Gamma(H+1/2)} \left(\int_{-\infty}^{\tau^N(s)} \left((s-\xi)^{H-1/2} - (\tau^N(s)-\xi)^{H-1/2} \right) \, \mathrm{d}W_\xi + \int_{\tau^N(s)}^{s} (s-\xi)^{H-1/2} \, \mathrm{d}W_\xi \right). \quad (9.13)$$

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

$$\begin{aligned}
&\int_0^{t_j} \left(f(s, X^N_{\tau^N(s)}, Y_s) - f(\tau^N(s), X^N_{\tau^N(s)}, Y_{\tau^N(s)}) \right) \, \mathrm{d}s \\
&\quad = \frac{1}{\Gamma(H+1/2)} \int_0^{t_j} \int_{-\infty}^{\tau^N(s)} \left((s-\xi)^{H-1/2} - (\tau^N(s)-\xi)^{H-1/2} \right) \, \mathrm{d}W_\xi \, \mathrm{d}s \\
&\qquad + \frac{1}{\Gamma(H+1/2)} \int_0^{t_j} \int_{\tau^N(s)}^{s} (s-\xi)^{H-1/2} \, \mathrm{d}W_\xi \, \mathrm{d}s \\
&\quad = \frac{1}{\Gamma(H+1/2)} \int_{-\infty}^{0} \int_0^{t_j} \left((s-\xi)^{H-1/2} - (\tau^N(s)-\xi)^{H-1/2} \right) \, \mathrm{d}s \, \mathrm{d}W_\xi \\
&\qquad + \frac{1}{\Gamma(H+1/2)} \int_0^{t_j} \int_{\tau^N(\xi)+\Delta t_N}^{t_j} \left((s-\xi)^{H-1/2} - (\tau^N(s)-\xi)^{H-1/2} \right) \, \mathrm{d}s \, \mathrm{d}W_\xi \\
&\qquad + \frac{1}{\Gamma(H+1/2)} \int_0^{t_j} \int_{\xi}^{\tau^N(\xi)+\Delta t_N} (s-\xi)^{H-1/2} \, \mathrm{d}s \, \mathrm{d}W_\xi.
\end{aligned} \quad (9.14)$$

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

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

Thus,

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

Exchanging the order of integration yields

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

Hence,

$$\begin{aligned}\left|\int_0^{t_j} \left((s-\xi)^{H-1/2} - (\tau^N(s)-\xi)^{H-1/2}\right)\,\mathrm{d}s\right| & \\ &\le (1/2-H)\int_0^{t_j} \Delta t_N(\sigma-\xi)^{H-3/2}\,\mathrm{d}\sigma.\end{aligned}$$

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

$$\begin{aligned}&\mathbb{E}\left[\left|\int_{-\infty}^0\int_0^{t_j} \left((s-\xi)^{H-1/2} - (\tau^N(s)-\xi)^{H-1/2}\right)\,\mathrm{d}s\,\mathrm{d}W_\xi\right|\right] \\ &\quad\le \left(\int_{-\infty}^0 \left(\int_0^{t_j} \left((s-\xi)^{H-1/2} - (\tau^N(s)-\xi)^{H-1/2}\right)\,\mathrm{d}s\right)^2\,\mathrm{d}\xi\right)^{1/2} \\ &\quad\le \Delta t_N \left(\int_{-\infty}^0 \left((1/2-H)\int_0^{t_j}(\sigma-\xi)^{H-3/2}\,\mathrm{d}\sigma\right)^2\,\mathrm{d}\xi\right)^{1/2} \\ &\quad\le (1/2-H)\Delta t_N \left(\int_{-\infty}^0 \left(\int_0^{T}(\sigma-\xi)^{H-3/2}\,\mathrm{d}\sigma\right)^2\,\mathrm{d}\xi\right)^{1/2}.\end{aligned}$$

Therefore,

$$\begin{aligned}\frac{1}{\Gamma(H+1/2)}\Delta t_N\mathbb{E}\left[\left|\int_{-\infty}^0\int_0^{t_j} \left((s-\xi)^{H-1/2} - (\tau^N(s)-\xi)^{H-1/2}\right)\,\mathrm{d}s\,\mathrm{d}W_\xi\right|\right] & \\ &\le c_H^{(1)}\Delta t_N, \quad (9.15)\end{aligned}$$

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

The second term is similar,

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

Thus,

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

Hence,

$$\begin{aligned}&\mathbb{E}\left[\left|\int_0^{t_j}\int_{\tau^N(\xi)+\Delta t_N}^{t_j}\left((s-\xi)^{H-1/2} - (\tau^N(s)-\xi)^{H-1/2}\right)\,\mathrm{d}s\,\mathrm{d}W_\xi\right|\right] \\ &\quad\leq \left(\int_0^{t_j}\left(\int_{\tau^N(\xi)+\Delta t_N}^{t_j}\left((s-\xi)^{H-1/2} - (\tau^N(s)-\xi)^{H-1/2}\right)\,\mathrm{d}s\right)^2\,\mathrm{d}\xi\right)^{1/2} \\ &\quad\leq \Delta t_N(1/2-H)\left(\int_0^{t_j}\left(\int_{\tau^N(\xi)+\Delta t_N}^{T}(\sigma-\xi)^{H-3/2}\,\mathrm{d}\sigma\right)^2\,\mathrm{d}\xi\right)^{1/2}.\end{aligned}$$

Therefore,

$$\begin{aligned}\frac{1}{\Gamma(H+1/2)}\mathbb{E}\left[\left|\int_0^{t_j}\int_{\tau^N(\xi)+\Delta t_N}^{t_j}\left((s-\xi)^{H-1/2} - (\tau^N(s)-\xi)^{H-1/2}\right)\,\mathrm{d}s\,\mathrm{d}W_\xi\right|\right] \\ \leq c_H^{(2)}\Delta t_N, \quad (9.16)\end{aligned}$$

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

For the last term, we have

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

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,

$$\frac{1}{\Gamma(H+1/2)}\mathbb{E}\left[\left|\int_0^{t_j}\int_\xi^{\tau^N(\xi)+\Delta t_N}(s-\xi)^{H-1/2}\,\mathrm{d}s\,\mathrm{d}W_\xi\right|\right]\leq c_H^{(3)}\Delta t_N^{H+1/2},\tag{9.17}$$

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

Putting the three estimates (9.15), (9.16), (9.17) in (9.14) we find that

$$\mathbb{E}\left[\left|\int_0^{t_j}\left(f(s,X^N_{\tau^N(s)},Y_s)-f(\tau^N(s),X^N_{\tau^N(s)},Y_{\tau^N(s)})\right)\,\mathrm{d}s\right|\right]$$
$$\leq c_H^{(4)}\Delta t_N+c_H^{(3)}\Delta t_N^{H+1/2},\tag{9.18}$$

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 9.1. *Consider the equation (9.10) where $\{B_t^H\}_t$ is a fractional Brownian motion (fBm) with Hurst parameter $0<H<1$. Suppose the initial condition X_0 satisfies*

$$\mathbb{E}[|X_0|^2]<\infty.\tag{9.19}$$

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,

$$\max_{j=0,\ldots,N}\mathbb{E}\left[\left|X_{t_j}-X^N_{t_j}\right|\right]\leq c_1\Delta t_N+c_2\Delta t_N^{H+1/2},\qquad\forall N\in\mathbb{N},\tag{9.20}$$

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

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

$$X_t=e^{-t}X_0+\int_0^t e^{-(t-s)}B_s^H\,\mathrm{d}s,\tag{9.21}$$

H	p	$p_{\min}$	$p_{\max}$
0.1	0.618829	0.558371	0.679264
0.2	0.714901	0.652603	0.777104
0.3	0.809808	0.751078	0.868671
0.4	0.882603	0.822044	0.942833
0.5	0.997408	0.93889	1.05586
0.7	1.00102	0.945619	1.05755
0.9	1.00198	0.936873	1.06707

TABLE 3. Hurst parameter H, order p of strong convergence, and 95% confidence interval $p_{\min}$ and $p_{\max}$, for a number of Hurst values.

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.

More precisely, the Euler approximation is implemented for (9.10) with several values of H. We fix the time interval as $[0, T] = [0.0, 1.0]$, the initial condition as $X_0 \sim \mathcal{N}(0, 1)$, set the resolution for the target approximation to $N_{\text{tgt}} = 2^{18}$, choose the time steps for the convergence test as $\Delta t = 1/N$, with $N = 2^6, \ldots, 2^9$, and use $M = 200$ samples for the Monte-Carlo estimate of the strong error. The fBm noise term is generated with the $\mathcal{O}(N)$ fast Fourier transform (FFT) method of Davies and Harte, as presented in [11] (see also [19, Section 14.4]). Table 3 shows the obtained convergence estimates, for a number of Hurst parameters, which is illustrated in Figure 5, matching the theoretical estimate of $p = \min\{H + 1/2, 1\}$.

9.4. **Population dynamics with harvest.** Here, we consider a population dynamics problem modeled by a logistic equation with random coefficients, loosely inspired by [19, Section 15.2], with an extra term representing harvest:

$$\frac{\mathrm{d}X_t}{\mathrm{d}t} = \Lambda_t X_t(1 - \frac{X_t}{r}) - \alpha H_t \frac{X_t}{r + X_t}, \tag{9.22}$$

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

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

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.

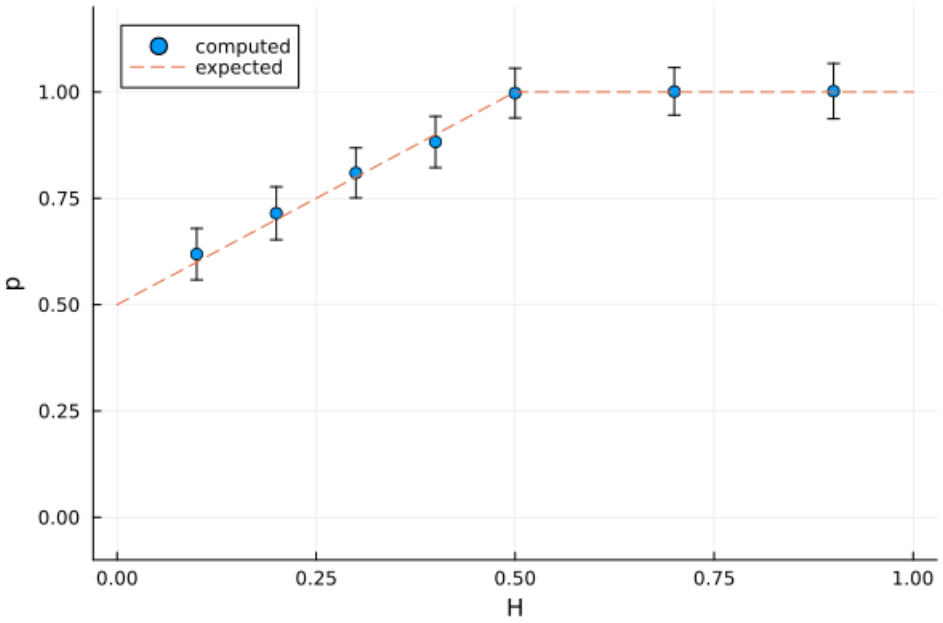

FIGURE 5. Order p of strong convergence for each value of the Hurst parameter H (scattered plot) along with the theoretical value $p = \min\{H + 1/2, 1\}$ (dashed line).

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 (9.22) 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 8.1 hold and the Euler method is of strong order 1.

Below, we simulate numerically the solutions of the above problem, with $\gamma = 0.8$, $\varepsilon = 0.3$, $r = 1.0$, and $\alpha = \gamma r = 0.64$ 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 $R = 1$. We take $M = 200$ samples for the Monte-Carlo estimate of the strong error of convergence. For the target solution, we solve the equation with a time mesh with $N_{\mathrm{tgt}} = 2^{18}$, while for the approximations we take $N = 2^i$, for $i = 4, \ldots, 9$.

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 4 shows the estimated strong error obtained for each mesh resolution, while Figure 6 illustrates the order of convergence, estimated to be close enough to the theoretical value of strong order 1. Finally, Figure 7 shows an approximation sequence of a sample path.

9.5. **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, inspired by the model in [7] (see also [30, Chapter 18] and [21] with this and other models).

There are a number of models for earthquake-type forcing, such as the ubiquotous Kanai-Tajimi and the Clough-Penzien models, where the noise has a characteristic

N	dt	error	std err
16	0.0625	0.00554	0.000253
32	0.0312	0.00269	0.000126
64	0.0156	0.00122	5.65e-5
128	0.00781	0.000637	2.87e-5
256	0.00391	0.000318	1.59e-5
512	0.00195	0.000169	8.22e-6

TABLE 4. 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 $M = 200$ sample paths for each fixed time step, with the target solution calculated with 262144 points. The order of strong convergence is estimated to be $p = 1.011$, with the 95% confidence interval $[0.9759, 1.0453]$.

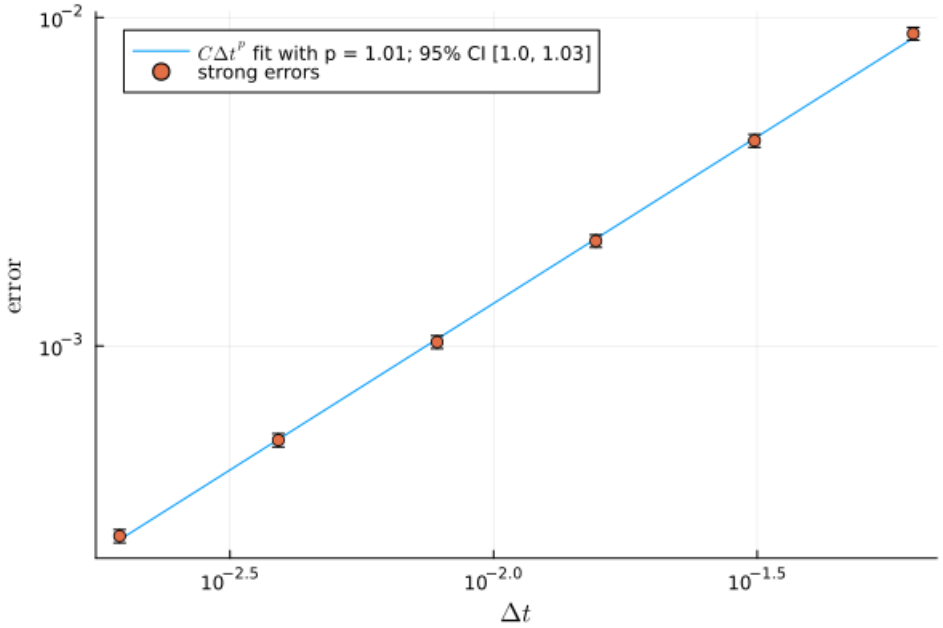

FIGURE 6. Order of convergence of the strong error of the Euler method for the population dynamics equation (9.22), based on Table 4.

spectral density, determined by the mechanical properties of the ground layer. The ideia, from [24], 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

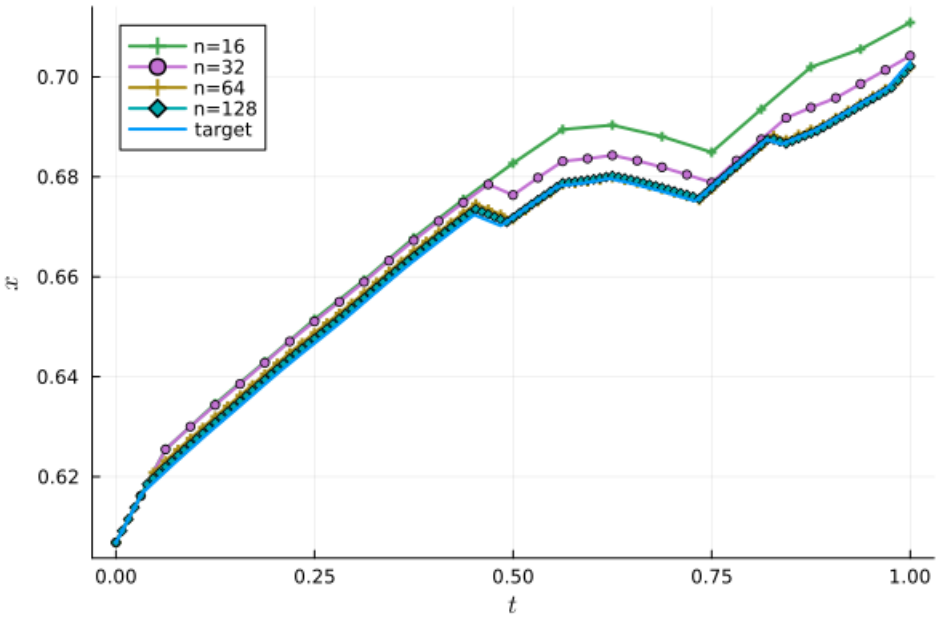

FIGURE 7. An approximation sequence of a sample path solution of the population dynamics equation (9.22).

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. We chose the later so we can illustrate the improved convergence for such type of 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

$$\ddot{X}_t + 2\zeta_0\omega_0\dot{X}_t + \omega_0^2 X_t = -\ddot{M}_t. \tag{9.23}$$

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 [7], 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:

$$M_t = \sum_{i=1}^{k} \gamma_i (t - \tau_i)_+^2 e^{-\delta_i (t-\tau_i)} \cos(\omega_i (t - \tau_i)), \tag{9.24}$$

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.

N	dt	error	std err
64	0.0312	2.09	0.18
128	0.0156	0.962	0.0761
256	0.00781	0.437	0.0356
512	0.00391	0.212	0.0175

TABLE 5. 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 $p = 1.106$, with the 95% confidence interval $[1.0181, 1.1935]$.

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. We set $T = 2.0$, as the final time. For the transport process, we set $k = 12$ and define the random parameters as the following exponential and uniform distributions: $\tau_i \sim$ Exponential(0.25), $\gamma_i \sim$ Uniform$(0.0, 4.0)$, $\delta_i \sim$ Uniform$(8.0, 12.0)$, and $\omega_i \sim$ Uniform$(8\pi, 32\pi)$.

For the mesh parameters, we set $N_{\text{tgt}} = 2^{18}$ and $N_i = 2^i$, for $i = 6, \ldots, 9$. For the Monte-Carlo estimate of the strong error, we choose $M = 100$. Table 5 shows the estimated strong error obtained with this setup, while Figure 8 illustrates the order of convergence. Figure 9 shows a sample ground motion and the corresponding ground acceleration and its envelope.

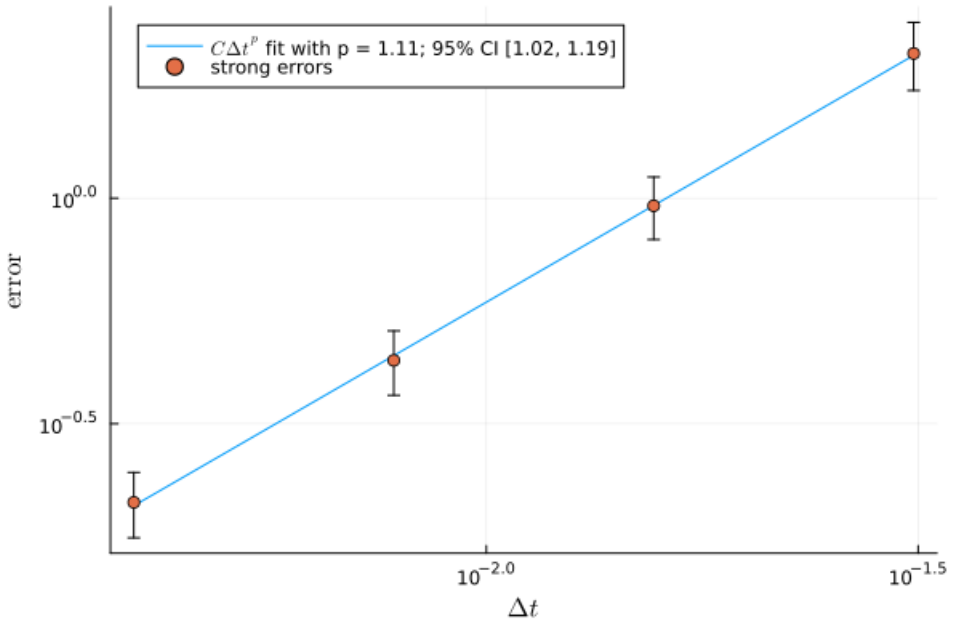

FIGURE 8. Order of convergence of the strong error of the Euler method for the mechanical structure model (9.23), based on Table 5.

9.6. **A toggle-switch model for gene expression.** Here, we consider the toggle-switch model in [2, Section 7.8], originated from [38]; see also [36].

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.

We consider the following simple model as discussed in [2, 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 by the system of equations

$$\begin{cases} \dfrac{\mathrm{d}X_t}{\mathrm{d}t} = \left(A_t + \dfrac{X_t^4}{a^4 + X_t^4}\right)\left(\dfrac{b^4}{b^4 + Y_t^4}\right) - \mu X_t, \\ \dfrac{\mathrm{d}Y_t}{\mathrm{d}t} = \left(B_t + \dfrac{Y_t^4}{c^4 + Y_t^4}\right)\left(\dfrac{d^4}{d^4 + X_t^4}\right) - \nu Y_t, \end{cases} \tag{9.25}$$

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 $\mathrm{d}B_t = (\mu_1 + \mu_2 \sin(\vartheta t))B_t\,\mathrm{d}t + \sigma \sin(\vartheta t)B_t\,\mathrm{d}W_t$.

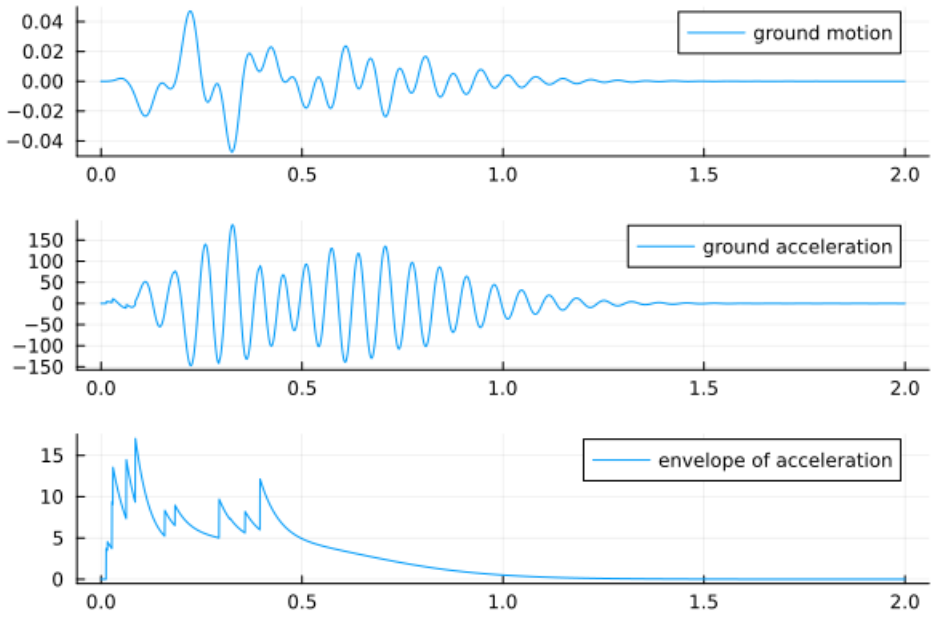

FIGURE 9. Sample ground motion and the corresponding excitation (ground acceleration) and envelope of excitation (just the exponential decay, without the oscillations) for the mechanical structure model (9.23).

In the simulations below, we use parameters similar to those in [2, 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 $B_0 = 0.2$;. The convergence estimate is done over the interval $[0, 5.0]$, while some illustrative solutions are over a longer interval $[0, 10.0]$.

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.

For the mesh parameters, we set $N_{\text{tgt}} = 2^{18}$ and $N_i = 2^i$, for $i = 5, \ldots, 9$. For the Monte-Carlo estimate of the strong error, we choose $M = 100$. Table 6 shows the estimated strong error obtained with this setup, while Figure 10 illustrates the order of convergence. Figure 11 shows a sample solution, while Figure 12 illustrates the two components (A_t, B_t) of a sample noise.

N	dt	error	std err
32	0.156	0.623	0.0354
64	0.0781	0.3	0.0169
128	0.0391	0.146	0.00818
256	0.0195	0.0716	0.00403
512	0.00977	0.0354	0.00202

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 $p = 1.034$, with the 95% confidence interval $[0.9865, 1.0809]$.

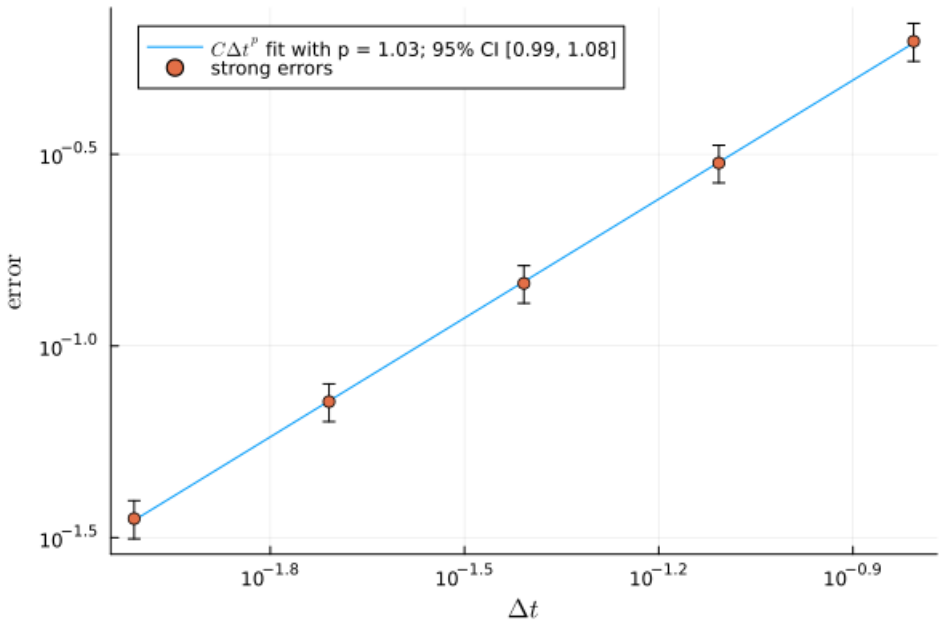

FIGURE 10. Order of convergence of the strong error of the Euler method for the toggle-switch model (9.25), based on Table 6.

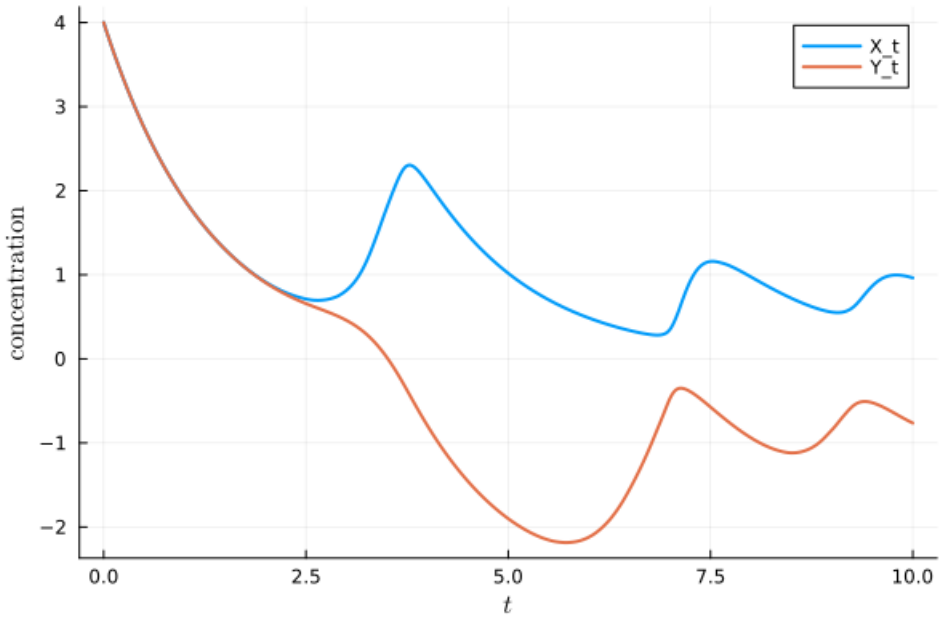

FIGURE 11. Evolution of a sample solution of the toggle-switch model (9.25).

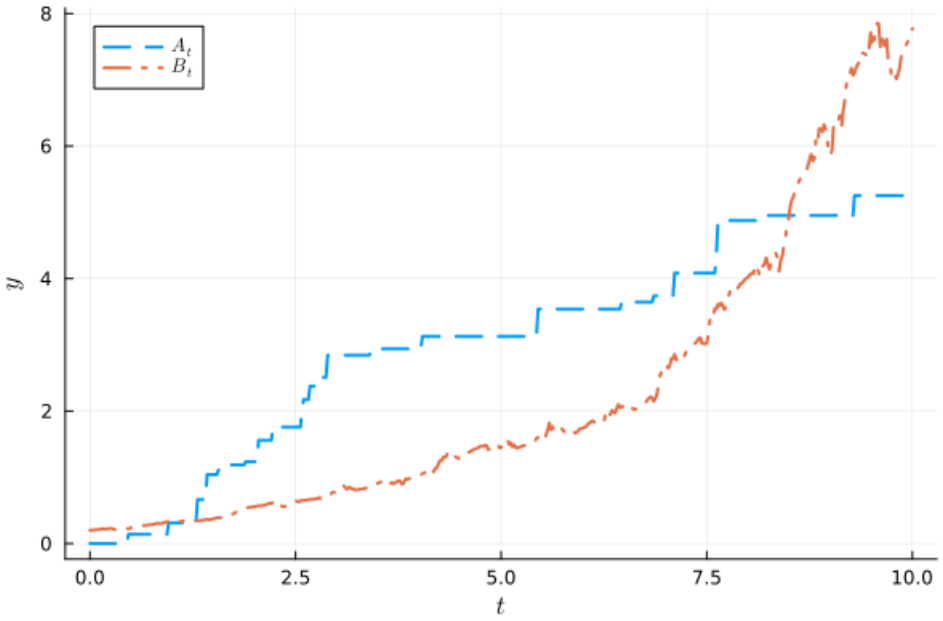

FIGURE 12. Noise sample paths for the toggle-switch model (9.25).

9.7. **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. [14]) given by

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

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

$$\frac{\mathrm{d}X_t}{\mathrm{d}t} = R_t X_t + R_t(C_t + O_t) + \nu O_t + \gamma. \tag{9.26}$$

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. [8, Chapter 3] in particular Table 3.1. Here, we consider the Dothan model ([8, 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 (9.26) 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 $O_0 = 0$, $\nu = 5$ and $\varepsilon = 0.8$, for the Ornstein-Uhlenbeck process $\{O_t\}_t$; $\lambda = 8.0$ and $C_i \sim \mathrm{Uniform}(0, 0.2)$, 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 $U_0 = X_0 + O_0 + R_0 = 1.2$. We set $T = 3.0$, as the final time.

For the mesh parameters, we set $N_{\mathrm{tgt}} = 2^{18}$ and $N_i = 2^i$, for $i = 6, \ldots, 9$. For the Monte-Carlo estimate of the strong error, we choose $M = 400$. Table 7 shows the estimated strong error obtained with this setup, while Figure 13 illustrates the order of convergence. Figure 14 shows a sample path of the noise, which is composed of three processes, while Figure 15 shows a sample path of the surplus.

9.8. **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 [35] and [13]. 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 [12] and [26]. It models inhomogeneous population

N	dt	error	std err
64	0.0469	0.223	0.0167
128	0.0234	0.115	0.00878
256	0.0117	0.0579	0.00405
512	0.00586	0.0286	0.00209

TABLE 7. 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 $p = 0.987$, with the 95% confidence interval $[0.9019, 1.0727]$.

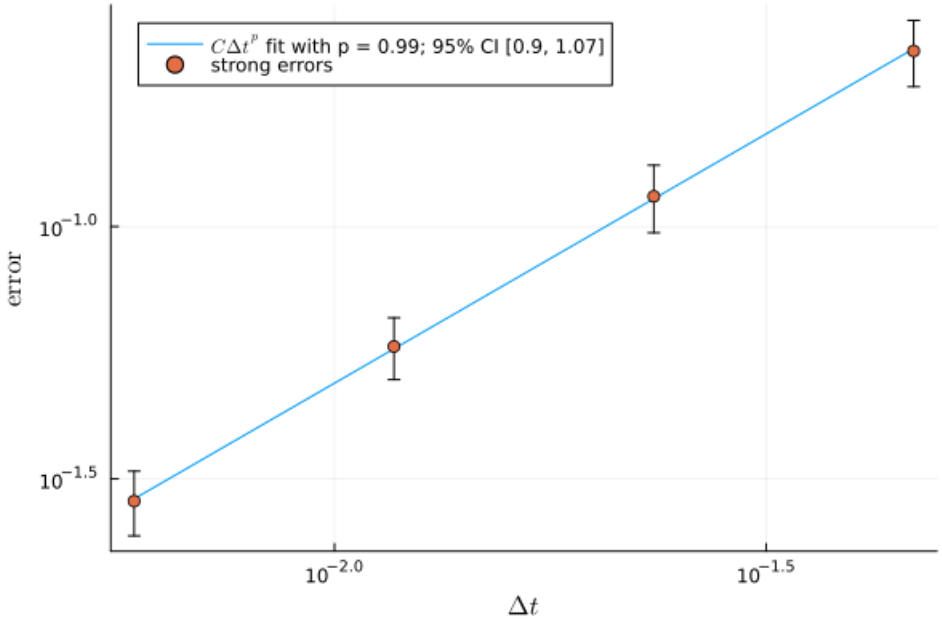

FIGURE 13. Order of convergence of the strong error of the Euler method for the actuarial risk model (9.26), based on Table 7.

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.

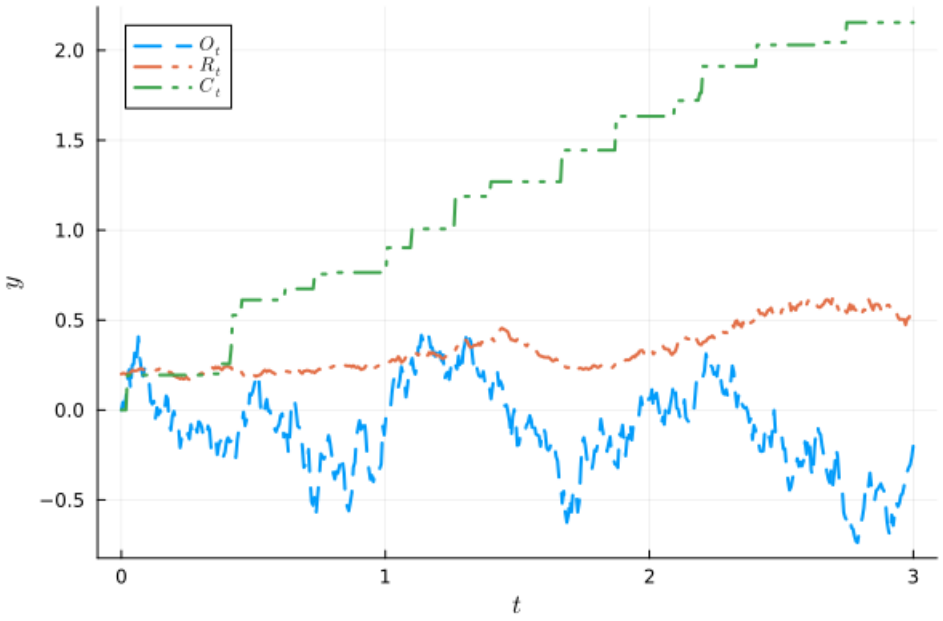

FIGURE 14. Sample noises for the risk model (9.26).

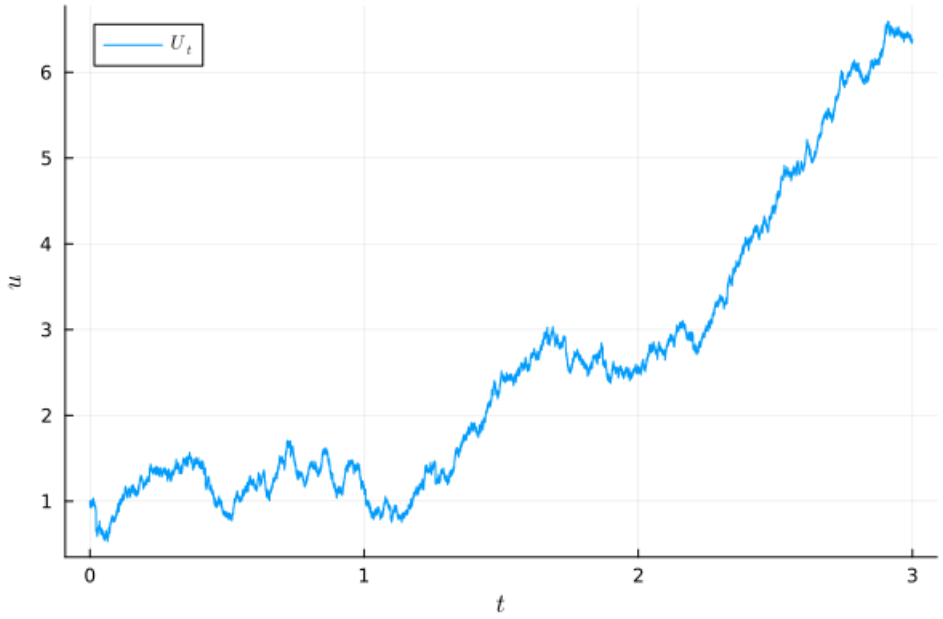

FIGURE 15. Sample surplus solution for the risk model (9.26).

The equation takes the form

$$\frac{\partial u}{\partial t} = \mu \frac{\partial^2 u}{\partial x^2} + \lambda u \left(1 - \frac{u}{u_m}\right), \quad (t, x) \in (0, \infty) \times (0, 1), \tag{9.27}$$

endowed with the boundary conditions

$$\frac{\partial u}{\partial x}(t,0) = -Y_t, \quad \frac{\partial u}{\partial x}(t,1) = 0, \tag{9.28}$$

and a given a initial condition

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

The unknown $u(t,x)$ represents the density of a given quantity 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 $\lambda_0 = 3.0$, base rate $a = 0.3$, exponential decay rate $\delta = 5.0$, and jump law following an exponential distribution with scale $\theta = 1/1.8$; and $\{O_t\}_t$ is an Orstein-Uhlenbeck process with $O_0 = 0.0$, time-scale $\tau = 0.005$, drift $\nu = 1/\tau$ and diffusion $\sigma = \tilde{\sigma}/\tau$, where $\tilde{\sigma} = 0.1$. 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 $0 \leq x \leq 1$ satisfies

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

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 9.4, 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 9.4, 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

N	dt	error	std err
32	0.0625	128.0	20.4
128	0.0156	30.2	4.82
512	0.00391	6.53	1.04

TABLE 8. 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 $p = 1.072$, with the 95% confidence interval $[0.9558, 1.1882]$.

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 $K_{\text{tgt}} = 2^9$.

For the Monte-Carlo estimate of the strong error, we choose $M = 40$. Table 8 shows the estimated strong error obtained with this setup, while Figure 16 illustrates the order of convergence.

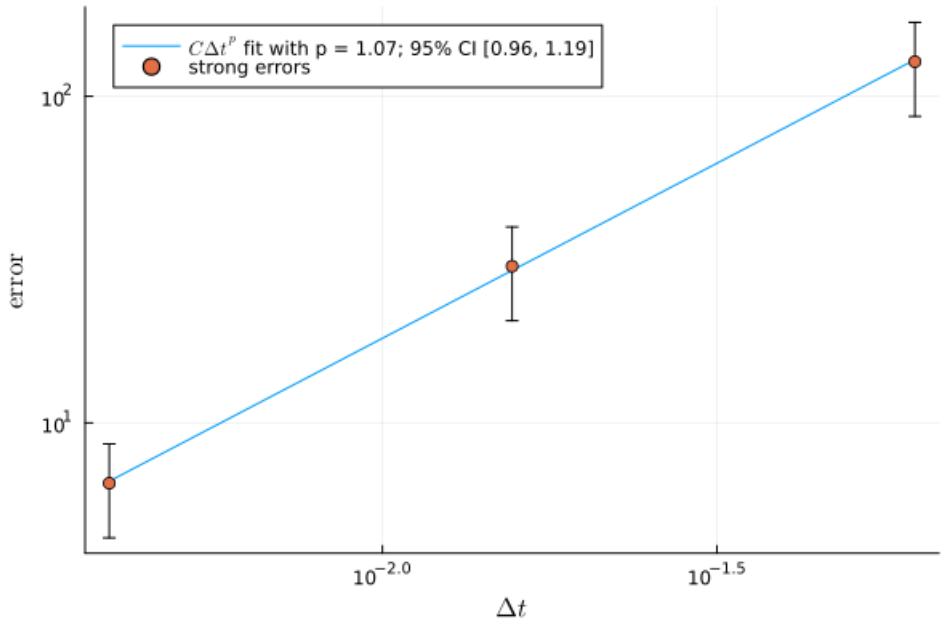

FIGURE 16. Order of convergence of the strong error of the Euler method for the Fisher-KPP model (9.27)-(9.28), based on Table 8.

References

[1] C. Aliprantis and K. Border, *Infinite dimensional analysis*, A hitchhiker's guide. Third edition. Springer, Berlin, 2006.

[2] Y. Asai, *Numerical Methods for Random Ordinary Differential Equations and their Applications in Biology and Medicine.* Dissertation, Institut für Mathematik, Goethe Universität Frankfurt am Main, 2016.

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

[4] C. Bender, An Itô formula for generalized functionals of a fractional Brownian motion with arbitrary Hurst parameter, *Stochastic Processes and their Applications,* 104 (2003), 81–106.

[5] J. Bezanson, A. Edelman, S. Karpinski, and V. B. Shah, Julia: A fresh approach to numerical computing, *Siam Review,* 59 (2017), no. 1, 65–98.

[6] F. Biagini, Y. Hu, B. Øksendal, and T. Zhang. *Stochastic Calculus for Fractional Brownian Motion and Applications,* Springer-Verlag, London, 2008.

[7] J. L. Bogdanoff, J. 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 (1961), no. 2, 293–310.

[8] , D. Brigo, F. Mercurio, *Interest Rate Models - Theory and Practice,* Springer-Verlag, Berlin, Heidelberg, 2006.

[9] D. S. Clark, Short proof of a discrete Gronwall inequality, *Discrete Applied Mathematics,* Vol. 16 (1987) no. 3, 279–281.

[10] E. A. Coddington and N. Levinson, *Theory of Ordinary Differential Equations,* New York: McGraw-Hill, 1987.

[11] A. B. Dieker and M. Mandjes, On spectral simulation of fractional Brownian motion, *Probability in the Engineering and Informational Sciences,* 17 (2003), 417–434.

[12] R. A. Fisher, The wave of advance of advantageous genes, *Annals of Eugenics,* 7 (1937), no. 4, 355–369.

[13] M. I. Freidlin and A. D. Wentzell, Reaction-diffusion equations with randomly perturbed boundary conditions, *Ann. Probab.* 20 (1992), no. 2, 963–986.

[14] H. U. Gerber, E. S. W. Shiu, On the time value of ruin, North American Actuarial J. vol. 2 (1998), no. 1, 48–72.

[15] H. Gjessing, H. Holden, T. Lindstrøn, B. Øksendal, J. Ubøe, and T.-S. Zhang, The Wick product, *Vol. 1 Proceedings of the Third Finnish-Soviet Symposium on Probability Theory and Mathematical Statistics,* Turku, Finland, August 13–16, 1991, edited by H. Niemi, G. Högnas, A. N. Shiryaev and A. V. Melnikov, Berlin, Boston: De Gruyter, 1993, pp. 29-67.

[16] V. Girault and P.-A. Raviart, *Finite-Element Approximation of the Navier-Stokes Equations,* Lecture Notes in Mathematics, vol. 749, Springer-Verlag, Berlin, Heidelberg, 1981.

[17] T. H. Gronwall, Note on the derivatives with respect to a parameter of the solutions of a system of differential equations, *Ann. of Math.* (2) 20 (1919), 292–296.

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

[19] X. Han and P. E. Kloeden, *Random Ordinary Differential Equations and Their Numerical Solution,* Probability Theory and Stochastic Modelling, vol. 85, Springer Singapore, 2017.

[20] D. J. Higham and P. E. Kloeden, *An Introduction to the Numerical Simulation of Stochastic Differential Equations,* Volume 169 of Other Titles in Applied Mathematics, SIAM, 2021.

[21] G. W. Housner and Paul C. Jennings, Generation of artificial Earthquakes, Journal of the Engineering Mechanics Division, 90 (1964), no. 1.

[22] A. Jentzen and P.E. Kloeden, *Taylor Approximations of Stochastic Partial Differential Equations,* CBMS Lecture series, SIAM, Philadelphia, 2011.
[23] A. Jentzen, P.E. Kloeden, and A. Neuenkirch, Pathwise approximation of stochastic differential equations on domains: Higher order convergence rates without global Lipschitz coefficients, *Numer. Math.* 112 (2009), no. 1, 41–64.
[24] K. Kanai, Semi-empirical formula for the seismic characteristics of the ground, *Bull. Earthq. Res. Inst.,* Vol. 35 (1957), University of Tokyo.
[25] A. Jentzen, P.E. Kloeden,, Pathwise convergent higher order numerical schemes for random ordinary differential equations, *Proc. R. Soc. A* 463 (2007), 2929–2944.
[26] A. N. Kolmogorov, I. G. Petrovskii, N. S. Piskunov, A study of the diffusion equation with increase in the amount of substance, and its application to a biological problem. *Bull. Moscow Univ. Math. Mech.* 1 (1937), 1–26.
[27] H.-H. Kuo, *Introduction to Stochastic Integration,* Universitext, Springer New York, NY, 2006.
[28] B. B. Mandelbrot and J. W. Van Ness, Fractional Brownian motions, fractional noises and applications, *Siam Review,* Vol. 10 (1968), no. 4, 422–437.
[29] Y. S. Mishura, *Stochastic calculus for fractional Brownian motion and related processes,* Lecture Notes in Mathematics 1929, Springer-Verlag, Berlin, Heidelberg, 2008.
[30] T. Neckel and F. Rupp, *Random Differential Equations in Scientific Computing,* Versita, London, 2013.
[31] B. Øksendal, *Stochastic Differential Equations - An Introduction with Applications,* Universitext, Springer-Verlag Berlin Heidelberg, 2003.
[32] P. E. Protter, *Stochastic Integration and Differential Equations,* 2nd Edition, Springer-Verlag, Berlin Heidelberg New York, 2005.
[33] C. Rackauckas and Q. Nie, DifferentialEquations.jl - a performant and feature-rich ecosystem for solving differential equations in Julia, *The Journal of Open Research Software*, 5 (2017), no. 1, 1–15.
[34] 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_conv_em`.
[35] R. B. Salako and W. Shen, Long time behavior of random and nonautonomous Fisher-KPP equations: Part I - stability of equilibria and spreading speeds, *J. Dyn. Diff. Eqs.,* 33 (2021), 1035–1070.
[36] M. Strasser, F. J. Theis, and C. Marr, Stability and multiattractor dynamics of a toggle switch based on a two-stage model of stochastic gene expression, *Biophysical J.,* 102 (2012), 19–29.
[37] H. Tajimi, A statistical method of determining the maximum response of a building during an Earthquake, *Proceedings of the Second World Conference on Earthquake Engineering,* Tokyo and Kyoto, Japan, vol. II, 1960.
[38] B. Verd, A. Crombach, and J. Jaeger, Classification of transient behaviours in a time-dependent toggle switch model, *BMC Systems Biology,* 8 (2014), no. 43.
[39] P. Wang, Y. Cao, X. Han, and P. Kloeden, Mean-square convergence of numerical methods for random ordinary differential equations, *Numerical Algorithms,* vol. 87 (2021), 299–333.

(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`