.

IMPROVED ERROR ESTIMATE FOR THE ORDER OF STRONG CONVERGENCE OF THE EULER METHOD FOR RANDOM ORDINARY DIFFERENTIAL EQUATIONS

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 $\mathrm{d}X_t/\mathrm{d}t = f(t, X_t, Y_t)$ 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. Here, it is proved that, in many typical cases, further conditions on the noise can be exploited so that the strong convergence is actually of order 1, regardless of the Hölder regularity of the sample paths. This applies for instance to additive or multiplicative Itô process noises (such as Wiener, Ornstein-Uhlenbeck, and geometric Brownian motion processes); to point-process noises (such as Poisson point processes and Hawkes self-exciting processes, which even have jump-type discontinuities); and to transport-type processes with sample paths of bounded variation. The result is based on a novel approach, estimating the global error as an iterated integral over both large and small mesh scales, and 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 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)**:

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

on a time interval $I = [0, T]$, with $T > 0$, and where the noise $\{Y_t\}_{t\in I}$ is a given stochastic process. This can be a scalar or a system of equations and the noise can also be either scalar or vector valued. The sample space is denoted by Ω.

Date: June 16, 2023.

2000 *Mathematics Subject Classification.* 76D05, 76D06, 35Q30, 37L99.

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).

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 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. There are other important notions of convergence, such as weak convergence, p-th mean convergence, and pathwise convergence (see e.g. [17, 18, 20]), but we focus on strong convergence, here.

Under certain regularity conditions on f, it is proved in [16] 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 *pathwise* sense with order θ with respect to the time step. Under global uniformity conditions on the f, the same approach in [16] also yields the convergence of order θ in the *strong* sense (1.3).

Our aim is to show that, in many 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, or an Itô process, for which the convergence is of strong order 1. It is also the case for 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 $0 < H < 1/2$, which is still higher than the Hölder exponent H of the sample paths.

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. [21, 20]), but those are not the subject of interest here. The possibility of extending the improved order of strong convergence to pathwise convergence seems feasible for the case of sample paths of bounded variation but not so much for Itô process noises for which the Itô isometry is of fundamental importance. These will be investigated in a future oportunity.

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 3.1), as it is done for

approximations of stochastic differential equations, namely

$$\begin{aligned} X_{t_j} - X_{t_j}^N = X_0 - X_0^N &+ \int_0^{t_j} \left(f(s, X_s, Y_s) - f(s, X_{\tau^N(s)}, Y_s) \right) \, \mathrm{d}s \\ &+ \int_0^{t_j} \left(f(s, X_{\tau^N(s)}, Y_s) - f(s, X^N_{\tau^N(s)}, Y_s) \right) \, \mathrm{d}s \\ &+ \int_0^{t_j} \left(f(s, X^N_{\tau^N(s)}, Y_s) - f(\tau^N(s), X^N_{\tau^N(s)}, Y_{\tau^N(s)}) \right) \, \mathrm{d}s, \end{aligned} \tag{1.4}$$

for $j = 1, \ldots, N$, where τ^N is a piecewise constant function with jumps at the mesh points t_j, defined in (3.2).

The first term vanishes due to the initial condition $X_0^N = X_0$. The second term only depends on the solution and can be easily 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 uniformly globally Lipschitz continuous with respect to x. With those, we obtain the following basic bound for the global error (see Lemma 4.2)

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

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

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

either in the sense of a Riemann-Stieltjes integral or of an Itô integral. The first sense fits the case of noises with bounded total variation, while the second one fits the case of an Itô process noise. In any case, we bound the global error term using the Fubini

Theorem,

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

Then, we find that

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

In the case of an Itô process noise, we assume

$$\mathrm{d}F_t = A_t \, \mathrm{d}t + B_t \, \mathrm{d}W_t,$$

with adapted processes $\{A_t\}_t$, $\{B_t\}_t$, which may actually depend on $\{Y_t\}_t$, so that multiplicative noise is allowed. Then, in this case, we bound the right hand side using the Lyapunov inequality and the Itô isometry:

$$\mathbb{E}\left[\left\| \int_0^{t_j} (\tau^N(\xi) + \Delta t_N - \xi) \, \mathrm{d}F_\xi \right\| \right]$$
$$\leq \int_0^{t_j} (\tau^N(\xi) + \Delta t_N - \xi) \mathbb{E}[\|A_\xi\|] \, \mathrm{d}\xi + \left(\int_0^{t_j} (\tau^N(\xi) + \Delta t_N - \xi)^2 \mathbb{E}[\|B_\xi\|^2] \, \mathrm{d}\xi \right)^{1/2}$$
$$\leq \Delta t_N \left(\int_0^{t_j} \mathbb{E}[\|A_\xi\|] \, \mathrm{d}\xi + \left(\int_0^{t_j} \mathbb{E}[\|B_\xi\|^2] \, \mathrm{d}\xi \right)^{1/2} \right).$$

which yields the strong order 1 convergence, provided the integrals are finite.

In the case of noises with bounded variation, we may actually relax the above condition and assume the steps are bounded by a process $\{\bar{F}_t\}_{t \in I}$ with monotonic non-decreasing sample paths,

$$\left\| f(s, X^N_{\tau^N(s)}, Y_s) - f(\tau^N(s), X^N_{\tau^N(s)}, Y_{\tau^N(s)}) \right\| \leq \bar{F}_s - \bar{F}_{\tau^N(s)}.$$

Using the monotonicity, this implies

$$\mathbb{E}\left[\left\| \int_0^{t_j} (\tau^N(\xi) + \Delta t_N - \xi) \, \mathrm{d}\bar{F}_\xi \right\| \right] \leq \Delta t_N \left(\mathbb{E}[\bar{F}_{t_j}] - \mathbb{E}[\bar{F}_0] \right),$$

yielding, again, strong order 1 convergence.

These two cases are treated in Section 5 and Section 6, with the bounded variation case in Section 5, and the Itô process noise case in Section 6.

The core results in these sections are Lemma 5.1 and Theorem 5.1, for the bounded variation case, and Lemma 6.1 and Theorem 6.1, for the Itô process noise case. The

conditions in such results, however, are not readily verifiable. With that in mind, Theorem 5.2 and Theorem 6.2 give more explicit conditions for each of these two cases. Essentially, $f = f(t, x, y)$ is required to have minimal regularity in the sense of differentiability and growth conditions, while the noise $\{Y_t\}_{t\in I}$ is either required to have sample paths of bounded variation or to be an Itô process noise.

These two types of noises can also appear at same time, in a given equation or system of equations, as treated in Section 7. This can be regarded as a vector-valued noise, where the components of the noise may either be of bounded variation or of Itô type. See Theorem 7.1 and Theorem 7.2.

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 a system of linear equations with all sorts of noises, emcompassing 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 (8.11) and (8.13))

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

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 [31], such as a logistic model of population dynamics with random coefficients, loosely inspired by [17, 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 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 [5] (see also [27, Chapter 18], [19], and [22], 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 [12] and [6]; and a Fisher-KPP partial differential equation with random boundary conditions, as inspired by the works of [32] and [11] (see also [10] and [23]), 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 solution for RODEs, we refer the reader to [17, Section 2.1] and [27, 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) *f is globally Lipschitz continuous on x, uniformly in t and y, i.e. there exists a constant $L_X \geq 0$ such that*

$$\|f(t,x_1,y) - f(t,x_2,y)\| \leq L_X \|x_1 - x_2\|, \quad \forall t \in I,\ \forall x_1, x_2, y \in \mathbb{R}^k. \tag{2.1}$$

(ii) *The mapping $(t,x) \mapsto f(t,x,Y_t)$ satisfies the Carathéodory conditions:*
   (a) *The mapping $x \mapsto f(t,x,Y_t(\omega))$ is continuous in $\mathbb{R}^d$, for almost every $(t,\omega) \in I \times \Omega$;*
   (b) *The mapping $t \mapsto f(t,x,Y_t(\omega))$ is Lebesgue measurable in $t \in I$, for each $x \in \mathbb{R}^d$ and each sample path $t \mapsto Y_t(\omega)$;*
   (c) *The bound $\|f(t,x,Y_t)\| \leq M_t + L_X\|x\|$ holds for all $t \in I$ and all $x \in \mathbb{R}^d$, where $\{M_t\}_{t\in I}$ is a real stochastic process with Lebesgue integrable sample paths $t \mapsto M_t(\omega)$ on $t \in I$.*

Under these assumptions, for each sample value in Ω, the integral equation

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

has a unique solution, 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 [8, Theorem 1.1]). Moreover, the mapping $(t,\omega) \mapsto X_t(\omega)$ is measurable (see [17, Section 2.1.2]) and, hence, give rise to a well-defined stochastic process $\{X_t\}_{t\in I}$.

Each sample path solution $t \mapsto X_t(\omega)$ is bounded by

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

For the strong convergence of the Euler approximation, we also need to control the expectation of the solution, among other things. With that in mind, we have the following useful result.

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

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

and

$$\int_0^T \mathbb{E}[|M_s|]\, \mathrm{d}s < \infty. \tag{2.5}$$

Then,

$$\mathbb{E}[\|X_t\|] \leq \left(\mathbb{E}[\|X_0\|] + \int_0^t \mathbb{E}[|M_s|]\, \mathrm{d}s\right) e^{L_X t}, \quad t \in I. \tag{2.6}$$

Proof. Thanks to (2.3), 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 continuous, then the integrand in (2.2) is 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 (2.5) 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. But we leave these cases to be handled in the applications.

3. Integral formula for the global pathwise error

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

Lemma 3.1. *Under Standing Hypothesis 2.1, the Euler approximation (1.2) for any pathwise solution of the random ordinary differential equation (1.1) satisfies the global error formula*

$$\begin{aligned} X_{t_j} - X_{t_j}^N = X_0 - X_0^N &+ \int_0^{t_j} \left(f(s, X_s, Y_s) - f(s, X_{\tau^N(s)}, Y_s)\right)\, \mathrm{d}s \\ &+ \int_0^{t_j} \left(f(s, X_{\tau^N(s)}, Y_s) - f(s, X^N_{\tau^N(s)}, Y_s)\right)\, \mathrm{d}s \\ &+ \int_0^{t_j} \left(f(s, X^N_{\tau^N(s)}, Y_s) - f(\tau^N(s), X^N_{\tau^N(s)}, Y_{\tau^N(s)})\right)\, \mathrm{d}s, \end{aligned} \tag{3.1}$$

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

$$\tau^N(t) = \max_j \{j\Delta t_N;\ j\Delta t_N \le t\} = \left[\frac{t}{\Delta t_N}\right] \Delta t_N = \left[\frac{tN}{T}\right] \frac{T}{N}. \tag{3.2}$$

Proof. Under 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 yields

$$\begin{aligned} X_{t_j} - X^N_{t_j} =& X_{t_{j-1}} - X^N_{t_{j-1}} \\ =& \int_{t_{j-1}}^{t_j} \left(f(s, X_s, Y_s) - f(s, X_{t_{j-1}}, Y_s)\right)\, \mathrm{d}s \\ &+ \int_{t_{j-1}}^{t_j} \left(f(s, X_{t_{j-1}}, Y_s) - f(s, X^N_{t_{j-1}}, Y_s)\right)\, \mathrm{d}s \\ &+ \int_{t_{j-1}}^{t_j} \left(f(s, X^N_{t_{j-1}}, Y_s) - f(t_{j-1}, X^N_{t_{j-1}}, Y_{t_{j-1}})\right)\, \mathrm{d}s. \end{aligned} \tag{3.3}$$

Now we iterate the time steps (3.3) to find that

$$\begin{aligned} X_{t_j} - X^N_{t_j} = X_0 - X^N_0 + \sum_{i=1}^{j} \Bigg(& \int_{t_{i-1}}^{t_i} \left(f(s, X_s, Y_s) - f(s, X_{t_i}, Y_s)\right)\, \mathrm{d}s \\ &+ \int_{t_{i-1}}^{t_i} \left(f(s, X_{t_{i-1}}, Y_s) - f(s, X^N_{t_{i-1}}, Y_s)\right)\, \mathrm{d}s \\ &+ \int_{t_{i-1}}^{t_i} \left(f(s, X^N_{t_{i-1}}, Y_s) - f(t_{i-1}, X^N_{t_{i-1}}, Y_{t_{i-1}})\right)\, \mathrm{d}s \Bigg). \end{aligned}$$

Using the jump function τ^N defined by (3.2), we rewrite the above expression as (3.1). □

Remark 3.1. Strictly speaking, we only need condition (ii) from Standing Hypothesis 2.1 in order to deduce (4.3), but since we need (i) for the strong convergence anyways, it is simpler to state the result as in Lemma 4.2.

4. 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 [14] (see also [7]). Its proof follows from [14, Lemma V.2.4] by taking $n = j$, $a_n = e_j$, $b_n = 0$, $c_n = b$, and $\lambda = a$.

Lemma 4.1 (Discrete Gronwall Lemma)**.** *Let $(e_j)_j$ be a (finite or infinite) sequence of positive numbers starting at $j = 0$ and satisfying*

$$e_j \leq a \sum_{i=0}^{j-1} e_i + b, \tag{4.1}$$

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

$$e_j \leq b e^{aj}, \qquad \forall j. \tag{4.2}$$

We are now ready to start proving our basic estimate for the global pathwise error.

Lemma 4.2. *Under Standing Hypothesis 2.1, the global error (3.1) is estimated as*

$$\begin{aligned} \|X_{t_j} - X^N_{t_j}\| \leq \Bigg(\|X_0 - X^N_0\| + L_X \int_0^{t_j} \|X_s - X_{\tau^N(s)}\| \, \mathrm{d}s \\ \left\| \int_0^{t_j} \left(f(s, X^N_{\tau^N(s)}, Y_s) - f(\tau^N(s), X^N_{\tau^N(s)}, Y_{\tau^N(s)}) \right) \, \mathrm{d}s \right\| \Bigg) e^{L_X t_j}. \end{aligned} \tag{4.3}$$

for $j = 1, \ldots, N$, where τ^N is given by (3.2).

Proof. We estimate the first two integrals in (3.1). For the first one, we use (2.1), so that

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

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

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

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

$$\|f(s, X_t, Y_s) - f(s, X^N_t, Y_s)\| \leq L_X \|X_t - X^N_t\|,$$

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

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

With these two estimates, we bound (3.1) as

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

This can be cast in the form of (4.1). Then, using the discrete version Lemma 4.1 of the Gronwall Lemma, we obtain (4.3). □

The first term in the right hand side of (4.3) 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 (4.3) 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 4.3. *Under Standing Hypothesis 2.1, it follows that*

$$\int_0^{t_j} \left\| X_s - X_{\tau^N(s)} \right\| \, \mathrm{d}s \leq \Delta t_N \int_0^{t_j} (M_s + L_X \|X_s\|) \, \mathrm{d}s. \tag{4.4}$$

Proof. By assumption, we have $\|f(t, X_t, Y_t)\| \leq M_t + L_X \|X_t\|$, for all $t \in I$ and all sample paths. Thus,

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

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

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

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 (4.4). □

Combining the two previous results we obtain the following:

Proposition 4.1. *Under Standing Hypothesis 2.1, suppose further that (2.4) and (2.5) hold and that, for some constant $C_0 \geq 0$,*

$$\mathbb{E}[\|X_0 - X_0^N\|] \leq C_0 \Delta t_N, \qquad N \in \mathbb{N}. \tag{4.5}$$

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

$$\mathbb{E}\left[\|X_{t_j} - X_{t_j}^N\|\right] \leq \left(C_0 \Delta t_N + \Delta t_N L_X \left(\mathbb{E}[\|X_0\|] + \int_0^{t_j} \mathbb{E}[M_\xi] \, \mathrm{d}\xi\right) e^{L_X t_j}\right.$$
$$\left. \mathbb{E}\left[\left\| \int_0^{t_j} \left(f(s, X_{\tau^N(s)}^N, Y_s) - f(\tau^N(s), X_{\tau^N(s)}^N, Y_{\tau^N(s)})\right) \, \mathrm{d}s \right\|\right]\right) e^{L_X t_j}. \tag{4.6}$$

Proof. Estimate (4.6) is obtained by taking the expectation of (4.3) in Lemma 4.2 and properly estimating the first two terms on the right hand side. The first term is handled with the assumption (4.5). We just need to take care of the second term.

Under Standing Hypothesis 2.1, estimate Lemma 4.3 applies and inequality (4.4) holds. Using (2.4) and (2.5), that inequality yields

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

Using now (2.3), we obtain

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

Thus,

$$\int_0^{t_j} \mathbb{E}[\|X_s - X_{\tau^N(s)}\|] \, \mathrm{d}s \leq \Delta t_N \left(\mathbb{E}[\|X_0\|] + \int_0^{t_j} \mathbb{E}[M_\xi] \, \mathrm{d}\xi\right) e^{L_X t_j}. \tag{4.7}$$

Now we look at Lemma 4.2. Taking the expectation of the global error formula (4.3) gives

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

Using now estimate (4.7) and condition (4.5), we find (4.6), completing the proof. □

5. The case of noise with sample paths of bounded variation

Here, the noise $\{Y_t\}_{t\in I}$ is *not* assumed to be an Itô process noise, but, instead, that the steps can be controlled by a monotonic nondecreasing process with finite expected growth. This fits well the typical case of point processes, such as renewal-reward processes, Hawkes process, and the like.

More precisely, we have the following result.

Lemma 5.1. *Besides Standing Hypothesis 2.1, suppose that, for all* $0 \le s \le T$,

$$\left\| f(s, X^N_{\tau^N(s)}, Y_s) - f(\tau^N(s), X^N_{\tau^N(s)}, Y_{\tau^N(s)}) \right\| \le \bar{F}_s - \bar{F}_{\tau^N(s)}, \tag{5.1}$$

where $\{\bar{F}_t\}$ *is a real-valued stochastic process with monotonic nondecreasing sample paths and with*

$$\mathbb{E}[\bar{F}_t] \text{ uniformly bounded on } t \in I. \tag{5.2}$$

Then,

$$\mathbb{E}\left[\left\| \int_0^t \left(f(s, X^N_{\tau^N(s)}, Y_s) - f(\tau^N(s), X^N_{\tau^N(s)}, Y_{\tau^N(s)}) \right) \, \mathrm{d}s \right\|\right] \le (\mathbb{E}[\bar{F}_t] - \mathbb{E}[\bar{F}_0])\Delta t_N, \tag{5.3}$$

for all $0 \le t \le T$ *and every* $N \in \mathbb{N}$.

Proof. Let $N \in \mathbb{N}$. From the assumption (5.1) we have

$$\mathbb{E}\left[\left\| f(s, X^N_{\tau^N(s)}, Y_s) - f(\tau^N(s), X^N_{\tau^N(s)}, Y_{\tau^N(s)}) \right\|\right] \le \mathbb{E}[\bar{F}_s] - \mathbb{E}[\bar{F}_{\tau^N(s)}],$$

for every $0 \le s \le T$. Thus, upon integration,

$$\begin{aligned}
&\mathbb{E}\left[\left\| \int_0^t \left(f(s, X^N_{\tau^N(s)}, Y_s) - f(\tau^N(s), X^N_{\tau^N(s)}, Y_{\tau^N(s)}) \right) \, \mathrm{d}s \right\|\right] \\
&\quad \le \int_0^t \mathbb{E}\left[\left\| f(s, X^N_{\tau^N(s)}, Y_s) - f(\tau^N(s), X^N_{\tau^N(s)}, Y_{\tau^N(s)}) \right\|\right] \, \mathrm{d}s \\
&\quad \le \int_0^t \left(\mathbb{E}[\bar{F}_s] - \mathbb{E}[\bar{F}_{\tau^N(s)}] \right) \, \mathrm{d}s.
\end{aligned}$$

Now we need to bound the right hand side above. When $0 \le t \le t_1 = \Delta t_N$, we have $\tau^N(s) = 0$ for all $0 \le s < t_1$, so that,

$$\int_0^t (\mathbb{E}[\bar{F}_s] - \mathbb{E}[\bar{F}_{\tau^N(s)}]) \, \mathrm{d}s = \int_0^t (\mathbb{E}[\bar{F}_s] - \mathbb{E}[\bar{F}_0]) \, \mathrm{d}s.$$

Using the monotonicity of $\{\bar{F}_t\}$ and the condition that $t \le \Delta t_N$,

$$\begin{aligned}
\int_0^t (\mathbb{E}[\bar{F}_s] - \mathbb{E}[\bar{F}_{\tau^N(s)}]) \, \mathrm{d}s \le \int_0^t (\mathbb{E}[\bar{F}_t] - \mathbb{E}[\bar{F}_0]) \, \mathrm{d}s \\
= (\mathbb{E}[\bar{F}_t] - \mathbb{E}[\bar{F}_0])t \le (\mathbb{E}[\bar{F}_t] - \mathbb{E}[\bar{F}_0])\Delta t_N.
\end{aligned}$$

When $\Delta t_N \le t \le T$, we split the integration of the second term at time $s = t_1 = \Delta t_N$ and write

$$\int_0^t (\mathbb{E}[\bar{F}_s] - \mathbb{E}[\bar{F}_{\tau^N(s)}])\,\mathrm{d}s = \int_0^t \mathbb{E}[\bar{F}_s]\,\mathrm{d}s - \int_0^{t_1} \mathbb{E}[\bar{F}_{\tau^N(s)}]\,\mathrm{d}s - \int_{t_1}^t \mathbb{E}[\bar{F}_{\tau^N(s)}]\,\mathrm{d}s.$$

Using the monotonicity together with the fact that $s - \Delta t_N \le \tau^N(s) \le s$ for all $\Delta t_N \le s \le T$,

$$\begin{aligned}\int_0^t (\mathbb{E}[\bar{F}_s] - \mathbb{E}[\bar{F}_{\tau^N(s)}])\,\mathrm{d}s &\le \int_0^t \mathbb{E}[\bar{F}_s]\,\mathrm{d}s - \int_0^{\Delta t_N} \mathbb{E}[\bar{F}_0]\,\mathrm{d}s - \int_{\Delta t_N}^t \mathbb{E}[\bar{F}_{s-\Delta t_N}]\,\mathrm{d}s \\ &= \int_0^t \mathbb{E}[\bar{F}_s]\,\mathrm{d}s - \int_0^{\Delta t_N} \mathbb{E}[\bar{F}_0]\,\mathrm{d}s - \int_0^{T-\Delta t_N} \mathbb{E}[\bar{F}_s]\,\mathrm{d}s \\ &= \int_{t-\Delta t_N}^t \mathbb{E}[\bar{F}_s]\,\mathrm{d}s - \mathbb{E}[\bar{F}_0]\Delta t_N.\end{aligned}$$

Using again the monotonicity yields

$$\int_0^t (\mathbb{E}[\bar{F}_s] - \mathbb{E}[\bar{F}_{\tau^N(s)}])\,\mathrm{d}s \le \int_{t-\Delta t_N}^t \mathbb{E}[\bar{F}_t]\,\mathrm{d}s - \mathbb{E}[\bar{F}_0]\Delta t_N = (\mathbb{E}[\bar{F}_t] - \mathbb{E}[\bar{F}_0])\Delta t_N.$$

Putting the estimates together and using the boundedness (5.2) proves (5.3). □

With Lemma 5.1 at hand, combined with the results in the previous sections, we prove our first main result.

Theorem 5.1. *Under Standing Hypothesis 2.1, suppose also that (2.4), (2.5), (4.5), (5.1), and (5.2) hold. Then, the Euler scheme (1.2) is of strong order 1, i.e.*

$$\max_{j=0,\dots,N} \mathbb{E}\left[\left\|X_{t_j} - X^N_{t_j}\right\|\right] \le C\Delta t_N, \qquad \forall N \in \mathbb{N}, \tag{5.4}$$

for a constant C given by

$$C = \left(C_0 + L_X\left(\mathbb{E}[\|X_0\|] + \int_0^T \mathbb{E}[M_\xi]\,\mathrm{d}\xi\right)e^{L_X T} + (\mathbb{E}[\bar{F}_T] - \mathbb{E}[\bar{F}_0])\right)e^{L_X T}. \tag{5.5}$$

Proof. Under Standing Hypothesis 2.1, Lemma 4.2 applies and the global error estimate (4.3) holds.

Thanks to (2.4), (2.5), and (4.5), the Proposition 4.1 applies and the global error is bounded according to (4.6).

With assumptions (5.1) and (5.2), Lemma 5.1 applies and the last term in (4.6) is bounded according to (5.3). Using (5.3) in (4.6) yields

$$\begin{aligned}\mathbb{E}\left[\|X_{t_j} - X^N_{t_j}\|\right] \le \Bigg(C_0\Delta t_N + \Delta t_N L_X\left(\mathbb{E}[\|X_0\|] + \int_0^{t_j} \mathbb{E}[M_\xi]\,\mathrm{d}\xi\right)e^{L_X t_j} \\ + (\mathbb{E}[\bar{F}_{t_j}] - \mathbb{E}[\bar{F}_0])\Delta t_N\Bigg)e^{L_X t_j}.\end{aligned}$$

Since this holds for every $j = 0, \dots, N$, we obtain the desired estimate (5.4). □

The conditions of Theorem 5.1, especially (5.1)-(5.2), are not readily verifiable, but the following result gives more explicit conditions.

Theorem 5.2. *Suppose 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 $\partial_y f$ with at most linear growth in x and y, i.e.*

$$\|\partial_t f(t,x,y)\| \leq C_1 + C_2\|x\| + C_3\|y\|, \quad \|\partial_y f(t,x,y)\| \leq C_4 + C_5\|x\| + C_6\|y\|, \tag{5.6}$$

in $(t,x,y) \in I \times \mathbb{R}^d \times \mathbb{R}^k$, for suitable constants $C_1, C_2, C_3, C_4 \geq 0$. Assume, further, that the sample paths of $\{Y_t\}_{t\in I}$ are of bounded variation $V(\{Y_t\}_{t\in I}; I)$, on I, with finite quadratic mean,

$$\mathbb{E}[V(\{Y_t\}_{t\in I}; I)^2] < \infty, \tag{5.7}$$

and such that

$$\mathbb{E}[\|Y_0\|^2] < \infty. \tag{5.8}$$

Moreover, it is assumed that

$$\mathbb{E}[\|X_0\|^2] < \infty. \tag{5.9}$$

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

$$\max_{j=0,\dots,N} \mathbb{E}\left[\left\|X_{t_j} - X_{t_j}^N\right\|\right] \leq C\Delta t_N, \qquad \forall N \in \mathbb{N}, \tag{5.10}$$

for a suitable constant $C \geq 0$.

Proof. Notice that

$$\begin{aligned}\|f(t,x,y)\| &\leq \|f(t,x,y) - f(t,0,y)\| + \|f(t,0,y) - f(0,0,y)\| + \|f(0,0,y) - f(0,0,0)\| \\ &\leq L_X\|x\| + C_1 + C_3\|y\| + C_4 + C_6\|y\|.\end{aligned}$$

Thus,

$$\|f(t,x,Y_t)\| \leq M_t + L_X\|x\|,$$

where

$$M_t = C_1 + C_4 + (C_3 + C_6)\|Y_t\|.$$

Since the sample paths of $\{Y_t\}_{t\in I}$ are of bounded variation, the process $\{M_t\}_{t\in I}$ has integrable sample paths. This means that we are under Standing Hypothesis 2.1. Moreover, we have that

$$\mathbb{E}[\|Y_t\|] \leq \mathbb{E}[\|Y_0\|] + \mathbb{E}[\|Y_t - Y_0\|] \leq \mathbb{E}[\|Y_0\|] + \mathbb{E}[V(\{Y_t\}_{t\in I}; I)].$$

Then, thanks to the Lyapunov inequality $\mathbb{E}[\|Y_t\|] \leq \mathbb{E}[\|Y_t\|^2]^{1/2}$ and the assumptions (5.7) and (5.8), we see that $\{M_t\}_{t\in I}$ satisfies (2.5). By assumption, (2.4) also holds, so that, from (2.3), we have

$$K_X = \sup_{t\in I} \mathbb{E}[\|X_t\|^2] < \infty.$$

Now, in order to apply Theorem 5.1, it remains to verify (5.1)-(5.2).

Since the noise is of bounded variation and $f = f(t,x,y)$ is continuously differentiable in (t,y), we have $s \mapsto f(s, X_\tau, Y_s)$ of bounded variation, for each fixed τ, with

$$f(s,X_\tau,Y_s) - f(\tau,X_\tau,Y_\tau) = \int_\tau^s \partial_t f(\xi,X_\tau,Y_\xi)\,\mathrm{d}\xi + \int_\tau^s \partial_y f(\xi,X_\tau,Y_\xi)\,\mathrm{d}Y_\xi.$$

More precisely, assuming $\{Y_t\}_{t\in I}$ has values in $\mathbb{R}^k$, $k\in\mathbb{N}$, we have each coordinate $t\mapsto (Y_t)_i$ with sample paths of bounded variation, and $\partial_y f = (\partial_{y_1} f, \dots, \partial_{y_k} f)$, so that

$$\int_\tau^s \partial_y f(\xi,X_\tau,Y_\xi)\,\mathrm{d}Y_\xi = \sum_{i=1}^k \int_\tau^s \partial_{y_i} f(\xi,X_\tau,Y_\xi)\,\mathrm{d}(Y_\xi)_i$$

Then, using (5.6),

$$\begin{aligned}\|f(s,X_\tau,Y_s) - f(\tau,X_\tau,Y_\tau)\| \\ \leq C_1(s-\tau) + C_2(s-\tau)\|X_\tau\| + (C_3 + C_4\|X_\tau\|)V(\{Y_t\}_{t\in I};\tau,s).\end{aligned}$$

Thus, (5.1) holds with

$$\bar{F}_t = (C_1 + C_2\|X^N_{\tau^N(t)}\|)t + (C_3 + C_4\|X^N_{\tau^N(t)}\|)V(\{Y_t\}_{t\in I};0,t).$$

It is clear that all the sample paths of $\{\bar{F}_t\}_{t\in I}$ are non-decreasing in $t\in I$, with $\bar{F}_0 = 0$. Moreover, thanks to (5.7), and using the Cauchy-Schwarz inequality in the last term, we have

$$\mathbb{E}[\bar{F}_T] \leq (C_1 + C_2K_1)T + (C_3 + C_4K_1)\mathbb{E}[V(\{Y_t\}_{t\in I};0,T)^2] < \infty.$$

Thus, Theorem 5.1 applies and we deduce the strong order 1 convergence of the Euler method. □

Remark 5.1. The conditions (5.7) and (5.9) on the finite mean square of the total variation of the noise and of the initial condition can be relaxed provided we have a better control on the growth of $\partial_y f(t,x,y)$ with respect to x. More precisely, if

$$\|\partial_y f(t,x,y)\| \leq C_4 + C_5\|x\|^{p-1} + C_6\|y\|,$$

and $\mathbb{E}[V(\{Y_t\}_{t\in I};T,0)^p] < \infty$, along with $\mathbb{E}[\|X_0\|^p] < \infty$, with $1\leq p<\infty$, then the process $\{\bar{F}_t\}_{t\in I}$ becomes

$$\bar{F}_t = (C_1 + C_2\|X^N_{\tau^N(t)}\|)t + (C_3 + C_4|X^N_{\tau^N(t)}|^{p-1})V(\{Y_t\}_{t\in I};0,t).$$

Applying the Hölder inequality yields

$$\bar{F}_t \leq (C_1 + C_2\|X^N_{\tau^N(t)}\|)t + C_3V(\{Y_t\}_{t\in I};0,t) + C_4\frac{p-1}{p}\|X^N_{\tau^N(t)}\|^p + \frac{C_4}{p}V(\{Y_t\}_{t\in I};0,t)^p.$$

With that, the required conditions on $\{\bar{F}_t\}_{t\in I}$ are met and we are allowed to apply Theorem 5.1 and deduce the strong order 1 convergence of the Euler method.

6. The case of an Itô process noise

Here, as explained in the Introduction, we assume the process given by $F_t = f(s, X_{\tau^N(s)}, Y_s)$ is an Itô process, which, in applications, follows from the Itô formula, assuming that $f = f(t, x, y)$ is sufficiently regular and that the noise $\{Y_t\}_{t\in I}$ is itself an Itô process.

We recall (see e.g. [28, Chapter 4]) that an Itô process is a process $\{Z_t\}_{t\in I}$ of the form

$$\mathrm{d}Z_t = A_t\,\mathrm{d}t + B_t\,\mathrm{d}W_t. \tag{6.1}$$

Here Z_t can be either a scalar or a vector-valued process. In the scalar case, $\{W_t\}_{t\geq 0}$ is a (scalar) Wiener process, while $\{A_t\}_{t\in I}$ and $\{B_t\}_{t\in I}$ are real-valued processes adapted to $\{W_t\}_{t\geq 0}$.

In the vector-valued case, $\{A_t\}_{t\in I}$ is a vector-valued process with the same dimension as Z_t, while $\{B_t\}_{t\in I}$ is a square-matrix-valued process with the number of rows and of columns matching the dimension of Z_t. The (multi-dimensional) Wiener process $\{W_t\}_{t\geq 0}$ is a vector-valued process $W_t = (W_t^{(j)})_j$ with the dimension matching the number of columns of B_t and the dimension of Z_t, and where each coordinate $\{W_t^{(j)}\}_{t\in I}$ is a Wiener process independent of the processes in the other coordinates. Both $\{A_t\}_{t\in I}$ and $\{B_t\}_{t\in I}$ are assumed to be adapted to each $\{W_t^{(j)}\}_{t\in I}$. Writing the *columns* of B_t as $B_t^{(j)}$, we can write the stochastic term in (6.1) as $B_t\,\mathrm{d}W_t = \sum_i B_t^{(j)}\,\mathrm{d}W_t^{(j)}$. More explicitly, writing in full coordinates $Z_t = (Z_t^{(i)})_i$, $A_t = (A_t^{(i)})_i$, and $B_t = (B_t^{(i,j)})_{i,j}$, we have (6.1) as the system

$$\mathrm{d}Z_t^{(i)} = A_t^{(i)}\,\mathrm{d}t + \sum_j B_t^{(i,j)}\,\mathrm{d}W_t^{(j)}, \quad \forall i. \tag{6.2}$$

With that in mind, we first have the following result.

Lemma 6.1. *Besides Standing Hypothesis 2.1, suppose that $F_t^N = f(t, X^N_{\tau^N(t)}, Y_t)$ is an Itô process, i.e. satisfying (6.1) with $Z_t = F_t^N$ and suitable processes $\{A_t\}_{t\in I}$ and $\{B_t\}_{t\in I}$. Assume, moreover, that*

$$\int_0^T \mathbb{E}[\|A_t\|]\,\mathrm{d}t < \infty, \quad \int_0^T \mathbb{E}[\|B_t\|^2]\,\mathrm{d}t < \infty. \tag{6.3}$$

Then,

$$\mathbb{E}\left[\left\|\int_0^t \left(f(s, X^N_{\tau^N(s)}, Y_s) - f(\tau^N(s), X^N_{\tau^N(s)}, Y_{\tau^N(s)})\right)\,\mathrm{d}s\right\|\right] \leq \Delta t_N \left(\int_0^t \mathbb{E}[\|A_\xi\|]\,\mathrm{d}\xi + \left(\int_0^t \mathbb{E}[\|B_\xi\|^2]\,\mathrm{d}\xi\right)^{1/2}\right), \tag{6.4}$$

for all $0 \leq t \leq T$ and every $N \in \mathbb{R}$.

Proof. We write

$$f(s, X^N_{\tau^N(s)}, Y_s) - f(\tau^N(s), X^N_{\tau^N(s)}, Y_{\tau^N(s)}) = \int_{\tau^N(s)}^{s} A_\xi \, \mathrm{d}\xi + \int_{\tau^N(s)}^{s} B_\xi \, \mathrm{d}W_\xi.$$

Upon integration,

$$\int_0^t \left(f(s, X^N_{\tau^N(s)}, Y_s) - f(\tau^N(s), X^N_{\tau^N(s)}, Y_{\tau^N(s)}) \right) \mathrm{d}s \\ = \int_0^t \left(\int_{\tau^N(s)}^{s} A_\xi \, \mathrm{d}\xi + \int_{\tau^N(s)}^{s} B_\xi \, \mathrm{d}W_\xi \right) \mathrm{d}s.$$

Exchanging the order of integration, according to Fubini's theorem (see e.g. [29, Section IV.6] for a suitable stochastic version of the Fubini Theorem), yields

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

For exchanging the order of integration in the stochastic integral, an important aspect is that, although the integral in s becomes an integral on the interval $[\xi, \tau^N(\xi)+\Delta t_N]$, which is posterior to ξ, the integrand does not depend on s and, hence, does not violate any non-anticipative condition for the validity of the Itô integral and of the stochastic Fubini Theorem.

Taking the absolute mean and using the Itô isometry [28] on the second term gives

$$\begin{aligned} &\mathbb{E}\left[\left\| \int_0^t \left(f(s, X^N_{\tau^N(s)}, Y_s) - f(\tau^N(s), X^N_{\tau^N(s)}, Y_{\tau^N(s)}) \right) \mathrm{d}s \right\| \right] \\ &\le \int_0^t |\tau^N(\xi) + \Delta t_N - \xi| \mathbb{E}[\|A_\xi\|] \, \mathrm{d}\xi + \left(\int_0^t (\tau^N(\xi) + \Delta t_N - \xi)^2 \mathbb{E}[\|B_\xi\|^2] \, \mathrm{d}\xi \right)^{1/2}. \end{aligned}$$

Since $|\tau^N(\xi) + \Delta t_N - \xi| \le \Delta t_N$, we find

$$\mathbb{E}\left[\left\| \int_0^t \left(f(s, X^N_{\tau^N(s)}, Y_s) - f(\tau^N(s), X^N_{\tau^N(s)}, Y_{\tau^N(s)}) \right) \mathrm{d}s \right\| \right] \\ \le \Delta t_N \left(\int_0^t \mathbb{E}[\|A_\xi\|] \, \mathrm{d}\xi + \left(\int_0^t \mathbb{E}[\|B_\xi\|^2] \, \mathrm{d}\xi \right)^{1/2} \right),$$

which completes the proof. □

Combining the estimate in Lemma 6.1 with the estimate for the global error we obtain the following main result.

Theorem 6.1. *Under Standing Hypothesis 2.1, suppose also that (2.4), (2.5), (4.5), (6.1), and (6.3) hold. Then, the Euler scheme (1.2) is of strong order 1, i.e.*

$$\max_{j=0,\dots,N} \mathbb{E}\left[\left\|X_{t_j} - X_{t_j}^N\right\|\right] \leq C\Delta t_N, \qquad \forall N \in \mathbb{N}, \tag{6.5}$$

for a constant C given by

$$C = \left(C_0 + L_X\left(\mathbb{E}[\|X_0\|] + \int_0^T \mathbb{E}[M_\xi]\,\mathrm{d}\xi\right)e^{L_X T} + \left(\int_0^T \mathbb{E}[\|A_\xi\|]\,\mathrm{d}\xi + \left(\int_0^T \mathbb{E}[\|B_\xi\|^2]\,\mathrm{d}\xi\right)^{1/2}\right)\right)e^{L_X T}. \tag{6.6}$$

Proof. Under Standing Hypothesis 2.1, the Lemma 4.2 applies and the global error estimate (4.3) holds. Thanks to (2.4), (2.5), and (4.5), the Proposition 4.1 applies and the global error is bounded according to (4.6).

With assumptions (6.1) and (6.3), the Lemma 6.1 applies and the last term in (4.6) is bounded according to (6.4). Using (6.4) in (4.6) yields

$$\mathbb{E}\left[\left\|X_{t_j} - X_{t_j}^N\right\|\right] \leq \left(C_0\Delta t_N + \Delta t_N L_X\left(\mathbb{E}[\|X_0\|] + \int_0^{t_j} \mathbb{E}[M_\xi]\,\mathrm{d}\xi\right)e^{L_X t_j} + \Delta t_N\left(\int_0^{t_j} \mathbb{E}[\|A_\xi\|]\,\mathrm{d}\xi + \left(\int_0^{t_j} \mathbb{E}[\|B_\xi\|^2]\,\mathrm{d}\xi\right)^{1/2}\right)\right)e^{L_X t_j}.$$

Since this holds for every $j = 0, \dots, N$, we obtain the desired estimate (6.5). □

In practice, conditions (6.1)-(6.3) follow from assuming sufficent regularity on $f = f(t, x, y)$ and that the noise is an Itô process. We assume the noise satisfies

$$\mathrm{d}Y_t = a(t, Y_t)\,\mathrm{d}t + b(t, Y_t)\,\mathrm{d}W_t. \tag{6.7}$$

In the scalar case ($k = 1$), both $a = a(t, y)$ and $b = b(t, y)$ are scalars, while in the multi-dimensional case ($k > 1$), the coefficient $a = a(t, y) = (a_i(t, y))_i$ is vector valued with the same dimension k as $y = (y_i)_i$, and $b = b(t, y) = (b_{ij}(t, y))_{ij}$ is matrix valued, with dimension $k \times k$.

Then, we have following result.

Theorem 6.2. *Let $f = f(t, x, y)$ be twice continuously differentiable with uniformly bounded derivatives. Suppose that the noise $\{Y_t\}_{t\in I}$ is an Itô process, satisfying (6.7), with $a = a(t, y)$ and $b = b(t, y)$ continuous and such that*

$$\|a(t, y)\| \leq A_M + A_Y\|y\|, \qquad \|b(t, y)\| \leq B_M + B_Y\|y\|, \tag{6.8}$$

for all y, with constants $A_M, A_Y, B_M, B_Y \geq 0$. Assume the bounds (2.4) and (4.5) hold, and that

$$\mathbb{E}[\|Y_0\|] < \infty. \tag{6.9}$$

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

$$\max_{j=0,\dots,N} \mathbb{E}\left[\left\|X_{t_j} - X_{t_j}^N\right\|\right] \leq C\Delta t_N, \qquad \forall N \in \mathbb{N}, \tag{6.10}$$

for a suitable constant $C \geq 0$.

Proof. Let us start by showing that the Standing Hypothesis 2.1 is valid. Since $f = f(t,x,y)$ is (twice) continuously differentiable with, in particular, bounded derivative in x, then it is uniformly globally Lipschitz in x. Since $a = a(t,y)$ and $b = b(t,y)$ are continuous, the noise has continuous sample paths. Thus, the remaining condition in Standing Hypothesis 2.1 to be verified is (iic).

Using the Itô formula with (6.7) in mind (see [28, Theorem 4.1.2] or [24, Section 7.4] for the scalar case of the Itô formula, with $g(t,y) = y^2$, and [28, Theorem 4.2.1] or, more generally, [24, Section 7.5] for the multi-dimensional case, with $g(t,y) = \|y\|^2$), we have

$$\mathrm{d}\|Y_t\|^2 = \left(2a(t,Y_t)\cdot Y_t + b(t,Y_t)^2\right)\,\mathrm{d}t + 2b(t,Y_t)Y_t\,\mathrm{d}W_t.$$

Thus,

$$\|Y_t\|^2 = \|Y_0^2\| + \int_0^t \left(2a(s,Y_s)\cdot Y_s + b(t,Y_s)^2\right)\,\mathrm{d}s + \int_0^t 2b(s,Y_s)Y_s\,\mathrm{d}W_s.$$

Taking the expectation,

$$\mathbb{E}[\|Y_t\|^2] = \mathbb{E}[\|Y_0\|^2] + \int_0^t \mathbb{E}\left[\left(2a(s,Y_s)\cdot Y_s + b(t,Y_s)^2\right)\right]\,\mathrm{d}s.$$

Using (6.8), this yields

$$\begin{aligned}\mathbb{E}[\|Y_t\|^2] &\leq \mathbb{E}[\|Y_0\|^2] + \int_0^t \left(2\mathbb{E}[(A_M + A_Y\|Y_s\|)\|Y_s\|] + \mathbb{E}[(B_M + B_Y\|Y_s\|)^2]\right)\,\mathrm{d}s\\ &\leq \mathbb{E}[\|Y_0\|^2] + \int_0^t \left((A_M^2 + (1+2A_Y)\mathbb{E}[\|Y_s\|^2]) + 2(B_M^2 + B_Y^2\mathbb{E}[\|Y_s\|^2])\right)\,\mathrm{d}s.\end{aligned}$$

By the classical Gronwall Lemma [15],

$$\mathbb{E}[\|Y_t\|^2] \leq \left(\mathbb{E}[\|Y_0\|^2] + (A_M^2 + 2B_M^2)t\right)e^{(1+2A_Y+2B_Y^2)t}.$$

Thus,

$$\sup_{t\in I}\mathbb{E}[\|Y_t\|^2] \leq \left(\mathbb{E}[\|Y_0\|^2] + (A_M^2 + 2B_M^2)T\right)e^{(1+2A_Y+2B_Y^2)T}. \tag{6.11}$$

Since $f = f(t,x,y)$ is Lipschitz in x and twice continuously differentiable in (t,y) with uniformly bounded first order derivatives, we have the bound

$$\|f(t,x,y)\| \leq \|f(0,0,0)\| + L_X\|x\| + L_T t + L_Y\|y\|.$$

Thus,

$$\|f(t,x,Y_t)\| \leq M_t + L_X\|x\|,$$

with

$$M_t = \|f(0,0,0)\| + L_T t + L_Y\|y\|.$$

Thanks to (6.11), we see that

$$\int_0^T M_t \, \mathrm{d}t < \infty.$$

Therefore, we are under the conditions of Standing Hypothesis 2.1.

Now, in view of Theorem 6.1, it remains to prove that $F_t^N = f(t, X^N_{\tau^N(t)}, Y_t)$ is an Itô process (6.1), with the bounds (6.3). The fact that it is an Itô process follows from the Itô formula and the smoothness of $f = f(t,x,y)$, as we shall see now. We first consider the case of a scalar noise, for the sake of clarity. Since $(t,y) \mapsto f(t,x,y)$ is twice continuously differentiable, for each fixed x, the Itô formula is applicable and yields, in the case of a scalar noise,

$$\begin{aligned}\mathrm{d}f(t,x,Y_t) = \left(\partial_t f(t,x,Y_t) + a(t,Y_t)\partial_y f(t,x,Y_t) + \frac{b(t,Y_t)^2}{2}\partial_{yy} f(t,x,Y_t)\right) \mathrm{d}t \\ + b(t,Y_t)\partial_y f(t,x,Y_t)\, \mathrm{d}W_t, \quad (6.12)\end{aligned}$$

for every fixed $x \in \mathbb{R}$. This means (6.1) holds with

$$A_t = \partial_t f(t,x,Y_t) + a(t,Y_t)\partial_y f(t,x,Y_t) + \frac{b(t,Y_t)^2}{2}\partial_{yy} f(t,x,Y_t)$$

and

$$B_t = b(t,Y_t)\partial_y f(t,x,Y_t).$$

For the integrability conditions, since $f = f(t,x,y)$ has uniformly bounded derivatives, we have

$$\|A_t\| \le L_T + L_Y(A_M + A_Y\|Y_t\|) + 2L_{YY}(B_M^2 + B_Y^2\|Y_t\|^2),$$

and

$$\|B_t\| \le L_Y(B_M + B_Y\|Y_t\|),$$

for a suitable constant $L_{YY} \ge 0$. Now, thanks to (6.11), we see that (6.3) is satisfied.

In the case of a multi-dimensional noise, we have $a(t,y) = (a_i(t,y))_i$ and $b(t,y) = (b_{ij}(t,y))_{ij}$, and (6.12) becomes

$$\begin{aligned}\mathrm{d}f(t,x,Y_t) = \Bigg(\partial_t f(t,x,Y_t) + \sum_i \partial_{y_i} f(t,x,Y_t) a_i(t,Y_t) \\ + \frac{1}{2}\sum_i \partial^2_{y_i} f(t,x,Y_t) b_{ii}(t,Y_t)^2\Bigg)\, \mathrm{d}t \\ + \sum_{i,j} \partial_{y_i} f(t,x,Y_t) b_{ij}(t,Y_t)\, \mathrm{d}W_t^{(j)}, \quad (6.13)\end{aligned}$$

so that (6.1)/(6.2) and (6.3) follow similarly.

Therefore, all the conditions of Theorem 6.1 are met and we obtain the strong order 1 convergence of the Euler method. □

Remark 6.1. When the diffusion coefficient $b = b(t, y) = b(t)$ in (6.7) is independent of y, the noise is an additive noise, and in this case the Euler scheme is well known to be of strong order 1 [18, Section 19.3.1]. In the more general case $b = b(t, y)$, however, the Euler scheme has always been regarded to be of order 1/2 [16] (see also [36] for mean square convergence). Here, though, we deduce, under the conditions of Theorem 6.2, that even if $b = b(t, y)$ depends on y, the strong convergence of the Euler scheme is actually of order 1.

7. The mixed case with Itô and bounded variation noises

Of course, it is possible to mix the two cases and have the following result combining Theorem 5.1 and Theorem 6.1.

Theorem 7.1. *Under Standing Hypothesis 2.1, suppose also that (2.4), (2.5), and (4.5) hold. Suppose, moreover, that $F_t^N = f(t, X^N_{\tau^N(t)}, Y_t)$ can be split into a sum $F_t^N = G_t^N + H_t^N$ where $\{G_t^N\}_{t\in I}$ satisfies (6.1) and (6.3) and where the steps of $\{H_t^N\}_{t\in I}$ are bounded by a real stochastic process $\{\bar{H}_t\}$ with monotonic non-decreasing sample paths, i.e.*

$$\|H_s^N - H^N_{\tau^N(s)}\| \le \bar{H}_s^N - \bar{H}^N_{\tau^N(s)} \tag{7.1}$$

with

$$\mathbb{E}[\bar{H}_t] \text{ uniformly bounded on } t \in I. \tag{7.2}$$

Then, the Euler scheme (1.2) is of strong order 1, i.e. there exists a constant $C \ge 0$ such that

$$\max_{j=0,\dots,N} \mathbb{E}\left[\left\|X_{t_j} - X^N_{t_j}\right\|\right] \le C\Delta t_N, \qquad \forall N \in \mathbb{N}. \tag{7.3}$$

We omit the proof since it is just a combination of Lemma 5.1 and Lemma 6.1. As a consequence, we also have the following more explicit result, which is a combination of Theorem 5.2 and Theorem 6.2.

Theorem 7.2. *Suppose that $f = f(t, x, y)$ is twice continuously differentiable with uniformly bounded derivatives. Assume, further, that the sample paths of $\{Y_t\}_{t\in I}$ are made of two independent components, one of bounded variation with finite quadratic mean, as in (5.7), and another an Itô process noise satisfying (6.7) and (6.9). Assume, moreover, that (5.9) holds. Then, the Euler scheme is of strong order 1, i.e.*

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

for a suitable constant $C \ge 0$.

8. Numerical examples

In this section, we illustrate the strong order 1 convergence with a few 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 [31].

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:

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

Then we fit the errors ϵ^N to the power law $C\Delta_N^p$, in order to find p, along with the 95% confidence interval.

Here are the main parameters for the error estimate:

(i) $M \in \mathbb{N}$ is the number of samples for the Monte Carlo estimate of the strong error, typically in the range $M = 100$ to $M = 1,000$.
(ii) The time interval $[0, T]$ for the initial-value problem, typically with $T = 1.0$.
(iii) The initial condition X_0, which is typically $X_0 \sim \mathcal{N}(0, 1)$.
(iv) A series of time steps $\Delta_i = T/(N_i - 1)$, usually with $N_i = 2^{n_i}$, most often with $n_i = 4, \dots, 10$, hence $N_i = 32, 64, \dots, 1024$.
(v) A number N_{tgt} of mesh points for a finer discretization to compute a target solution path, typically $N_{\text{tgt}} = \max_i\{N_i^2\}$, e.g. $N = 2^{20} = 1,048,576$, 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 $m = 1, \dots, M$, we first generate a discretization $\{Y_{t_j}\}_{j=0,N_{\text{tgt}}}$ of a sample path of the noise on the finest grid $\{t_j^{N_{\text{tgt}}}\}$, with N_{tgt} points, using either an exact distribution for the noise or an approximation in a much finer mesh.
(ii) Next, we use the values of the noise at the target time mesh to generate the target solution $\{X_{t_j}\}_{j=0,N_{\text{tgt}}}$, 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 step N_i in the selected range, we compute the Euler approximation using the computed noise values at the corresponding coarser mesh.
(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 $(\Delta_i, \epsilon^{N_i})$ to the power law $C\Delta_i^p$, via linear least-square regression in log scale, for suitable C and p, with p giving the order of convergence, which 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 mesh steps $(\Delta_{N_i})_i$, 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,

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

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 de 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 [3]. 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 [30], 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 checking the accuracy of the implementation, for publication purposes, becomes easier. All the source code for the numerical simulations presented below are in a Github repository [31].

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

$$\begin{cases} \dfrac{\mathrm{d}X_t}{\mathrm{d}t} = W_t X_t, & 0 \le t \le T, \\ X_t|_{t=0} = X_0. \end{cases} \tag{8.2}$$

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

Equation (8.2) has the explicit solution

$$X_t = e^{\int_0^t W_s \,\mathrm{d}s} X_0. \tag{8.3}$$

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 [17, Section 14.2] as

$$\int_0^\tau W_s \,\mathrm{d}s \sim \frac{\tau}{2}\Delta W + \sqrt{\frac{\tau^3}{12}}\mathcal{N}(0,1) = \frac{\tau}{2}\Delta W + \mathcal{N}\left(0, \frac{\tau^3}{12}\right). \tag{8.4}$$

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 (8.4), that

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

where

$$Z_i \sim \mathcal{N}\left(0, \frac{(t_{i+1}-t_i)^3}{12}\right) = \frac{\sqrt{(t_{i+1}-t_i)^3}}{\sqrt{12}}\mathcal{N}(0,1). \tag{8.5}$$

Thus, breaking down the sum over the mesh intervals:

$$\int_0^{t_j} W_s\,\mathrm{d}s = \sum_{i=0}^{j-1}\int_{t_i}^{t_{i+1}} W_s\,\mathrm{d}s = \sum_{i=0}^{j-1}\left(\frac{(W_{t_{i+1}} + W_{t_i})}{2}(t_{i+1}-t_i) + Z_i\right), \tag{8.6}$$

where $Z_0, \ldots, Z_{N-1}$ are independent random variables with distribution (8.5).

Notice the first term in (8.6) is the trapezoidal rule, with the second one being a higher-order correction term.

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

$$\mathbb{E}[e^{Z_i}] = e^{((t_{i+1}-t_i)^3)/24}. \tag{8.7}$$

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 (8.2) 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

$$X_{t_j} = X_0 e^{\sum_{i=0}^{j-1}\left(\frac{1}{2}(W_{t_i}+W_{t_{i+1}})(t_{i+1}-t_i)+Z_i\right)}, \tag{8.8}$$

for realizations Z_i drawn from a normal distributions given by (8.5). 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 a thousand sample paths for each chosen time step, with initial condition $X_0 \sim \mathcal{N}(0,1)$, on the interval $[0,T]$. Figure 2 illustrates the order of convergence.

8.2. **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.

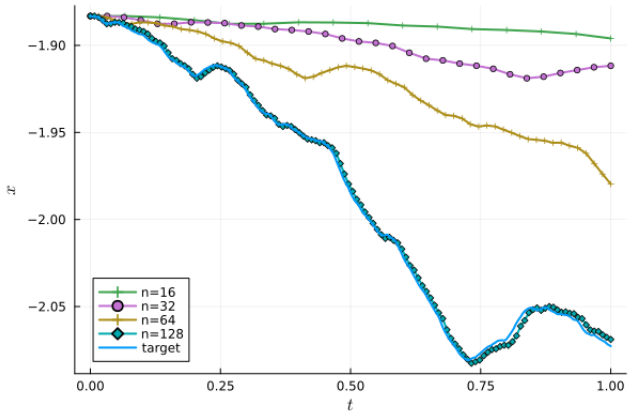

FIGURE 1. Euler approximation of $\mathrm{d}X_t/\mathrm{d}t = W_t X_t$ with $X_0 = 1.0$, on $[0, T]$, and a few sample paths of exact solutions compatible with the given realizations of the noise on the mesh points.

N	dt	error
16	0.0667	0.0194
32	0.0323	0.0108
64	0.0159	0.00546
128	0.00787	0.00278
256	0.00392	0.00139
512	0.00196	0.000687
1024	0.000978	0.000345
2048	0.000489	0.000172
4096	0.000244	8.62e-5
8192	0.000122	4.66e-5
16384	6.1e-5	2.45e-5

TABLE 1. Mesh points (N), time steps (dt), and strong error (error) of the Euler method for $\mathrm{d}X_t/\mathrm{d}t = W_t X_t$, 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 $M = 1000$ sample paths for each fixed time step, with the target solution calculated with 65536 points.

The system of equations takes the form

$$\begin{cases} \dfrac{\mathrm{d}\mathbf{X}_t}{\mathrm{d}t} = -\left\|\mathbf{Y}_t\right\|^2 \mathbf{X}_t + \mathbf{Y}_t, \qquad 0 \le t \le T, \\ \mathbf{X}_t|_{t=0} = \mathbf{X}_0, \end{cases} \tag{8.9}$$

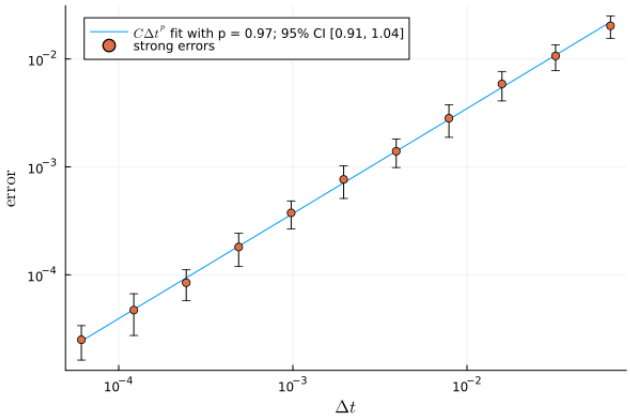

FIGURE 2. Order of convergence $p = 0.98$ of the strong error of the Euler method for $\mathrm{d}X_t/\mathrm{d}t = W_t X_t$, on $[0, T] = [0.0, 1.0]$, with $X_0 \sim \mathcal{N}(0, 1)$, and a standard Wiener process noise $\{W_t\}_t$, computed with $M = 1000$ sample paths, for each $\Delta t = 1/N$, $N = 16$, 32, 64, 128, 256, 512, 1024.

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 $y_0 = 0.2$;
(iv) A non-autonomous homogeneous linear Itô process (hlp) $\{H_t\}_t$ given by $\mathrm{d}H_t = (\mu_1 + \mu_2 \sin(\omega t))H_t \ \mathrm{d}t + \sigma \sin(\omega t)H_t \ \mathrm{d}W_t$ with $\mu_1 = 0.5$, $\mu_2 = 0.3$, $\sigma = 0.5$, $\omega = 3\pi$, 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 $\theta = 2.0$;

N	dt	error
64	0.0159	0.22
128	0.00787	0.106
256	0.00392	0.0525
512	0.00196	0.0262

TABLE 2. Mesh points (N), time steps (dt), and strong error (error) of the Euler method for $\mathrm{d}\mathbf{X}_t/\mathrm{d}t = -\|\mathbf{Y}_t\|^2\mathbf{X}_t + \mathbf{Y}_t$, with initial condition $\mathbf{X}_0 \sim \mathcal{N}(\mathbf{0}, \mathrm{I})$ and vector-valued noise $\{Y_t\}_t$ with all the implemented noises, on the time interval (0.0, 1.0), based on $M = 200$ sample paths for each fixed time step, with the target solution calculated with 262144 points.

(ix) A fractional Brownian motion (fBm) process with Hurst parameter $H = 0.6$ and initial condition $y_0 = 0.2$.

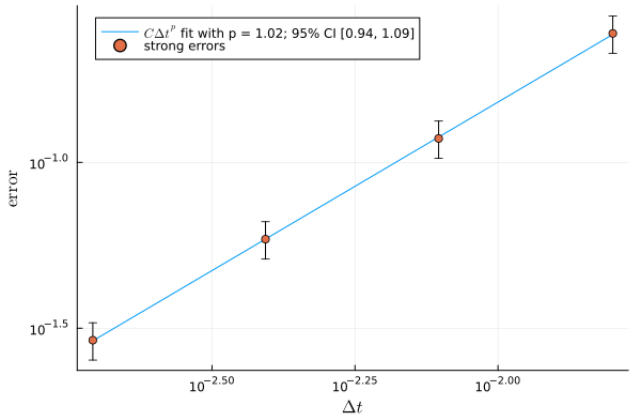

FIGURE 3. Order of convergence $p = 1.02$ of the strong error of the Euler method for $\mathrm{d}\mathbf{X}_t/\mathrm{d}t = -\|\mathbf{Y}_t\|^2\mathbf{X}_t + \mathbf{Y}_t$, on $[0, T] = [0.0, 1.0]$, with $\mathbf{X}_0 \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$, and a vector-valued noise with different types of noise processes, computed with $M = 200$ sample paths, for each $\Delta t = 1/N$, with $N = 64$, 128, 256, and 512.

Table 2 shows the estimated strong error obtained from 200 sample paths for each chosen time step, with initial condition $\mathbf{X}_0 \sim \mathcal{N}(\mathbf{0}, \mathrm{I})$, i.e. normally distributed on each coordinate, independently of the other coordinates, and on the time interval $[0.0, 1.0]$. Figure 3 illustrates the order of convergence.

Finally, Figure 4 illustrates 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

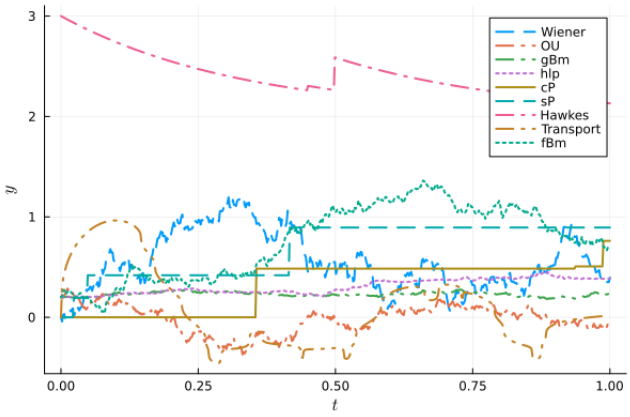

FIGURE 4. Sample paths of all the noises used in the linear system (8.9), mixing all different types of implemented noises.

SDE with an additive noise. In this case, the Euler-Maruyama approximation for the noise part of the SDE is distributionally exact, and the Euler method for the RODE becomes equivalent to the Euler-Maruyama method for the SDE, and it is known that the Euler-Maruyama method for an SDE with additive noise is of strong order 1 [18].

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.

As for the geometric Brownian motion process, it is worth saying that its state at a given time t depends solely on t and on the state of an associated Wiener process at time t, so that the RODE can be transformed into another RODE with an additive noise. However, the corresponding nonlinear term does not have a global Lipschitz bound, so the strong order 1 does not follow from that. Our results, however, apply without further assumptions.

Finally, the homogeneous linear Itô process is a multiplicative noise whose state at time t cannot be written explicitly as a function of t and W_t. It requires the previous history W_s of the associated Wiener process, for $0 \leq s \leq t$, hence the associated RODE cannot be transformed into a RODE with additive noise.

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

$$\begin{cases} \dfrac{\mathrm{d}X_t}{\mathrm{d}t} = -X_t + B_t^H, & 0 \leq t \leq T, \\ X_t|_{t=0} = X_0, \end{cases} \tag{8.10}$$

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 [4, Theorem 4.2.6], [2, Theorem 4.1]and [26, Theorem 2.7.4]), based on the Wick Itô Skorohod (WIS) integral (see [4, Chapter 4]). A corresponding WIS isometry is also needed (see e.g. [4, 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 (4.6) of Proposition 4.1, involving the steps of the term $f(t, x, y) = -x + y$, which in this case reduce to

$$f(s, X^N_{\tau^N(s)}, Y_s) - f(\tau^N(s), X^N_{\tau^N(s)}, Y_{\tau^N(s)}) = B^H_s - B^H_{\tau^N(s)}, \tag{8.11}$$

for $0 \le s \le T$. There are several ways to represent an fBm (see e.g. [4, 26]). We will use the formula [25, eq. (2.1)], [4, eq. (1.1)]

$$B^H_t = \frac{1}{\Gamma(H+1/2)} \left(\int_{-\infty}^0 \left((t-s)^{H-1/2} - (-s)^{H-1/2} \right) \, \mathrm{d}W_s \right. \\ \left. + \int_0^t (t-s)^{H-1/2} \, \mathrm{d}W_s \right), \tag{8.12}$$

where $\Gamma(\cdot)$ is the well-known Gamma function. For the step, (8.12) means that

$$B^H_s - B^H_{\tau^N(s)} = \frac{1}{\Gamma(H+1/2)} \left(\int_{-\infty}^{\tau^N(s)} \left((s-\xi)^{H-1/2} - (\tau^N(s)-\xi)^{H-1/2} \right) \, \mathrm{d}W_\xi \right. \\ \left. + \int_{\tau^N(s)}^s (s-\xi)^{H-1/2} \, \mathrm{d}W_\xi \right). \tag{8.13}$$

Then, using Fubini's Theorem to exchange the order of integration (see, again, [29, Section IV.6]),

$$\begin{aligned}
&\int_0^{t_j} \left(f(s, X^N_{\tau^N(s)}, Y_s) - f(\tau^N(s), X^N_{\tau^N(s)}, Y_{\tau^N(s)}) \right) \, \mathrm{d}s \\
&\quad = \frac{1}{\Gamma(H+1/2)} \int_0^{t_j} \int_{-\infty}^{\tau^N(s)} \left((s-\xi)^{H-1/2} - (\tau^N(s)-\xi)^{H-1/2} \right) \, \mathrm{d}W_\xi \, \mathrm{d}s \\
&\qquad + \frac{1}{\Gamma(H+1/2)} \int_0^{t_j} \int_{\tau^N(s)}^{s} (s-\xi)^{H-1/2} \, \mathrm{d}W_\xi \, \mathrm{d}s \\
&\quad = \frac{1}{\Gamma(H+1/2)} \int_{-\infty}^{0} \int_0^{t_j} \left((s-\xi)^{H-1/2} - (\tau^N(s)-\xi)^{H-1/2} \right) \, \mathrm{d}s \, \mathrm{d}W_\xi \\
&\qquad + \frac{1}{\Gamma(H+1/2)} \int_0^{t_j} \int_{\tau^N(\xi)+\Delta t_N}^{t_j} \left((s-\xi)^{H-1/2} - (\tau^N(s)-\xi)^{H-1/2} \right) \, \mathrm{d}s \, \mathrm{d}W_\xi \\
&\qquad + \frac{1}{\Gamma(H+1/2)} \int_0^{t_j} \int_{\xi}^{\tau^N(\xi)+\Delta t_N} (s-\xi)^{H-1/2} \, \mathrm{d}s \, \mathrm{d}W_\xi
\end{aligned} \tag{8.14}$$

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

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

Hence,

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

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

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

Therefore,

$$\frac{1}{\Gamma(H+1/2)}\Delta t_N\mathbb{E}\left[\left|\int_{-\infty}^{0}\int_{0}^{t_j}\left((s-\xi)^{H-1/2}-(\tau^N(s)-\xi)^{H-1/2}\right)\,\mathrm{d}s\,\mathrm{d}W_\xi\right|\right]\le C_H^{(1)}\Delta t_N,\quad(8.15)$$

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

The second term is similar,

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

Thus,

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

Hence,

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

Therefore,

$$\frac{1}{\Gamma(H+1/2)}\mathbb{E}\left[\left|\int_0^{t_j}\int_{\tau^N(\xi)+\Delta t_N}^{t_j}\left((s-\xi)^{H-1/2}-(\tau^N(s)-\xi)^{H-1/2}\right)\,\mathrm{d}s\,\mathrm{d}W_\xi\right|\right]$$
$$\leq C_H^{(2)}\Delta t_N,\quad(8.16)$$

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

For the last term, we have

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

so that, using the Lyapunov inequality and the Itô isometry

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

Therefore,

$$\frac{1}{\Gamma(H+1/2)}\mathbb{E}\left[\left|\int_0^{t_j}\int_\xi^{\tau^N(\xi)+\Delta t_N}(s-\xi)^{H-1/2}\,\mathrm{d}s\,\mathrm{d}W_\xi\right|\right]\leq C_H^{(3)}\Delta t_N^{H+1/2},\quad(8.17)$$

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

Putting the three estimates (8.15), (8.16), (8.17) in (8.14) we find that

$$\mathbb{E}\left[\left|\int_0^{t_j}\left(f(s,X^N_{\tau^N(s)},Y_s)-f(\tau^N(s),X^N_{\tau^N(s)},Y_{\tau^N(s)})\right)\,\mathrm{d}s\right|\right] \leq C_H^{(4)}\Delta t_N + C_H^{(3)}\Delta t_N^{H+1/2}, \quad (8.18)$$

where $C_H^{(4)} = C_H^{(1)} + C_H^{(2)}$. Using this estimate in Proposition 4.1 shows that the Euler method is of order $H+1/2$, when $0 < H < 1/2$, and is of order 1, when $1/2 \leq H < 1$, having in mind that $H = 1/2$ corresponds to the classical Wiener process.

In summary, we have proved the following result.

Theorem 8.1. *Consider the equation (8.10) where $\{B_t^H\}_t$ is a fractional Brownian motion (fBm) with Hurst parameter $0 < H < 1$. Suppose the initial condition X_0 satisfies*

$$\mathbb{E}[|X_0|^2] < \infty. \quad (8.19)$$

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,

$$\max_{j=0,\ldots,N}\mathbb{E}\left[\left|X_{t_j} - X^N_{t_j}\right|\right] \leq C_1\Delta t_N + C_2\Delta t_N^{H+1/2}, \qquad \forall N \in \mathbb{N}, \quad (8.20)$$

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

$$X_t = e^{-t}X_0 + \int_0^t e^{-(t-s)}B_s^H\,\mathrm{d}s, \quad (8.21)$$

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 (8.10) with several values of H. We fix the time interval as $[0, T] = [0.0, 1.0]$, set the resolution for the target approximation to $N_{\mathrm{tgt}} = 2^{19}$, choose the time steps for the convergence test as $\Delta t = 1/N$, $N = 64$, 128, 256, and 512, 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 [9] (see also [17, Section 14.4]). Table 3 shows the obtained convergence estimates, for a number of Hurst parameters, which is illustrated in Figure 5, matching reasonably well the theoretical estimate of $p = \min\{H + 1/2, 1\}$.

8.4. **Population dynamics with harvest.** Here, we consider a population dynamics problem modeled by a logistic equation with random coefficients, loosely inspired by [17, Section 15.2], with an extra term representing harvest:

$$\frac{\mathrm{d}X_t}{\mathrm{d}t} = G_tX_t(r - X_t) - H_t, \quad (8.22)$$

H	p
0.1	0.630713
0.2	0.759896
0.3	0.855504
0.4	0.942058
0.5	1.0012
0.7	1.00544
0.9	0.99782

TABLE 3. Hurst parameter H and order p of strong convergence for a number of Hurst values.

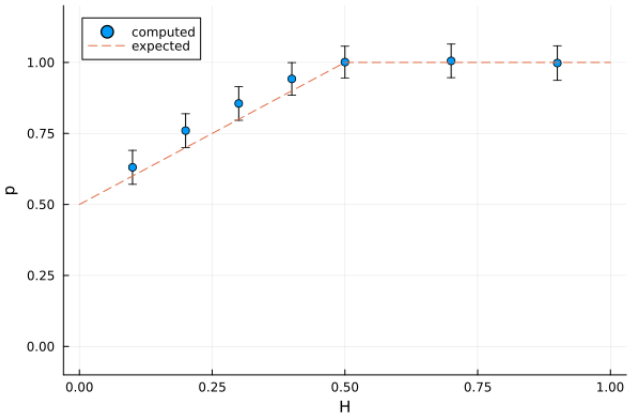

FIGURE 5. Order p of strong convergence for each value of the Hurst parameter H (scattered plot) along with the theoretical value $p = \min\{H + 1/2, 1\}$ (dashed line).

where $r > 0$ is constant; $\{G_t\}_{t\geq 0}$ is a stochastic process playing the role of a random growth parameter; and $\{H_t\}_{t\geq 0}$ is a nonnegative process playing the role of the harvest term. More specifically, $\{G_t\}_{t\geq 0}$ is given by

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

where $\gamma > 0$, $0 < \varepsilon < 1$, and $\{Z_t\}_{t\geq 0}$ is a geometric Brownian motion (gBm) process, hence of the form (6.7)-(6.8). A Wiener process is a natural choice, but we choose a gBm process instead since it is a multiplicative noise, thus the convergence would not be expected to be of strong order 1 based on previous works.

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, independent also of the Poisson process $\{N_t\}_{t\geq 0}$.

We suppose the initial condition is nonnegative and bounded almost surely:

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

for some $R > r$.

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

$$0 < \lambda - \varepsilon \leq G_t \leq \lambda + \varepsilon < 2\lambda, \qquad \forall t \geq 0.$$

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

$$f(t, x, y) = \begin{cases} \gamma(1 + \varepsilon \sin(y_1))x(r - x) - y_2, & x > 0, \\ 0, & x \leq 0. \end{cases}$$

The equation (8.22) 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 $Y_t = (Z_t, H_t)$.

Notice that $f(t, x, y) = 0$, for $x < 0$ and 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 interval. Hence, within the region of interest, all the conditions of Theorem 7.2 hold and the Euler method is of strong order 1.

Below, we simulate numerically the solutions of the above problem, with $\gamma = 1.0$, $\varepsilon = 0.3$, $r = 1.0$, and $\alpha = \gamma/2 = 0.5$. The geometric Brownian motion process $\{Z_t\}_{t\geq 0}$ 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 $M = 1000$ 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 $N_{\mathrm{tgt}} = 2^{18}$ points, while for the approximations we take $N = 2^i$, for $i = 4, \ldots, 9$.

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.

N	dt	error
16	0.0667	0.0118
32	0.0323	0.00553
64	0.0159	0.00269
128	0.00787	0.00133
256	0.00392	0.000657
512	0.00196	0.000325

TABLE 4. Mesh points (N), time steps (dt), and strong error (error) of the Euler method for population dynamics, with initial condition $X_0 \sim \text{Beta}(7.0, 5.0)$ and gBm and step process noises, on the time interval (0.0, 1.0), based on $m = 1000$ sample paths for each fixed time step, with the target solution calculated with 262144 points.

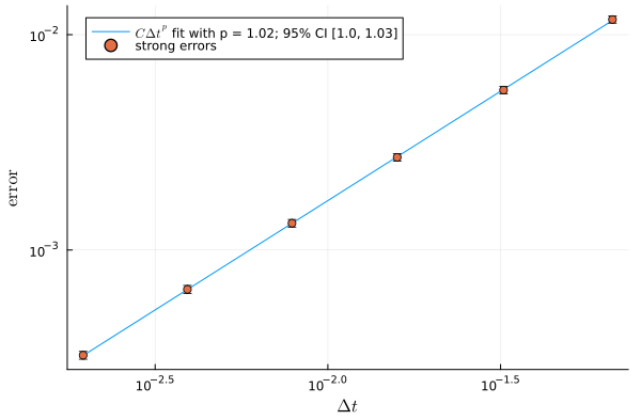

FIGURE 6. Order of convergence of the strong error of the Euler method for equation (8.22).

8.5. **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 [5] (see also [27, Chapter 18] and [19] 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 spectral density, determined by the mechanical properties of the ground layer. The ideia, from [22], is that the spectrum of the noise at bedrock is characterized by a constant pattern, while at the ground surface it is modified by the vibration property of the ground layer. This interaction between the bedrock and the ground layer is

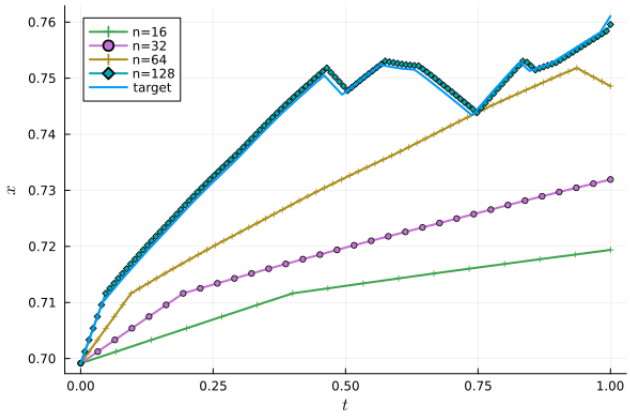

FIGURE 7. An approximation sequence of a sample path solution of the population dynamics equation (8.22).

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

$$\ddot{X}_t + 2\zeta_0\omega_0\dot{X}_t + \omega_0^2 X_t = -\ddot{M}_t. \tag{8.23}$$

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, $Y_t = \dot{M}_t$, generated by an Earthquake and its aftershocks, or any other 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 [5], 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, which is connected with a square power attact rate for the displacement itself. Finally, for simulation purposes, we model the displacement M_t instead of modeling directly the excitation $\ddot{M}_t$, but in such a way that the ground-motion excitation follows essentially the proposed signal.

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

$$M_t = \sum_{i=1}^{k} \gamma_i (t-\tau_i)_+^2 e^{-\delta_i (t-\tau_i)} \cos(\omega_i (t-\tau_i)), \tag{8.24}$$

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.

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 $s < 0$.

N	dt	error
64	0.0317	2.2
128	0.0157	0.932
256	0.00784	0.427
512	0.00391	0.205

TABLE 5. Mesh points (N), time steps (dt), and strong error (error) of the Euler method for the mechanical structure model under ground-shaking random excitations, with initial condition $X_0 = \mathbf{0}$ and transport process noise, on the time interval (0.0, 2.0), based on $m = 500$ sample paths for each fixed time step, with the target solution calculated with 262144 points.

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 $\tau_i \sim$ Exponential(0.25), $\gamma_i \sim$ Unif$(0.0, 4.0)$, $\delta_i \sim$ Unif$(8.0, 12.0)$, and $\omega_i \sim$ Unif$(8\pi, 32\pi)$.

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 $m = 500$. 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.

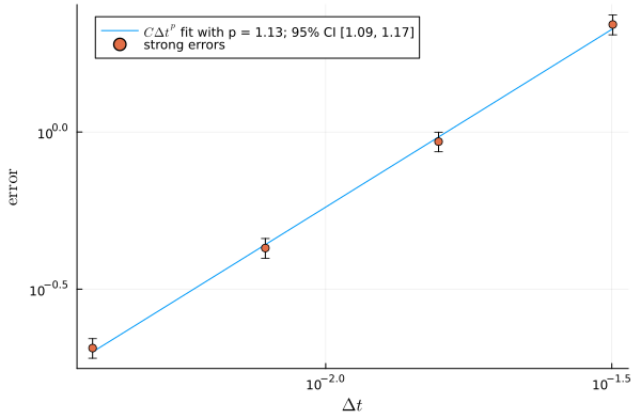

FIGURE 8. Order of convergence of the strong error of the Euler method for the mechanical structure model (8.23).

8.6. **A toggle-switch model for gene expression.** Here, we consider the toggle-switch model in [1, Section 7.8], originated from [35]; see also [33].

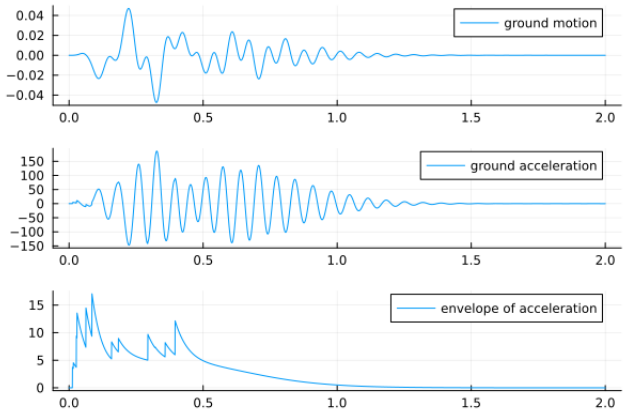

FIGURE 9. Sample ground motion and the corresponding excitation (ground acceleration) and envelope of excitation (just the exponential decay, without the oscillations) for the mechanical structure model (8.23).

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 [1, 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

$$\begin{cases} \dfrac{\mathrm{d}X_t}{\mathrm{d}t} = \left(A_t + \dfrac{X_t^4}{a^4 + X_t^4}\right)\left(\dfrac{b^4}{b^4 + Y_t^4}\right) - \mu X_t, \\ \dfrac{\mathrm{d}Y_t}{\mathrm{d}t} = \left(B_t + \dfrac{Y_t^4}{c^4 + Y_t^4}\right)\left(\dfrac{d^4}{d^4 + X_t^4}\right) - \nu Y_t, \end{cases} \tag{8.25}$$

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 a geometric Brownian motion process.

In the simulations below, we use parameters similar to those in [1, 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 a geometric Brownian motion process with initial value 0.2, drift 0.6 and diffusion

N	dt	error
16	0.267	1.91
32	0.129	0.871
64	0.0635	0.42
128	0.0315	0.205
256	0.0157	0.101
512	0.00783	0.0504

TABLE 6. Mesh points (N), time steps (dt), and strong error (error) of the Euler method for a toggle-switch model of gene regulation for each N, with initial condition $X_0 = 10.0, Y_0 = 10.0$ and compound Poisson process noises, on the time interval (0.0, 4.0), based on $m = 1000$ sample paths for each fixed time step, with the target solution calculated with 262144 points. The order of strong convergence is estimated to be $p = 1.03$

0.3. 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 $i = 4, \ldots, 9$. For the Monte-Carlo estimate of the strong error, we choose $m = 1000$. 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.

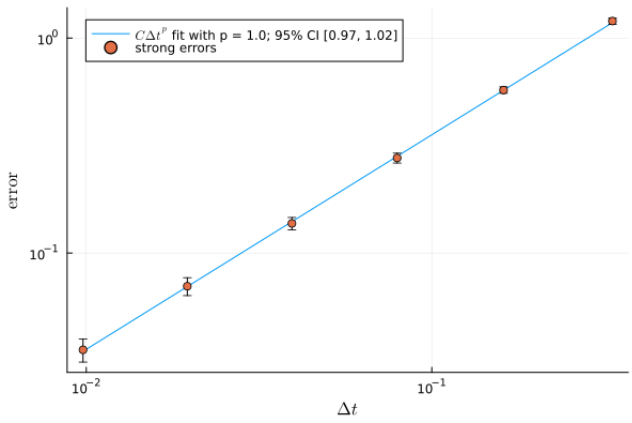

FIGURE 10. Order of convergence of the strong error of the Euler method for the toggle-switch model (8.25).

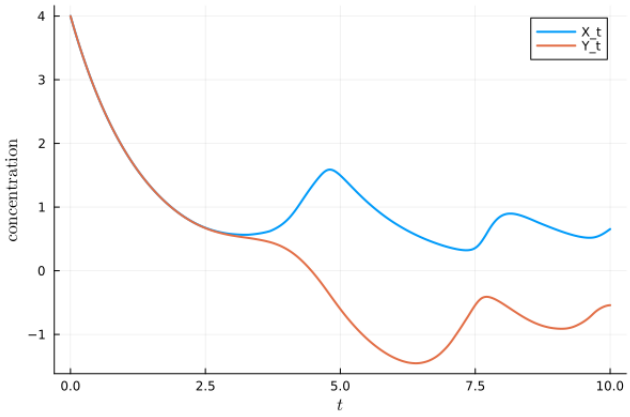

FIGURE 11. Evolution of a sample solution of the toggle-switch model (8.25).

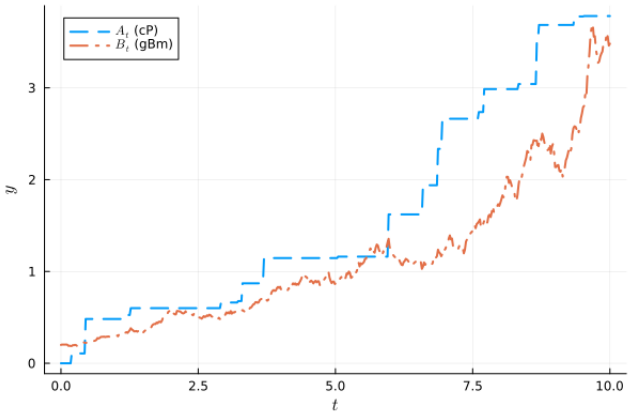

FIGURE 12. Noise sample paths for the toggle-switch model (8.25).

8.7. **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. [12]) 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 will consider 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

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

so the equation becomes

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

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

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

This leads us to the linear random ordinary differential equation

$$\frac{\mathrm{d}X_t}{\mathrm{d}t} = R_tX_t + R_t(C_t + O_t) + \nu O_t + \gamma. \tag{8.26}$$

This equation has the explicit solution

$$X_t = X_0e^{\int_0^t R_s\ \mathrm{d}s} + \int_0^t e^{\int_s^t R_\tau\ \mathrm{d}\tau}(R_s(C_s + O_s) + \nu O_s + \gamma)\ \mathrm{d}s.$$

As for the interest rate process $\{R_t\}_t$, there is a vast literature with models for it, see e.g. [6, Chapter 3] in particular Table 3.1. Here, we consider the Dothan

N	dt	error
64	0.0476	0.175
128	0.0236	0.0818
256	0.0118	0.043
512	0.00587	0.0212

TABLE 7. Mesh points (N), time steps (dt), and strong error (error) 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 $p = 1.0$

model ([6, Section 3.2.2] of the aforementioned reference), which consists simply of a geometric Brownian motion process

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

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

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

so that the equation (8.26) 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 $U_t = X_0 + O_0 + R_0 = 1.2$. 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.

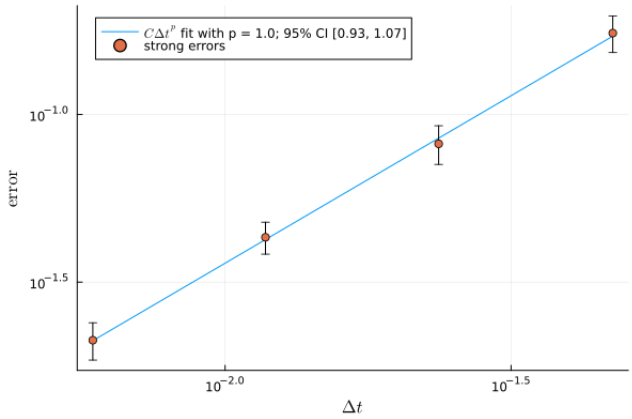

FIGURE 13. Order of convergence of the strong error of the Euler method for the actuarial risk model (8.26).

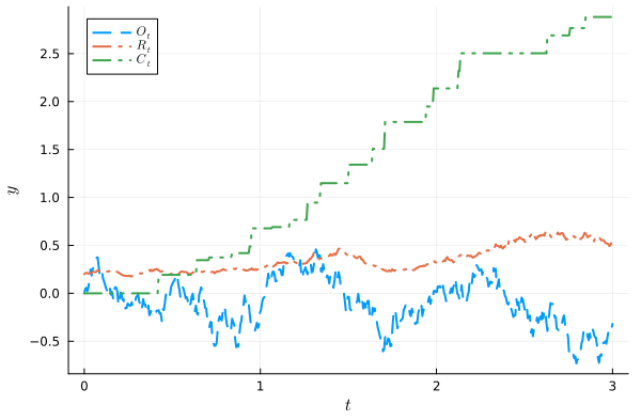

FIGURE 14. Sample noises for the risk model (8.26).

8.8. **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 [32] and [11]. 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.

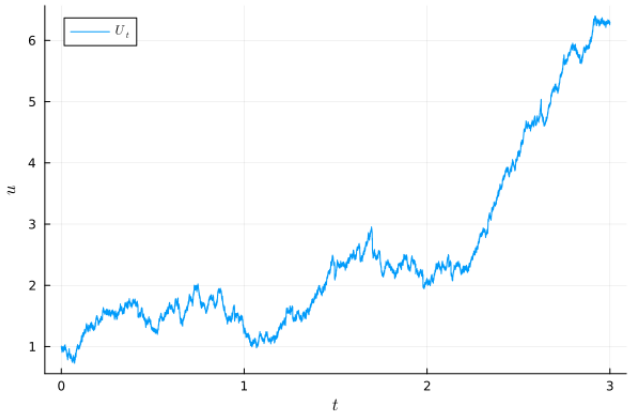

FIGURE 15. Sample surplus solution for the risk model (8.26).

The deterministic Fisher-KPP equations has its origins in [10] and [23]. It is a nonlinear parabolic equation of reaction-diffusion type, modeling inhomogeneous population growth displaying wave propagation, and many other phenomena such as combustion front wave propagation, physiollogy and crystallography pattern formation, and so on.

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

The equation takes the form

$$\frac{\partial u}{\partial t} = \mu \frac{\partial^2 u}{\partial x^2} + \lambda u \left(1 - \frac{u}{u_m}\right), \quad (t, x) \in (0, \infty) \times (0, 1), \tag{8.27}$$

endowed with the boundary conditions

$$\frac{\partial u}{\partial x}(t, 0) = -Y_t, \quad \frac{\partial u}{\partial x}(t, 1) = 0, \tag{8.28}$$

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 trains of incoming population.

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

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

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

The equation involves a nonlinear term which is not globally Lipschitz continuous, but, similiarly to the population dynamics model considered in Section 8.4, 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 certainly. The procedure is the same as that done in Section 8.4, so the details are omited.

The spatial discretization is done with finite differences on a grid with 17 mesh points. For the time-mesh parameters, we set $N_{\mathrm{tgt}} = 2^{15}\times 3^3\times 5 = 4,423,680$ and $N_i = 2^{10}, 2^7\times 3^2, 2^8\times 5, 2^9\times 3, 2^7\times 3\times 5, 2^{11}$. The reason for not using powers of two in this particular case is that, due to the spatial discretization, we cannot use too coarse a time mesh, for stability reasons, limiting the range of time-mesh parameters. Instead of going too high and have too slow computations, we populate the range of N_i's with values which are not power of two.

For the Monte-Carlo estimate of the strong error, we choose $m = 100$. Table 8 shows the estimated strong error obtained with this setup, while Figure 16 illustrates the order of convergence.

References

[1] 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.

[2] 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.

[3] 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.

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

[5] 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.

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

N	dt	error
1024	0.00196	0.243
1152	0.00174	0.214
1280	0.00156	0.194
1536	0.0013	0.165
1920	0.00104	0.131
2048	0.000977	0.122

TABLE 8. Mesh points (N), time steps (dt), and strong error (error) of the Euler method for Fisher-KPP equation for each N, with initial condition $X_0 = 0$ and Hawkes-modulated colored noise, on the time interval (0.0, 2.0), based on $m = 100$ sample paths for each fixed time step, with the target solution calculated with 4423680 points. The order of strong convergence is estimated to be $p = 0.98$

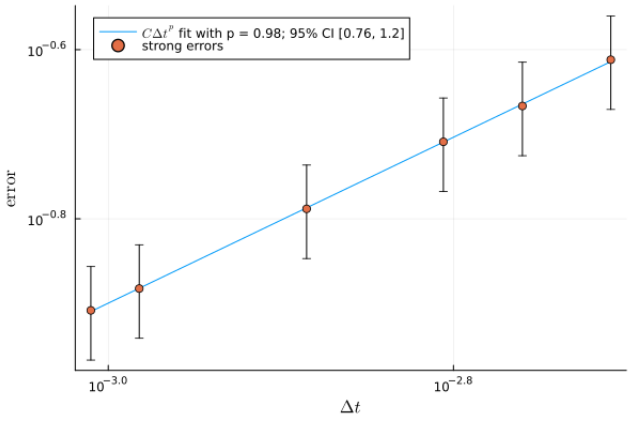

FIGURE 16. Order of convergence of the strong error of the Euler method for the Fisher-KPP model (8.27)-(8.28).

[7] D. S. Clark, Short proof of a discrete Gronwall inequality, *Discrete Applied Mathematics,* Vol. 16 (1987) no. 3, 279–281.
[8] E. A. Coddington and N. Levinson, *Theory of Ordinary Differential Equations,* New York: McGraw-Hill, 1987.
[9] A. B. Dieker and M. Mandjes, On spectral simulation of fractional Brownian motion, *Probability in the Engineering and Informational Sciences,* 17 (2003), 417–434.
[10] R. A. Fisher, The wave of advance of advantageous genes, *Annals of Eugenics,* 7 (1937), no. 4, 355–369.
[11] M. I. Freidlin and A. D. Wentzell, Reaction-diffusion equations with randomly perturbed boundary conditions, *Ann. Probab.* 20 (1992), no. 2, 963–986.
[12] H. U. Gerber, E. S. W. Shiu, On the time value of ruin, North American Actuarial J. vol. 2 (1998), no. 1, 48–72.

[13] 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.
[14] 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.
[15] 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.
[16] L. Grüne and P.E. Kloeden, Higher order numerical schemes for affinely controlled nonlinear systems, *Numer. Math.* 89 (2001), 669–690.
[17] X. Han and P. E. Kloeden, *Random Ordinary Differential Equations and Their Numerical Solution,* Probability Theory and Stochastic Modelling, vol. 85, Springer Singapore, 2017.
[18] 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.
[19] G. W. Housner and Paul C. Jennings, Generation of artificial Earthquakes, Journal of the Engineering Mechanics Division, 90 (1964), no. 1.
[20] A. Jentzen and P.E. Kloeden, *Taylor Approximations of Stochastic Partial Differential Equations,* CBMS Lecture series, SIAM, Philadelphia, 2011.
[21] 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.
[22] K. Kanai, Semi-empirical formula for the seismic characteristics of the ground, *Bull. Earthq. Res. Inst.,* Vol. 35 (1957), University of Tokyo.
[23] 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.
[24] H.-H. Kuo, *Introduction to Stochastic Integration,* Universitext, Springer New York, NY, 2006.
[25] B. B. Mandelbrot and J. W. Van Ness, Fractional Brownian motions, fractional noises and applications, *Siam Review,* Vol. 10 (1968), no. 4, 422–437.
[26] Y. S. Mishura, *Stochastic calculus for fractional Brownian motion and related processes,* Lecture Notes in Mathematics 1929, Springer-Verlag, Berlin, Heidelberg, 2008.
[27] T. Neckel and F. Rupp, *Random Differential Equations in Scientific Computing,* Versita, London, 2013.
[28] B. Øksendal, *Stochastic Differential Equations - An Introduction with Applications,* Universitext, Springer-Verlag Berlin Heidelberg, 2003.
[29] P. E. Protter, *Stochastic Integration and Differential Equations,* 2nd Edition, Springer-Verlag, Berlin Heidelberg New York, 2005.
[30] 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.
[31] 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`.
[32] 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.
[33] 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.

[34] 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.
[35] 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.
[36] 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 Tubingen, 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