.

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

RICARDO M. S. ROSA AND PETER E. KLOEDEN

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, we show that it is possible to exploit further conditions on the noise and prove that, in many typical cases, the strong convergence is actually of order 1, regardless of the Hölder regularity of the sample paths. This applies to additive or multiplicative Itô noises (such as Wiener, Ornstein-Uhlenbeck, and Geometric Brownian motion process); to point-process noises (such as Poisson point processes and Hawkes self-exciting processes, which are not even continuous and have jump-type discontinuities); and to transport-type processes. The order 1 convergence is based on a novel approach, resting on three main ideas: First, we do not estimate directly the local error and, instead, add up the local steps and work directly with an accumulated global error, leading to formula with an iterated integral, the outer one spanning the whole interval and the inner one spanning a time step. Secondly, we use Fubini theorem to switch the order of the iterated integral, moving the critical regularity to the large scale time, easing the regularity requirement on the small scale of the time step. Finally, we assume either a control of the total variation of the sample paths of the noise (as in many point processes and transport process) or that the noise is an Itô process (such as Wiener, Ornstein-Uhlenbeck, and Geometric Brownian motion) in order to bound the large scale term. We complement the work with examples with fractional Brownian motion noise with Hurst parameter $0 < H < 1/2$ for which the order of convergence is $H + 1/2$, hence below the attained order 1 in the examples above, but still above the H order of convergence obtained in previous works.

1. INTRODUCTION

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

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

Date: February 10, 2023.

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

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

on a time interval $I = [0, T]$, with $T > 0$, and where the noise $\{Y_t\}_{t\in I}$ is a given stochastic process. The sample space is denoted by Ω.

The Euler method for solving this initial value problem consists in approximating the solution on a uniform time mesh $t_j = j\Delta t_N$, $j = 0, \ldots, N$, 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

$$X_{t_j}^N = X_{t_{j-1}}^N + \Delta t_N f(t_{j-1}, X_{t_{j-1}}^N, Y_{t_{j-1}}), \qquad j = 1, \ldots, N, \tag{1.2}$$

with the initial condition

$$X_0^N = X_0. \tag{1.3}$$

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.

When the noise $\{Y_t\}_{t\in I}$ has θ-Hölder continuous sample paths, it can be show [10], under suitable regularity conditions on f, that the Euler scheme converges strongly with order θ with respect to the time step, i.e. there exists a constant $C \geq 0$ such that

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

where $\mathbb{E}[\cdot]$ indicates the expectation of a random variable on Ω.

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 convergence, with the sample 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 only when $1/2 \leq H < 1$, dropping to order $H + 1/2$, when $0 < H < 1/2$, which is still higher the Hölder exponent H of the sample paths.

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, namely (see Lemma 3.1)

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

for $j = 1, \ldots, N$, where τ^N is a piecewise constant function with jumps at the mesh points t_j (see Equation (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 continuity with respect to x. With those, we obtain the following basic bound for the global error (see Lemma 4.2)

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

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

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

Instead, we look at the whole noise error

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

and assume that the steps of the process given by $F_t = f(t, X_{\tau^N(t)}^N, Y_t)$ can be controlled in a suitable global way. In order to give the main idea, let us assume for the moment 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 noise. In any case, we bound the global error term using the Fubini Theorem,

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

Then, we find that

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

In the case of an Itô noise, we assume

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

with adapted processes $\{A_t\}_t$, $\{B_t\}_t$, and we bound the right hand side using the Lyapunov inequality and the Itô isometry:

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

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,

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

Using the monotonicity, this yields

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

yielding, again, strong order 1 convergence.

These two cases are treated in Section 5 (for the bounded variation case; see Lemma 5.1 and Theorem 5.1) and Section 6 (for the Itô noise case; see Lemma 6.1 and Theorem 6.1).

The conditions in Theorem 5.1 and Theorem 6.1 are not readily verifiable, but Theorem 5.2 and Theorem 6.2 give more explicit conditions for each of the two cases. Essentially, $f = f(t, x, y)$ is required to have minimal regularity in the sense of differentiability and growth conditions and the noise $\{Y_t\}_{t\in I}$ is either required to have sample paths of bounded variation or to be an Itô noise.

We complement this work with a few explicit examples and their numerical implementation, illustrating the strong order 1 convergence in the cases above. 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 have a general proof of this order of convergence in the case of fBm noise, but, in the example considered, we essentially have

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

In this case,

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

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

2. Pathwise solutions

For the notion and main results on pathwise solution for RODEs, we refer the reader to [11, Section 2.1] (see also [17, 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:

Hypothesis 2.1. *We consider a function $f = f(t, x, y)$ defined on $I \times \mathbb{R} \times \mathbb{R}$ and a real-valued stochastic process $\{Y_t\}_{t \in I}$, where $I = [0, T]$, $T > 0$. We make the following standing hypotheses.*

(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*
$$
|f(t, x_1, y) - f(t, x_2, y)| \leq L_X |x_1 - x_2|, \quad \forall t \in I, \ \forall x_1, x_2, y \in \mathbb{R}. \tag{2.1}
$$
(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 on $x \in \mathbb{R}$, 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 $x \in \mathbb{R}$ 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 $x \in \mathbb{R}$, 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 [6, Theorem 1.1]). Moreover,

the mapping $(t, \omega) \mapsto X_t(\omega)$ is measurable (see [11, 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 above, among other things. With that in mind, we have the following useful result.

Lemma 2.1. *Under 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 coincide 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 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*

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

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

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

Proof. Under 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 $t_j = t_{j-1} + \Delta t_N$.

For notational simplicity, we momentarily write $t = t_{j-1}$ and $\tau = \Delta t_N$, so that $t_j = t + \tau$. The exact pathwise solution satisfies

$$
X_{t+\tau} = X_t + \int_t^{t+\tau} f(s, X_s, Y_s) \,\mathrm{d}s.
$$

The Euler step is given by

$$
X_{t+\tau}^N = X_t^N + \tau f(t, X_t^N, Y_t).
$$

Subtracting, we obtain

$$
X_{t+\tau} - X_{t+\tau}^N = X_t - X_t^N + \int_t^{t+\tau} \big(f(s, X_s, Y_s) - f(t, X_t^N, Y_t) \big) \,\mathrm{d}s.
$$

We arrange the integrand as

$$
\begin{aligned}
f(s, X_s, Y_s) - f(t, X_t^N, Y_t) = & f(s, X_s, Y_s) - f(s, X_t, Y_s) \\
& + f(s, X_t, Y_s) - f(s, X_t^N, Y_s) \\
& + f(s, X_t^N, Y_s) - f(t, X_t^N, Y_t).
\end{aligned}
$$

This yields

$$\begin{aligned}X_{t+\tau} - X_{t+\tau}^N =& X_t - X_t^N \\ &= \int_t^{t+\tau} \left(f(s, X_s, Y_s) - f(s, X_t, Y_s)\right)\,\mathrm{d}s \\ &\quad + \int_t^{t+\tau} \left(f(s, X_t, Y_s) - f(s, X_t^N, Y_s)\right)\,\mathrm{d}s \\ &\quad + \int_t^{t+\tau} \left(f(s, X_t^N, Y_s) - f(t, X_t^N, Y_t)\right)\,\mathrm{d}s.\end{aligned}$$

Going back to the notation $t = t_{j-1}$ and $t + \tau = t_j$, the above identity reads

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

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

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

Using the jump function τ^N defined by (3.2), we may rewrite the above expression as in (3.1). $\square$

Remark 3.1. Strictly speaking, we only need condition (ii) from 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, under minimal hypotheses, that is the basis for the estimates in specific cases. For that, we use a discrete version of the Grownwall lemma. Here we state a particular case of a result that can be found in [8] (see also [5]).

Lemma 4.1 (Discrete Gronwall Lemma)**.** *Let $(e_j)_j$ be a (finite or infinite) sequence of positive numbers 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 $a, b > 0$. Then,

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

Proof. This follows from [8, Lemma V.2.4] by taking $n = j$, $a_n = e_j$, $b_n = 0$, $c_n = b$, and $\lambda = a$. For the sake of completeness, we give a direct proof for this particular case.

The result is trivially true for $j = 0$. Suppose, by induction, that the result is true up to $j - 1$. Then,

$$e_j \leq a \sum_{i=0}^{j-1} b e^{ai} + b = b \left(a \sum_{i=0}^{j-1} e^{ai} + 1 \right).$$

Using that $1 + a \leq e^a$, we have $a \leq e^a - 1$, hence

$$e_j \leq b \left((e^a - 1) \sum_{i=0}^{j-1} e^{ia} + 1 \right).$$

Using that $\sum_{i=0}^{j-1} \theta^i = (\theta^j - 1)(\theta - 1)$, with $\theta = e^a$, we see that

$$(e^a - 1) \sum_{i=0}^{j-1} e^{ia} \leq e^{ja} - 1,$$

so that

$$e_j \leq b e^{ja},$$

which completes the induction. □

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

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

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

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,

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

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

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

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

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

With these two estimates, we bound (3.1) as

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

Using the discrete version of the Gronwall Lemma, we prove (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 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 non-negative, 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

Proposition 4.1. *Under 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 $j = 0, \ldots, N$,

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

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

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

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

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

Using now estimate (4.7) and condition (4.5), we find (4.6), which completes 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 noise and f is not assumed to be differentiable, but, instead, that the steps can be controlled by monotonic nondecreasing processes with finite expected growth. This fits well with 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 Hypothesis 2.1, suppose that, for all $0 \leq s \leq T$,*

$$
|f(s, X_{\tau^N(s)}^N, Y_s) - f(\tau^N(s), X_{\tau^N(s)}^N, Y_{\tau^N(s)})| \leq \bar{F}_s - \bar{F}_{\tau^N(s)}, \tag{5.1}
$$

where $\{\bar{F}_t\}$ is a real 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,

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

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

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

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

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

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

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

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

When $\Delta t_N \leq t \leq T$, 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 $s - \Delta t_N \leq \tau^N(s) \leq s$ for all $\Delta t_N \leq s \leq T$,

$$\begin{aligned}\int_0^t (\mathbb{E}[\bar F_s] - \mathbb{E}[\bar F_{\tau^N(s)}])\;\mathrm{d}s &\leq \int_0^t \mathbb{E}[\bar F_s]\;\mathrm{d}s - \int_0^{\Delta t_N} \mathbb{E}[\bar F_0]\;\mathrm{d}s - \int_{\Delta t_N}^t \mathbb{E}[\bar F_{s-\Delta t_N}]\;\mathrm{d}s \\ &= \int_0^t \mathbb{E}[\bar F_s]\;\mathrm{d}s - \int_0^{\Delta t_N} \mathbb{E}[\bar F_0]\;\mathrm{d}s - \int_0^{T-\Delta t_N} \mathbb{E}[\bar F_s]\;\mathrm{d}s \\ &= \int_{t-\Delta t_N}^t \mathbb{E}[\bar F_s]\;\mathrm{d}s - \mathbb{E}[\bar F_0]\Delta t_N.\end{aligned}$$

Using again the monotonicity yields

$$\int_0^t (\mathbb{E}[\bar F_s] - \mathbb{E}[\bar F_{\tau^N(s)}])\;\mathrm{d}s \leq \int_{t-\Delta t_N}^t \mathbb{E}[\bar F_t]\;\mathrm{d}s - \mathbb{E}[\bar F_0]\Delta t_N = (\mathbb{E}[\bar F_t] - \mathbb{E}[\bar F_0])\Delta t_N.$$

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

Theorem 5.1. *Under Hypothesis 2.1, suppose also that* (2.4), (2.5), (4.5), (5.1), *and* (5.2) *hold. Then, the Euler scheme* (1.2)-(1.3) *is of strong order 1, i.e.*

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

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 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 (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

$$\begin{aligned}\mathbb{E}\left[|X_{t_j} - X_{t_j}^N|\right] \leq \Bigg(C_0\Delta t_N + \Delta t_N L_X\left(\mathbb{E}[|X_0|] + \int_0^{t_j} \mathbb{E}[M_\xi]\;\mathrm{d}\xi\right)e^{L_X t_j} \\ + (\mathbb{E}[\bar F_{t_j}] - \mathbb{E}[\bar F_0])\Delta t_N\Bigg)e^{L_X t_j}.\end{aligned}$$

Since this holds for every $j = 0, \ldots, N$, we obtain the desired (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 partial derivatives $\partial_t f$ 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 $(t, x, y) \in I \times \mathbb{R} \times \mathbb{R}$, *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 with

$$\mathbb{E}[|X_0|^2] < \infty. \tag{5.8}$$

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

$$\max_{j=0,\ldots,N} \mathbb{E}\left[\left|X_{t_j} - X_{t_j}^N\right|\right] \leq C\Delta t_N, \qquad \forall N \in \mathbb{N}, \tag{5.9}$$

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 the Hypothesis 2.1. Moreover, thanks to (5.7), we see that

$$\mathbb{E}[|Y_t|] \leq \mathbb{E}[|Y_t|^2] \leq \mathbb{E}[V(\{Y_t\}_{t\in I}; I)^2] < \infty.$$

Then, thanks to the Lyapunov inequality $\mathbb{E}[|Y_t|] \leq \mathbb{E}[|Y_t|^2]^{1/2}$, 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). We have

$$\begin{aligned}|f(s,X_\tau,Y_s) - f(\tau,X_\tau,Y_\tau)| = \left|\int_\tau^s \partial_t f(\xi, X_\tau, Y_\xi)\,\mathrm{d}\xi + \int_\tau^s \partial_y f(\xi, X_\tau, Y_\xi)\,\mathrm{d}Y_\xi\right| \\ \leq C_1(s-\tau) + C_2(s-\tau)|X_\tau| + (C_3 + C_4|X_\tau|)V(\{Y_t\}_{t\in I}; \tau, s).\end{aligned}$$

Thus, (5.1) holds with

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

It is clear that all the sample paths of $\{F_t\}_{t\in I}$ are monotonic 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 (5.8) 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 the $\partial_y f(t, x, y)$ 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.

Remark 5.2. One particular example that easily yields (5.6) is when $f = f(t, x, y)$ is semi-separable, i.e.

$$f(t, x, y) = a(t, y)h(x) + b(t, y), \tag{5.10}$$

where $a = a(t, y)$ and $b = b(t, y)$ are continuously differentiable on $I \times \mathbb{R}$ with uniformly bounded first derivatives, $a = a(t, y)$ itself is uniformly bounded, and $h = h(x)$ is globally Lipschitz continous on $\mathbb{R}$.

Since $a = a(t, x)$ is uniformly bounded and $h = h(x)$ is globally Lipschitz continuous, it follows that $f = f(t, x, y)$ is uniformly globally Lipschitz continuous in x. Moreover, it is continuously differentiable in (t, y), with partial derivatives $\partial_t f$ and $\partial_y f$ given by

$$\partial_t f = \partial_t a(t, y)h(x) + \partial_t b(t, y), \qquad \partial_y f = \partial_y a(t, y)h(x) + \partial_t b(t, y)$$

Since the partial derivatives of $a = a(t, y)$ and $b = b(t, y)$ are uniformly bounded and h is Lipschitz, it follows that the partial derivatives $\partial_t f$ and $\partial_y f$ have at most linear growth. Thus, (5.6) is satisfies and Theorem 5.2 applies. But this special form (5.10) is by no means necessary, and the result applies to more general terms $f = f(t, x, y)$, as stated in the theorem.

6. The case of an Itô 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 assuming that $f = f(t, x, y)$ is sufficiently regular and that the noise $\{Y_t\}_{t\in I}$ is itself an Itô process. With that in mind, we first have the following result.

Lemma 6.1. *Besides Hypothesis 2.1, suppose that $F_t^N = f(t, X_{\tau^N(t)}^N, Y_t)$ is an Itô noise, satisfying*

$$\mathrm{d}F_t^N = A_t \,\mathrm{d}t + B_t \,\mathrm{d}W_t, \tag{6.1}$$

for a Wiener process $\{W_t\}_{t\geq 0}$ and stochastic processes $\{A_t\}_{t\in I}$, $\{B_t\}_{t\in I}$ adapted to $\{W_t\}_{t\geq 0}$ and such that

$$\int_0^T \mathbb{E}[|A_t|] \,\mathrm{d}t < \infty, \quad \int_0^T \mathbb{E}[|B_t|^2] \,\mathrm{d}t < \infty. \tag{6.2}$$

Then,

$$\begin{aligned}\mathbb{E}\left[\left|\int_0^t \left(f(s, X_{\tau^N(s)}^N, Y_s) - f(\tau^N(s), X_{\tau^N(s)}^N, Y_{\tau^N(s)})\right) \,\mathrm{d}s\right|\right]& \\ \leq \Delta t_N \left(\int_0^t \mathbb{E}[|A_\xi|] \,\mathrm{d}\xi + \left(\int_0^t \mathbb{E}[|B_\xi|^2] \,\mathrm{d}\xi\right)^{1/2}\right),&\end{aligned} \tag{6.3}$$

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

Proof. We write

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

Upon integration,

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

Exchanging the order of integration, according to Fubini's theorem, yields

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

Taking the absolute mean and using the Itô isometry [14] on the second term gives

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

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

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

which completes the proof. □

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

Theorem 6.1. *Under Hypothesis 2.1, suppose also that* (2.4), (2.5), (4.5), (6.1), *and* (6.2) *hold. Then, the Euler scheme* (1.2)-(1.3) *is of strong order 1, i.e.*

$$\max_{j=0,\ldots,N} \mathbb{E}\left[\left|X_{t_j} - X^N_{t_j}\right|\right] \leq C\Delta t_N, \qquad \forall N \in \mathbb{N}, \tag{6.4}$$

for a constant C given by

$$\begin{aligned}
C = &\left(C_0 + L_X\left(\mathbb{E}[|X_0|] + \int_0^T \mathbb{E}[M_\xi]\;\mathrm{d}\xi\right)e^{L_X T}\right. \\
&\qquad\qquad \left. + \left(\int_0^T \mathbb{E}[|A_\xi|]\;\mathrm{d}\xi + \left(\int_0^T \mathbb{E}[|B_\xi|^2]\;\mathrm{d}\xi\right)^{1/2}\right)\right)e^{L_X T}
\end{aligned} \tag{6.5}$$

Proof. Under 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 (6.2), Lemma 6.1 applies and the last term in (4.6) is bounded according to (6.3). Using (6.3) in (4.6) yields

$$\begin{aligned}
\mathbb{E}\left[|X_{t_j} - X^N_{t_j}|\right] \leq &\left(C_0\Delta t_N + \Delta t_N L_X\left(\mathbb{E}[|X_0|] + \int_0^{t_j} \mathbb{E}[M_\xi]\;\mathrm{d}\xi\right)e^{L_X t_j}\right. \\
&\qquad \left. + \Delta t_N\left(\int_0^{t_j} \mathbb{E}[|A_\xi|]\;\mathrm{d}\xi + \left(\int_0^{t_j} \mathbb{E}[|B_\xi|^2]\;\mathrm{d}\xi\right)^{1/2}\right)\right)e^{L_X t_j}.
\end{aligned}$$

Since this holds for every $j = 0, \ldots, N$, we obtain the desired (6.4). □

In practice, conditions (6.1)-(6.2) follows from assuming sufficent regularity on $f = f(t, x, y)$ and an Itô noise, as given by the 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 noise,*

$$\mathrm{d}Y_t = a(t, Y_t)\,\mathrm{d}t + b(t, Y_t)\,\mathrm{d}W_t, \tag{6.6}$$

with $a = a(t, y)$ and $b = b(t, y)$ continuous and satisfying

$$|a(t, y)| \leq A_M + A_Y|y|, \qquad |b(t, y)| \leq B_M + B_Y|y|. \tag{6.7}$$

Assume the bounds (2.4), (4.5), and

$$\mathbb{E}[|Y_0|] < \infty \tag{6.8}$$

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

$$\max_{j=0,\ldots,N} \mathbb{E}\left[\left|X_{t_j} - X_{t_j}^N\right|\right] \leq C\Delta t_N, \qquad \forall N \in \mathbb{N}, \tag{6.9}$$

for a suitable constant $C \geq 0$.

Proof. Let us start by showing that 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 the noise has continuous sample paths. Thus, the remaining condition in Hypothesis 2.1 to be verified is (iic).

From (6.6) and (6.7), we have

$$Y_t = \int_0^t a(s, Y_s)\,\mathrm{d}s + \int_0^s b(t, Y_s)\,\mathrm{d}W_s.$$

Using the Itô formula, we have

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

Thus

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

Taking the expectation,

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

Using (6.7), this yields

$$\begin{aligned}\mathbb{E}[|Y_t|^2] &\leq \mathbb{E}[|Y_0|^2] + \int_0^t \left(2\mathbb{E}[(A_M + A_Y|Y_s|)|Y_s|] + \mathbb{E}[(B_M + B_Y|Y_s|)^2]\right)\,\mathrm{d}s \\ &\leq \mathbb{E}[|Y_0|^2] + \int_0^t \left(4(A_M^2 + (1 + A_Y)\mathbb{E}[|Y_s|^2]) + 2(B_M^2 + B_Y^2\mathbb{E}[|Y_s|^2])\right)\,\mathrm{d}s\end{aligned}$$

By the classical Gronwall Lemma [9],

$$\mathbb{E}[|Y_t|^2] \leq \left(\mathbb{E}[|Y_0|^2] + (4A_M^2 + 2B_M^2)t\right) e^{(4(1+A_Y)+2B_Y^2)t}.$$

Thus,

$$\sup_{t\in I} \mathbb{E}[|Y_t|^2] \leq \left(\mathbb{E}[|Y_0|^2] + (4A_M^2 + 2B_M^2)T\right) e^{(4(1+A_Y)+2B_Y^2)T}. \tag{6.10}$$

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 (6.10), we see that

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

Therefore, we are under the condition of (2.1).

Now, in view of Theorem 6.1, it remains to prove that $F_t^N = f(t, X_{\tau^N(t)}^N, Y_t)$ is an Itô noise (6.1), with the bounds (6.2). The fact that it is an Itô noise follows from the Itô formula and the smoothness of $f = f(t, x, y)$. Indeed, since $(t, y) \mapsto f(t, x, y)$ is twice continuously differentiable, for each fixed x, the Itô formula is applicable and yields

$$\begin{aligned} \mathrm{d}f(t, x, Y_t) = \left(\partial_t f(t, x, Y_t) + a(t, Y_t)\partial_y f(t, x, Y_t) + \frac{b(t, Y_t)^2}{2}\partial_{yy} f(t, x, Y_t)\right) \,\mathrm{d}t \\ + b(t, Y_t)\partial_y f(t, x, Y_t) \,\mathrm{d}W_t, \end{aligned} \tag{6.11}$$

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

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

It remains to show that $\{A_t\}_{t\in I}$ is mean integrable and that $\{B_t\}_{t\in I}$ is square mean integrable. Since $f = f(t, x, y)$ has uniformly bounded derivatives, we have

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

and

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

for a suitable constants $L_{YY} \geq 0$. Now, thanks to (6.10), we see that (6.2) is satisfied.

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

Remark 6.1. When the diffusion term $b = b(t, y) = b(t)$, in (6.6), is actually independent of y, then the noise is an additive noise and in this case the Euler scheme is well known to be of strong order 1 [12]. In the more general $b = b(t, y)$ case, however, the Euler scheme has always been regarded to be of order 1/2 [10] (see also [18] 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 Hypothesis 2.1, suppose also that* (2.4), (2.5), (4.5). *Suppose, moreover, that* $F_t^N = f(t, X_{\tau^N(t)}^N, Y_t)$ *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* (6.2) *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.*

$$|H_s^N - H_{\tau^N(s)}^N| \leq \bar{H}_s^N - \bar{H}_{\tau^N(s)}^N \tag{7.1}$$

with

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

Then, the Euler scheme (1.2)-(1.3) *is of strong order 1, i.e. there exists a constant* $C \geq 0$ *such that*

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

We ommit 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 components, one of bounded variation with finite quadratic mean, as in* (5.7), *and another an Itô noise satisfying* (6.6) *and* (6.8). *Assume, moreover, that* (5.8) *holds. Then, the Euler scheme is of strong order 1, i.e.*

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

for a suitable constant $C \geq 0$.

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 also illustrate the $H+1/2$ order of convergence in the case of a fractional Brownian motion noise with Hurst parameter $0 < H < 1/2$.

For estimating the order of convergence, we use the Monte Carlo method, computing a number of numerical approximations $\{X_{t_j}^N(\omega_m)\}_{j=0,\ldots,N}$, of sample path solutions $\{X_t(\omega_m)\}_{t\in I}$, for samples $\omega_1, \ldots, \omega_M$, and taking the maximum in time of the average of their absolute differences at the mesh points:

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

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 $M = 1,000$ or $M = 10,000$.
(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 $\Delta_i = T/(N_i - 1)$, with $N_i = 2^{n_i}$, most often with $n_i = 4, \ldots, 10$, hence $N_i = 32, 64, \ldots, 1024$.
(v) A number N_{fine} of mesh points for a fine discretization to compute a target solution path, typically $N_{\text{fine}} = \max_i\{N_i^2\}$, e.g. $N = 2^{20} = 1,048,576$;
(vi) The target solution path is either an exact pathwise solution, when available, or a higher-order approximation, thanks to the choice of N_{fine}.

And here is the method:

(i) For each sample $m = 1, \ldots, M$, we first generate a discretization $\{Y_{t_j}\}_{j=0,N_{\text{fine}}}$ of a sample path of the noise on the finest grid $\{t_j^{N_{\text{fine}}}\}$, with N_{fine} points, using an exact distribution for the noise.
(ii) Next, we use the values of the noise at the finest time mesh to generate the target solution $\{X_{t_j}\}_{j=0,N_{\text{fine}}}$, still on the fine mesh. This is constructed either using the Euler approximation itself, keeping in mind that the mesh is sufficiently fine, or by an exact distributions of the solution, when available.
(iii) Then, for each time step $N_i = 2^{n_i}$ in the selected range, we compute the Euler approximation using the computed noise values at the corresponding coarse mesh.
(iv) We then compare each approximation $\{X_{t_j}^{N_i}\}_{j=0,\ldots,N_i}$ to the values of the target path on that coarse mesh and update the strong error

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

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

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

(vi) Finally, 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.

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 performant and feature-rich package for solving differential equations [15], including random and stochastic differential equations. For the sake of transparency, however, all the numerics presented below were implemented explicitly, complementing the aforementioned package, and made available in the github repository [16].

8.1. **Linear homogeneous equation.** We start by considering the Euler approximation of one of the simplest random ordinary differential equation, that of a linear homogenous equation with a Wiener process as the coefficient:

$$\begin{cases} \dfrac{\mathrm{d}X_t}{\mathrm{d}t} = W_t X_t, \qquad 0 \leq t \leq T, \\ X_t\big|_{t=0} = X_0, \end{cases} \tag{8.2}$$

This 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 solutions 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.

First we break down the sum into parts:

$$\int_0^{t_j} W_s \,\mathrm{d}s = \sum_{i=0}^{j-1} \int_{t_i}^{t_{i+1}} W_s \,\mathrm{d}s. \tag{8.4}$$

On each mesh interval $[t_i, t_{i+1}]$, we consider the process

$$B_t^i = W_t - W_{t_i} - \frac{t - t_i}{t_{i+1} - t_i}(W_{t_{i+1}} - W_{t_i}) \tag{8.5}$$

which is a Brownian bridge on that mesh interval, vanishing at the extremes, and independent of W_{t_i} and $W_{t_{i+1}}$. Then,

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

where

$$Z_i = \int_{t_i}^{t_{i+1}} B_s^i \,\mathrm{d}s. \tag{8.6}$$

Notice the first term is the trapezoidal rule while the second term is a Gaussian with zero mean. We need to compute the variance of Z_i to completely characterize it. By translation, it suffices to consider a Brownian bridge $\{B_t\}_{t\in[0,\tau]}$ on an interval $[0,\tau]$, with $\tau = \Delta t_N$. This is obtained from $B_t = W_t - (t/\tau)W_\tau$. We have, since $\mathbb{E}[W_t W_s] = \min\{t, s\}$, that

$$\mathbb{E}[B_t B_s] = \min\{t, s\} - \frac{ts}{\tau}.$$

Hence,

$$\begin{aligned}
\mathbb{E}\left[\left(\int_0^\tau B_s \,\mathrm{d}s\right)^2\right] &= \mathbb{E}\left[\int_0^\tau \int_0^\tau B_s B_t \,\mathrm{d}s \,\mathrm{d}\right] \\
&= \int_0^\tau \int_0^\tau \mathbb{E}[B_s B_t] \,\mathrm{d}s \,\mathrm{d}t \\
&= \int_0^\tau \int_0^\tau \left(\min\{t, s\} - \frac{ts}{\tau}\right) \mathrm{d}s \,\mathrm{d}t \\
&= \int_0^\tau \int_0^t s \,\mathrm{d}s \,\mathrm{d}t + \int_0^\tau \int_t^\tau t \,\mathrm{d}s \,\mathrm{d}t - \int_0^\tau \int_0^\tau \frac{ts}{\tau} \,\mathrm{d}s \,\mathrm{d}t \\
&= \int_0^\tau \frac{t^2}{2} \,\mathrm{d}t + \int_0^\tau t(\tau - t) \,\mathrm{d}t - \int_0^\tau \frac{t\tau^2}{2\tau} \,\mathrm{d}t \\
&= \frac{\tau^3}{12}.
\end{aligned}$$

Back to Z_i, this means that

$$Z_i \sim \mathcal{N}\left(0, \frac{(t_{i+1} - t_i)^3}{12}\right) = \frac{\sqrt{(t_{i+1} - t_i)^3}}{\sqrt{12}} \mathcal{N}(0, 1). \tag{8.7}$$

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

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

This is the contribution of this random variable to the mean of the exact solution. But we actually need to use the exact $e^{\sum_i Z_i}$ for a more reliable estimate.

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

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

FIGURE 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 solutios compatible with the given realizations of the noise on the mesh points.

N	dt	error
16	0.0667	0.0416
32	0.0159	0.0201
64	0.0323	0.0101
128	0.00787	0.00516
256	0.00392	0.00261
512	0.00196	0.00129
1024	0.000978	0.000637
2048	0.000489	0.000314

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 $X_0 \sim \mathcal{N}(0, 1)$, on the time interval $[0, 1]$, based on 1000 samples for each fixed time step, with the exact solution calculated with $2^{16} = 65536$ points.

for realizations Z_i drawn from a normal distributions given by (8.7). Figure 8.1 shows an approximate solution and a few sample paths of possible exact solutions associated with the given realizations of the noise on the mesh points.

The Table 8.1 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 $[0, T]$. The Figure 8.1 illustrates the order of convergence.

Remark 8.1. The extra multiplicative exponential term involving $\sum_i Z_i$ is estimated to be, since these random variables are independent, of order

$$\begin{aligned}
\mathbb{E}\left[\left|e^{\sum_i Z_i} - 1\right|\right] \sim \mathbb{E}\left[\left|\sum_i Z_i\right|\right] &\leq \left(\mathbb{E}\left[\left(\sum_i Z_i\right)^2\right]\right)^{1/2} \\
&= \left(\sum_i \mathbb{E}\left[Z_i^2\right]\right)^{1/2} \sim \left(\sum_i \Delta t_N^3\right)^{1/2} = \left(t_j \Delta t_N^2\right)^{1/2} \sim \Delta t_N
\end{aligned}$$

hence of strong order 1 in Δt_N. Therefore, this term does not affect the order of convergence of the Euler method, but it is of crucial importance when analysing higher order methods.

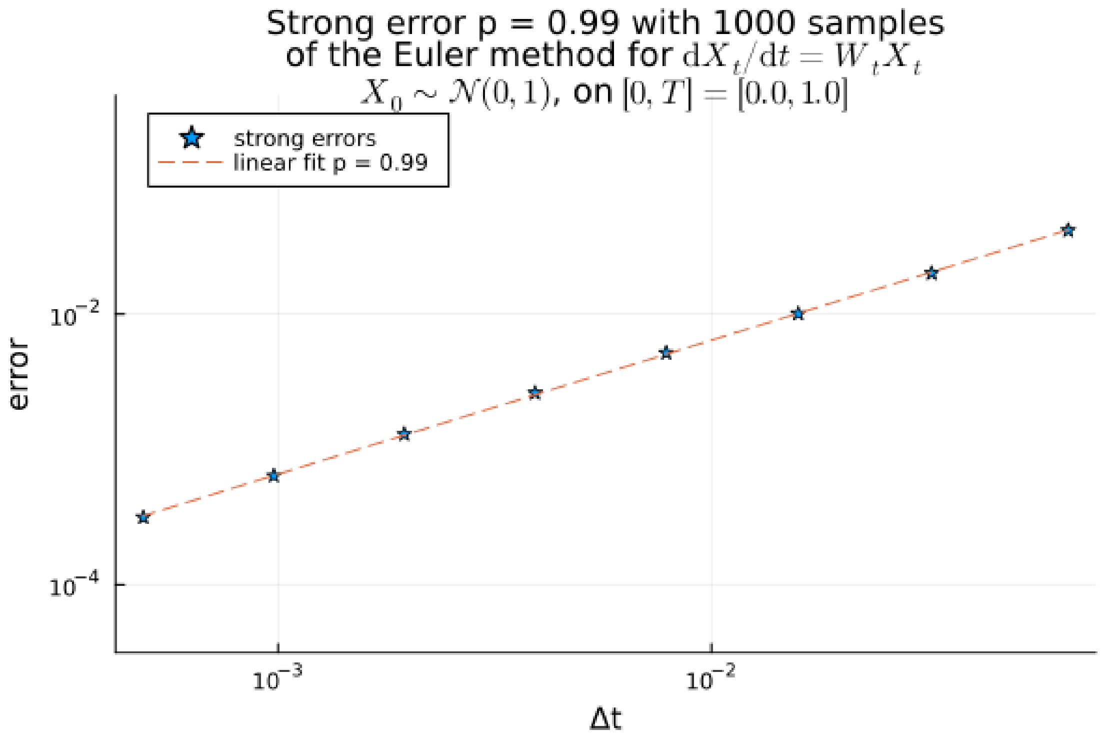

FIGURE 2. Convergence of the strong error for the Euler approximation of $\mathrm{d}X_t/\mathrm{d}t = W_t X_t$, with initial condition $X_0 \sim \mathcal{N}(0,1)$, on the time interval $[0,1]$, based on 1000 sample paths, for each of the time steps in $\mathrm{d}t = 1/N$, with $N = 16, 32, 64, 128, 256, 512, 1024, 2048$.

8.2. **Linear inhomogeneous equation.** We now consider the following linear equation with noise in the nonhomogenous term:

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

This has the explicit solution

$$X_t = e^{-t}X_0 + \int_0^t e^{-(t-s)} W_s \,\mathrm{d}s. \tag{8.11}$$

As in Section 8.1, when we compute an approximate solution via Euler's method, we only draw the realizations $\{W_{t_i}\}_{i=0}^n$ of a sample path, on the mesh points. We can find the exact distribution for $\int_0^{t_j} e^s W_s \,\mathrm{d}s$ given the values of the noise at the mesh points and use that to draw feasible exact solutions to estimate the strong error.

First we break down the sum into parts:

$$\int_0^{t_j} e^s W_s \, \mathrm{d}s = \sum_{i=0}^{j-1} \int_{t_i}^{t_{i+1}} e^s W_s \, \mathrm{d}s. \tag{8.12}$$

On each mesh interval $[t_i, t_{i+1}]$, we consider, again, the Brownian bridge

$$B_t^i = W_t - W_{t_i} - \frac{t - t_i}{t_{i+1} - t_i}(W_{t_{i+1}} - W_{t_i}). \tag{8.13}$$

Then,

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

where

$$Z_i = \int_{t_i}^{t_{i+1}} e^s B_s^i \, \mathrm{d}s. \tag{8.14}$$

As before, the second term is a Gaussian with zero mean, and we need to compute its variance to completely characterize it. By translation, it suffices to consider a Brownian bridge $\{B_t\}_{t \in [0,\tau]}$ on an interval $[0, \tau]$, with $\tau = \Delta t_N$. This is obtained from $B_t = W_t - (t/\tau) W_\tau$. We have, since $\mathbb{E}[W_t W_s] = \min\{t, s\}$, that

$$\mathbb{E}[B_t B_s] = \min\{t, s\} - \frac{ts}{\tau}.$$

Hence,

$$\begin{aligned}
\mathbb{E}\left[\left(\int_0^\tau e^s B_s \, \mathrm{d}s \right)^2 \right] &= \mathbb{E}\left[\int_0^\tau \int_0^\tau e^s e^t B_s B_t \, \mathrm{d}s \, \mathrm{d} \right] \\
&= \int_0^\tau \int_0^\tau e^s e^t \mathbb{E}[B_s B_t] \, \mathrm{d}s \, \mathrm{d}t \\
&= \int_0^\tau \int_0^\tau e^s e^t \left(\min\{t, s\} - \frac{ts}{\tau} \right) \, \mathrm{d}s \, \mathrm{d}t \\
&= \int_0^\tau \int_0^t e^s e^t s \, \mathrm{d}s \, \mathrm{d}t + \int_0^\tau \int_t^\tau e^s e^t t \, \mathrm{d}s \, \mathrm{d}t - \int_0^\tau \int_0^\tau e^s e^t \frac{ts}{\tau} \, \mathrm{d}s \, \mathrm{d}t \\
&= \int_0^\tau e^t (t e^t - e^t + 1) \, \mathrm{d}t + \int_0^\tau t e^t (e^\tau - e^t) \, \mathrm{d}t \\
&\qquad - \int_0^\tau \frac{t e^t}{\tau} \left(\tau e^\tau - e^\tau + 1 \right) \, \mathrm{d}t \\
&= \frac{\tau^3}{12}.
\end{aligned}$$

FIGURE 3. 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 solutios compatible with the given realizations of the noise on the mesh points.

1	2	3
4	5	6

TABLE 2. blah

Back to Z_i, this means that

$$Z_i \sim \mathcal{N}\left(0, \frac{(t_{i+1} - t_i)^3}{12}\right) = \frac{\sqrt{(t_{i+1} - t_i)^3}}{\sqrt{12}} \mathcal{N}(0, 1). \tag{8.15}$$

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,

$$\mathbb{E}[e^{Z_i}] = e^{((t_{i+1} - t_i)^3)/24}. \tag{8.16}$$

This is the contribution of this random variable to the mean of the exact solution. But we actually need to use the exact $e^{\sum_i Z_i}$ for a more reliable estimate.

Hence, once an Euler approximation of (8.10) 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

$$X_t = X_0 e^{\sum_{i=0}^{j-1} \left(\frac{1}{2}\left(W_{t_i} + W_{t_{i+1}}\right)(t_{i+1} - t_i) + Z_i\right)}, \tag{8.17}$$

for realizations Z_i drawn from a normal distributions given by (8.15). Figure 8.2 shows an approximate solution and a few sample paths of possible exact solutions associated with the given realizations of the noise on the mesh points.

The Table 8.2 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 $[0, T]$. The Figure 8.2 illustrates the order of convergence.

FIGURE 4. Convergence of the strong error for the Euler approximation of $\mathrm{d}X_t/\mathrm{d}t = W_t X_t$, with initial condition $X_0 \sim \mathcal{N}(0, 1)$, on the time interval $[0, 1]$, based on 1000 sample paths, for each of the time steps in $\mathrm{d}t = 1/N$, with $N = 16, 32, 64, 128, 256, 512, 1024$.

8.3. **Population dynamics with harvest.** Here, we consider a population dynamics modelled by a logistic equation (loosely inspired by [11, Section 15.2]) with harvest:

$$\frac{\mathrm{d}X_t}{\mathrm{d}t} = Z_t X_t (r - X_t) - H_t \tag{8.18}$$

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

$$Z_t = \lambda(1 + \varepsilon \sin(Y_t)),$$

where $0 < \varepsilon < 1$ and $\{Y_t\}_{t\geq 0}$ is a geometric Brownian motion process, hence of the form (6.6)-(6.7), and $\{H_t\}_{t\geq 0}$ is a point process of the form ...

We suppose the initial condition is non-negative and bounded almost surely:

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

for some $R > r$.

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

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

Define

$$f(t, x, z) = zx(r - x)$$

and notice that $f(t, x, z)x \geq 0$, for $x \geq 0$ and $z \geq 0$, and $f(t, x, z)x \leq 0$, for $x \geq r$ and $z \geq 0$. Hence the interval $[0, R]$ in x is positively invariant and the pathwise solutions of (8.18) are almost surely bounded as well, with

$$0 \leq X_t \leq R,$$

for all $t \geq 0$.

The function $f = f(t, x, z)$ is continuously differentiable infinitely many times and with

$$\left|\frac{\partial f}{\partial x}(t, x, z)\right| = |z(r - 2x)| \leq 2\lambda(2R - r),$$

for $|x| \leq R$ and $0 \leq z \leq 2\lambda$. In turn, the function $z = z(y) = \lambda(1 + \varepsilon \sin(y))$ is also continuously differentiable infinitely many times and is uniformly bounded along with all its derivatives.

The right hand side of (8.18) is not globally Lipschitz, but, for the sake of analysis, since X_t and Y_t are bounded, the right hand side can be modified to a twice continuously differentiable, uniformly globally Lipschitz function $\tilde{f}(t, x, y)$ that coincides with $f(t, x, y)$ for $(t, x, y) \in \mathbb{R} \times [0, R] \times [0, 2\lambda]$ and satisfies (2.1) with

$$L_X = 2\lambda(2R - r).$$

Thus, the RODE (8.18) with $0 \leq X_0 \leq R$ almost surely, for some $R > r$, is equivalent to the RODE

$$\frac{\mathrm{d}X_t}{\mathrm{d}t} = \tilde{f}(t, X_t, Y_t). \tag{8.19}$$

With $\tilde{f} = \tilde{f}(t, x, y)$, the Hypothesis 2.1 hold. Moreover, it follows from (2.3) (notice $M_t = 0$ here) that

$$|X_t| \leq |X_0| e^{2\lambda(2R-r)t} \leq R e^{2\lambda(2R-r)T}, \qquad 0 \leq t \leq T.$$

almost surely.

Therefore, all the conditions of Theorem 6.2 hold and the Euler method is of strong order 1.

8.4. **Earthquake and other impulse driven models.** See Neckel and Rupp pg 582 as a starting point of a model driven by a transport process as the source of ground motion excitation (both Kanai-Tajima and Bogdanoff, check out also the Clough-Penzien).

8.5. **Fractional Brownian motion noise.** Here, we consider again the linear equation

$$\begin{cases} \dfrac{\mathrm{d}X_t}{\mathrm{d}t} = -X_t + B_t^H, \qquad 0 \leq t \leq T, \\ X_t\big|_{t=0} = X_0, \end{cases} \tag{8.20}$$

except now the noise $\{B_t^H\}_t$ is assumed to be a fractional Brownian motion (fBm) with Hurst parameter $0 < H < 1$. We show 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 [13, 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.

Although the above linear equation has the explicit solution

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

we check the convergence numerically by solving the equation with the Euler method itself, but using as target the approximation in the finest mesh.

In this particular case, we can prove rigorously the strong order of convergence is $H + 1/2$, for $0 < H < 1/2$. Indeed, we need to estimate the last term in (4.6), involving the steps of the term $f(t, x, y) = -x + y$, which in this case reduce to

$$f(s, X_{\tau^N(s)}^N, Y_s) - f(\tau^N(s), X_{\tau^N(s)}^N, Y_{\tau^N(s)}) = B_s^H - B_{\tau^N(s)}^H,$$

for $0 \leq s \leq T$. Now, we use the expression

$$B_s^H = \int_0^s$$

Hence, the last term in (4.6) becomes

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

Acknowledgments

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, 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, T. Zhang. *Stochastic Calculus for Fractional Brownian Motion and Applications,* Springer-Verlag, London, 2008.

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

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

[7] H. Gjessing, H. Holden, T. Lindstrøn, B. Øksendal, J. Ubøe, 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.

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

[9] 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.

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

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

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

[13] Y. S. Mishura, *Stochastic calculus for fractional Brownian motion and related processes,* Lecture Notes in Mathematics 1929, Springer-Verlag, Berlin, Heidelberg, 2008.

[14] B. Øksendal, *Stochastic Differential Equations - An Introduction with Applications,* Universitext, Springer-Verlag Berlin Heidelberg 2003.

[15] C. Rackauckas, 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.

[16] P. Kloeden, 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`.

[17] F. Rupp, T. Neckel, *Random Differential Equations in Scientific Computing,* Versita, London, 2013.

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

(Ricardo M. S. Rosa) Instituto de Matemática, Universidade Federal do Rio de Janeiro, Brazil

(Peter E. Kloeden) Mathematics Department, University of Tubingen, Germany
Email address, R. M. S. Rosa: rrosa@im.ufrj.br
Email address, P. E. Kloeden: kloeden@math.uni-frankfurt.de