.

CONDITIONS FOR THE STRONG ORDER 1 CONVERGENCE OF THE EULER-MARUYAMA APPROXIMATION FOR RANDOM ORDINARY DIFFERENTIAL EQUATIONS

RICARDO M. S. ROSA AND PETER E. KLOEDEN

Abstract. It is well known that the Euler-Maruyama method of approximating 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 $f = f(t, x, w)$ is sufficiently regular. Here, we show that, in common situations, it is possible to exploit "hidden" conditions on the noise and prove that the strong convergence is actually of order 1, regardless of much regularity on the sample paths. This applies to Itô process noises (such as Wiener, Ornstein-Uhlenbeck, and Geometric Brownian process), which are Hölder continuous, and to point processes (such as Poisson point processes and Hawkes self-exciting processes), which are not even continuous and have jump-type discontinuities, as well as to transport processes. The order 1 convergence follows from not estimating directly the local error, but, instead, adding up the local steps and estimating the compound error. In the case of an Itô noise, the compound error is then estimated via Itô formula and the Itô isometry. In the case of a point process or a transport process, a monotonic bound is exploited. We HOPEFULLY complement the result by giving examples where some of the conditions are not met and the order of convergence seems indeed to be less than 1.

1. Introduction

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

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

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

The Euler-Maruyama 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-Maruyama

---

Date: October 26, 2022.

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

Key words and phrases. random ordinary differential equations, Euler-Maruyama method, strong convergence, Itô process, point process.

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 [1], under further suitable conditions, that the Euler-Maruyama scheme converges strongly with order θ with the time step, i.e. there exists a constant $C > 0$ 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 "hidden" conditions that yield in fact a strong order 1 convergence, even when the sample paths are still Hölder continuous or have jump discontinuities. This is the case, for instance, when the noise is an Itô noise, or when the equation is semi-separable and the noise is a point process or a transport process.

More precisely, for the semi-separable case, we assume f is of the form

$$f(t, x, y) = a(t, y)h(x) + b(t, y).$$

In this case, we assume the processes $\{a(t, Y_t)\}_{t\in I}$ and $\{b(t, Y_t)\}_{t\in I}$ have their steps bounded by a monotonic process, which typically happens for point processes, i.e.

$$|a(t+\tau, Y_{t+\tau}) - a(t, Y_t)| \leq A_t, \quad |b(t+\tau, Y_{t+\tau}) - b(t, Y_t)| \leq B_t,$$

where $\{A_t\}_{t\in I}$ and $\{B_t\}_{t\in I}$ have monotonically non-decreasing sample paths. Under further suitable conditions (see Corollary 6.1), we show that the Euler-Maruyama method is of strong order 1, i.e. (1.4) holds with $\theta = 1$.

Even if the structure of the equation is not exactly semi-separable but it is somehow possible to bound the steps $|f(t+\tau, X_t, Y_{t+\tau}) - a(t, X_t, Y_t)|$ by a suitable process with monotonic non-decreasing sample paths, it is possible to prove the strong order 1 convergence (see Theorem 6.1).

For the Itô noise case, we consider a general equation of the form (1.1), with a noise defined as an **Itô process** $\{Y_t\}_{t\geq 0}$, satisfying

$$\mathrm{d}Y_t = A_t \,\mathrm{d}t + B_t \,\mathrm{d}W_t, \tag{1.5}$$

We are not solving for Y_t, nor approximating it numerically, otherwise we would actually need to consider a system of stochastic differential equations. Instead, we assume it is a known process that can be computed analytically, such as a Wiener process, an Ornstein-Uhlenbeck process, or a geometric Brownian motion. With those in mind, A_t and B_t may be originally given in terms of $\{W_t\}_{t\geq 0}$ and $\{Y_t\}_{t\in I}$, but the general assumption is only given in terms of $\{A_t\}_{t\in I}$ and $\{B_t\}_{t\in I}$.

In the case that $f = f(t, x, y)$ is twice continuously differentiable, the Itô formula applies and we show, under suitable conditions on $\{A_t\}_{t\in I}$, $\{B_t\}_{t\in I}$ and the derivatives of f, that the Euler-Maruyama method is of strong order 1, i.e. (1.4) holds with $\theta = 1$.

In order to make the main idea clear cut, here are the options we have for estimating the error:

(i) If the local error e_j, at the jth time step, is bounded as

$$\mathbb{E}[|e_j|] \lesssim \Delta t_N^{3/2},$$

as usual for a 1/2-Hölder noise, then adding them up leads to

$$\sum \mathbb{E}[|e_j|] \lesssim N\Delta t_N^{3/2} = T\Delta t_N^{1/2}.$$

(ii) If we use the Itô isometry locally, we still get the local error as

$$\mathbb{E}[|e_j|] \leq \mathbb{E}[|e_j|^2]^{1/2} \lesssim \left(\Delta t_N^{2(3/2)}\right)^{1/2} = \Delta t_N^{3/2},$$

and adding that up still leads to an error of order Δt_N^{θ}.

(iii) If, instead, we first add the terms up, then $\sum e_j$ becomes an integral over $[0, T]$ with respect to the Wiener noise, so that we can use the Itô isometry on the added up term and obtain

$$\begin{aligned}\mathbb{E}\left[\left|\sum e_j\right|\right] \lesssim \left(\mathbb{E}\left[\left|\sum e_j\right|^2\right]\right)^{1/2} &= \left(\sum \mathbb{E}[|e_j|^2]\right)^{1/2} \\ &= \left(\sum \Delta t_N^3\right)^{1/2} = \left(\Delta t_N^2\right)^{1/2} = \Delta t_N.\end{aligned}$$

and we finally get the error to be of order 1.

2. Pathwise solution

For the notion and main results on pathwise solution for RODEs, we refer the reader to [2, Section 2.1]. We consider two sets of hypotheses, Standing Hypotheses 2.1 and Standing Hypotheses 2.2, each suitable to one of the two main cases we consider, namely the case in which the steps are bounded by processes with monotonic bounds and the case with Itô type noises.

We start with the following hypotheses, which imply the existence and uniqueness of pathwise solutions of the RODE (1.1) in the sense of Carathéodory.

Standing Hypotheses 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 $L_X > 0$ such that*

$$|f(t, x_1, y) - f(t, x_2, y)| \leq L_X|x_1 - x_2|, \quad \forall t \in [0, T],\ \forall x_1, x_2, y \in \mathbb{R}. \tag{2.1}$$

(ii) *We also assume that $(t, x) \mapsto f(t, x, Y_t)$ satisfies the Carathéodory conditions:*
   (a) *The mapping $x \mapsto f(t, x, y)$ is continuous on $x \in \mathbb{R}$, for almost every $(t, y) \in I \times \mathbb{R}$;*
   (b) *The mapping $t \mapsto f(t, x, Y_t)$ is Lebesgue measurable in $t \in [0, T]$, 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 each sample path $t \mapsto M_t(\omega)$ being absolutely continuous on $t \in [0, T]$.*

Under these assumptions, 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 each realization $X_0 = X_0(\omega)$ of the initial condition and each sample path $t \mapsto Y_t(\omega)$ of the noise process, where $\omega \in \Omega$.

The family of pathwise solutions, $(t, \omega) \mapsto X_t(\omega)$, is measurable (see [2, 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

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

Notice the absolute continuity of the sample path of $\{M_t\}_{t \in I}$ is not needed for this bound, but is required for the existence of sample path solutions in the Carathéodory theory (see [2, Theorem 2.3]).

Moreover, the requirement that its sample paths be absolutely continuous is not a direct requirement on the regularity of the noise. They are just bounds. For instance, if the sample paths of the noise are just continuous, then, each sample path is uniformly bounded on the compact time interval $[0, T]$ and, hence, we can bound the sample path of the noise by the constant path given by the maximum $M(\omega) = \max_{t \in [0,T]} |Y_t(\omega)|$. If the noise is a renewal process, with sample paths with jump discontinuities, we can bound its sample paths by the piecewise linear paths joining the discontinuity values. Similarly for other point processes, for transport processes, and so on.

For the strong convergence of the Euler-Maruyama approximation, we also need to control the expectation of the solution above, among other things. One common situation is when

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

in which case

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

In special *dissipative* cases, depending on the structure of the equation, we might not need the second condition and only require

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

Due to the diverse ways in which to control the solution, we leave further estimates to the applications.

When $f = f(t, x, y)$ is continuous on all three variables, as well as uniformly globally Lipscthiz continuous on x, and the sample paths of $\{Y_t\}_{t\geq 0}$ are continuous, then the integrand in (2.2) is continuous on 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 classical sense and the estimate (2.3) is still valid, provided the sample paths of $\{M_t\}_{t\in I}$ are at leat integrable.

This setting is assumed in the analysis of the case of Itô noise. With that in mind, we define, more precisely, a second set of standing hypotheses.

Standing Hypotheses 2.2. *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, in the sense of (2.1), with constant $L_X > 0$.*
(ii) *$f = f(t, x, y)$ is continuous on $I \times \mathbb{R} \times \mathbb{R}$.*
(iii) *For each $x \in \mathbb{R}$, the map $(t, y) \mapsto f(t, x, y)$ is twice continuously differentiable on $I \times \mathbb{R}$.*
(iv) *The sample paths $t \mapsto Y_t(\omega)$ of the noise process are continuous on I.*
(v) *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 each sample path $t \mapsto M_t(\omega)$ being integrable on $t \in [0, T]$.*

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 either the Standing Hypotheses 2.1 or the Standing Hypotheses 2.2, the Euler-Maruyama approximation (1.2) for any pathwise solution of the random ordinary differential equation (1.1) satisfies the global error formula*

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

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

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-Maruyama step is given by

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

Subtracting, we obtain

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

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

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

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

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

Using the jump function τ^N, we may rewrite the above expression as in (3.1). □

4. Basic estimate

Here we derive an estimate, under minimal hypotheses, that will be the basis for the estimates in specific cases.

Lemma 4.1. *Under either the Standing Hypotheses 2.1 or the Standing Hypotheses 2.2, the global error* (3.1) *is estimated as*

$$
\begin{aligned}
|X_{t_j} - X_{t_j}^N| \leq \Bigg(&|X_0 - X_0^N| + L_X \int_0^{t_j} |X_s - X_{\tau^N(s)}|\;\mathrm{d}s \\
&\left|\int_0^{t_j} \Big(f(s, X_{\tau^N(s)}^N, Y_s) - f(\tau^N(s), X_{\tau^N(s)}^N, Y_{\tau^N(s)})\Big)\;\mathrm{d}s\right| \Bigg) e^{L_X t_j}.
\end{aligned}
\tag{4.1}
$$

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 $t, s \in [0, T]$, 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 (8.2), so that

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

again for any $t, s \in [0, T]$, and, in particular, for $t = \tau^N(s)$. Hence,

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

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 (4.1). □

The first term in the right hand side of (4.1) usually vanishes since in general we take $X_0^N = X_0$, but it suffices to assume that

$$\mathbb{E}[|X_0^N - X_0|] \leq C_0 \Delta t_N, \qquad N \in \mathbb{N}, \tag{4.2}$$

for some constant C_0, which is useful for lower order approximations or for the discretization of random partial differential equations.

The third term in (4.1) is the more delicate one that will be handled differently in the next sections.

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

Lemma 4.2. *Under either the Standing Hypotheses 2.1 or the Standing Hypotheses 2.2, suppose further that, for all $0 \leq \tau \leq s \leq T$, we have*

$$|f(s, X_s, Y_s)| \leq F_t, \tag{4.3}$$

where $\{F_t\}_{t \in I}$ is a real random process. Then,

$$\int_0^{t_j} |X_s - X_{\tau^N(s)}| \,\mathrm{d}s \leq \int_0^{t_j} \int_{\tau^N(s)}^s F_\sigma \,\mathrm{d}\sigma \,\mathrm{d}s. \tag{4.4}$$

If, moreover,

$$\sup_{0\leq\tau<t\leq T}\frac{1}{t-\tau}\int_\tau^t \mathbb{E}[F_s]\,\mathrm{d}s<\infty, \tag{4.5}$$

then

$$\int_0^{t_j}\mathbb{E}[|X_s-X_{\tau^N(s)}|]\,\mathrm{d}s\leq K\Delta t_N, \tag{4.6}$$

where

$$K=T\sup_{0\leq\tau<t\leq T}\frac{1}{t-\tau}\int_\tau^t\mathbb{E}[F_s]\,\mathrm{d}s \tag{4.7}$$

Proof. Since $s-\tau^N(s)\leq\Delta t_N$ and using the control on $s\mapsto f(s,X_s,Y_s)$, we observe that

$$\begin{aligned}\left|X_s-X_{\tau^N(s)}\right| &= \left|\int_{\tau^N(s)}^s f(\sigma,X_\sigma,Y_\sigma)\,\mathrm{d}\sigma\right|\\ &\leq\int_{\tau^N(s)}^s F_\sigma\,\mathrm{d}\sigma.\end{aligned}$$

Integrating over $[0,t_j]$ and using integration by parts

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

Taking the expectation we find

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

□

Lemma 4.3. *Under either the Standing Hypotheses 2.1 or the Standing Hypotheses 2.2, suppose further that, for all $0\leq\tau\leq s\leq T$, we have*

$$|f(s,X_s,Y_s)|\leq F_t, \tag{4.8}$$

where $\{F_t\}_{t\in I}$ is a real random process. Then,

$$\int_0^{t_j}|X_s-X_{\tau^N(s)}|\,\mathrm{d}s\leq\int_0^{t_j}\int_{\tau^N(s)}^s F_\sigma\,\mathrm{d}\sigma\,\mathrm{d}s. \tag{4.9}$$

If, moreover,

$$\sup_{0\leq\tau<t\leq T}\frac{1}{t-\tau}\int_\tau^t\mathbb{E}[F_s]\,\mathrm{d}s<\infty, \tag{4.10}$$

then

$$\int_0^{t_j} \mathbb{E}[|X_s - X_{\tau^N(s)}|] \,\mathrm{d}s \leq K\Delta t_N, \tag{4.11}$$

where

$$K = T \sup_{0\leq\tau<t\leq T} \frac{1}{t-\tau}\int_\tau^t \mathbb{E}[F_s] \,\mathrm{d}s \tag{4.12}$$

Proof. Since $s - \tau^N(s) \leq \Delta t_N$ and using the control on $s \mapsto f(s, X_s, Y_s)$, we observe that

$$\begin{aligned} \left|X_s - X_{\tau^N(s)}\right| &= \left|\int_{\tau^N(s)}^s f(\sigma, X_\sigma, Y_\sigma) \,\mathrm{d}\sigma\right| \\ &\leq \int_{\tau^N(s)}^s F_\sigma \,\mathrm{d}\sigma. \end{aligned}$$

This implies (4.9).

Taking the expectation in (4.9) and using the assumption of $\{F_t\}_{t\in I}$ yields (4.11). □

Remark 4.1. A typical example that fits within the conditions of Lemma 4.3 is when $f = f(t, x, y)$ has a power law growth of the form,

$$|f(t,x,y)| \leq C(1 + |x|^a + |y|^b), \qquad \forall\,(t,x,y) \in [0,T]\times\mathbb{R}\times\mathbb{R},$$

with $a, b \geq 1$, and the corresponding moments of the solution and the noise are bounded, i.e.

$$\sup_{t\in I} \mathbb{E}[|X_t|^a] < \infty, \qquad \sup_{t\in I} \mathbb{E}[|X_t|^b] < \infty.$$

5. The case of an Itô noise

Here, we assume the noise $\{Y_t\}_{t\in I}$ is an **Itô process**, i.e. satisfying

$$\mathrm{d}Y_t = A_t \,\mathrm{d}t + B_t \,\mathrm{d}W_t, \tag{5.1}$$

where $\{W_t\}_{t\geq 0}$ is a Wiener process and $\{A_t\}_{t\in I}$ and $\{B_t\}_{t\in I}$ are stochastic processes adapted to $\{W_t\}_{t\geq 0}$. As mentioned in the Introduction, we are not solving for Y_t, otherwise we would actually have a system of stochastic differential equations. Instead, we assume it is a known process, with analytic solution to be used in the Euler-Maruyama approximation of (1.1). In theory, A_t and B_t are allowed to be originally given in terms of $\{W_t\}_{t\geq 0}$ and $\{Y_t\}_{t\geq 0}$. For example, Y_t may be a Wiener process, an Ornstein-Uhlenbeck process or a geometric Brownian process. At this point, we only assume that $\{A_t\}_{t\geq 0}$ and $\{B_t\}_{t\geq 0}$ satisfy

$$\mathbb{E}[|A_t|] \leq M_A, \quad \mathbb{E}[|B_t|] \leq M_B, \qquad \forall t \in [0,T]. \tag{5.2}$$

We also assume that $(t, y) \mapsto f(t, x, y)$ is twice continuously differentiable, for each fixed x, so the Itô formula is applicable and yields

$$\mathrm{d}f(t,x,Y_t) = \left(\partial_t f(t,x,Y_t) + A_t \partial_y f(t,x,Y_t) + \frac{B_t^2}{2}\partial_{yy} f(t,x,Y_t)\right)\,\mathrm{d}t \\ + B_t \partial_y f(t,x,Y_t)\,\mathrm{d}W_t, \quad (5.3)$$

for every fixed $x \in \mathbb{R}$.

We need to estimate the global error (4.1). The first term in the right hand side vanishes with the assumption that $X_0^N = X_0$. The second term is bounded according to Lemma 4.3. Our main concern here is in estimating the last term. For that, we have the following results.

Lemma 5.1. *Under the Standing Hypotheses 2.2, the bound*

$$\int_0^{t_j} \left(f(s, X^N_{\tau^N(s)}, Y_s) - f(\tau^N(s), X^N_{\tau^N(s)}, Y_{\tau^N(s)}) \right)\,\mathrm{d}s = I^1 + I^2, \qquad (5.4)$$

holds for each $j = 0, 1, \ldots, N$, where

$$I^1 = \int_0^{t_j} (\tau^N(s) + \Delta t_N - s) \\ \left(\partial_s f(s, X^N_{\tau^N(s)}, Y_s) + A_s \partial_y f(s, X^N_{\tau^N(s)}, Y_s) + \frac{B_s^2}{2} \partial_{yy} f(s, X^N_{\tau^N(s)}, Y_s) \right)\,\mathrm{d}s$$

and

$$I^2 = \int_0^{t_j} (\tau^N(s) + \Delta t_N - s) B_s \partial_y f(s, X^N_{\tau^N(s)}, Y_s)\,\mathrm{d}W_s.$$

Proof. We use the Itô formula on $Z_s = f(s, X_t^N, Y_s)$ and write, for any $0 \leq t < t+\tau \leq T$,

$$\int_t^{t+\tau} \left(f(s, X_t^N, Y_s) - f(t, X_t^N, Y_t) \right)\,\mathrm{d}s = \int_t^{t+\tau} \int_t^s \mathrm{d}Z_\xi\,\mathrm{d}s \\ = \int_t^{t+\tau} \int_t^s \left(\partial_\xi f(\xi, X_t^N, Y_\xi) + A_\xi \partial_y f(\xi, X_t^N, Y_\xi) + \frac{B_\xi^2}{2} \partial_{yy} f(\xi, X_t^N, Y_\xi) \right)\,\mathrm{d}\xi\,\mathrm{d}s \\ + \int_t^{t+\tau} \int_t^s B_\xi \partial_y f(\xi, X_t^N, Y_\xi)\,\mathrm{d}W_\xi\,\mathrm{d}s.$$

Using Fubini's Theorem, the first double integral is rewritten as

$$\int_t^{t+\tau} \int_\xi^{t+\tau} \left(\partial_\xi f(\xi, X_t^N, Y_\xi) + A_\xi \partial_y f(\xi, X_t^N, Y_\xi) + \frac{B_\xi^2}{2} \partial_{yy} f(\xi, X_t^N, Y_\xi) \right)\,\mathrm{d}s\,\mathrm{d}\xi \\ = \int_t^{t+\tau} (t + \tau - \xi) \left(\partial_\xi f(\xi, X_t^N, Y_\xi) + A_\xi \partial_y f(\xi, X_t^N, Y_\xi) + \frac{B_\xi^2}{2} \partial_{yy} f(\xi, X_t^N, Y_\xi) \right)\,\mathrm{d}\xi.$$

Using now Fubini's Theorem on the last integral, we find

$$\begin{aligned}\int_t^{t+\tau}\int_t^s B_\xi \partial_y f(\xi, X_t^N, Y_\xi)\;\mathrm{d}W_\xi\;\mathrm{d}s &= \int_t^{t+\tau}\int_\xi^{t+\tau} B_\xi \partial_y f(\xi, X_t^N, Y_\xi)\;\mathrm{d}s\;\mathrm{d}W_\xi \\ &= \int_t^{t+\tau}(t+\tau-\xi) B_\xi \partial_y f(\xi, X_t^N, Y_\xi)\;\mathrm{d}W_\xi.\end{aligned}$$

Thus, for $\tau = \Delta t_N$, $t = t_{j-1} = (j-1)\Delta t_N$, and $t+\tau = t_{j-1} + \Delta t_N = t_j$, we have

$$\int_{t_j}^{t_{j+1}} \Big(f(s, X_{t_j}^N, Y_s) - f(t, X_{t_j}^N, Y_t)\Big)\;\mathrm{d}s = I_j^1 + I_j^2,$$

where

$$I_j^1 = \int_{t_j}^{t_{j+1}} (t_{j+1}-\xi)\left(\partial_\xi f(\xi, X_{t_j}^N, Y_\xi) + A_\xi \partial_y f(\xi, X_{t_j}^N, Y_\xi) + \frac{B_\xi^2}{2}\partial_{yy} f(\xi, X_{t_j}^N, Y_\xi)\right)\;\mathrm{d}\xi,$$

and

$$I_j^2 = \int_{t_j}^{t_{j+1}} (t_{j+1}-\xi) B_\xi \partial_y f(\xi, X_{t_j}^N, Y_\xi)\;\mathrm{d}W_\xi.$$

Summing them up in j and writing s for ξ leads to (5.4). □

Lemma 5.2. *Under the Standing Hypotheses 2.2, suppose, moreover, that*

$$\begin{aligned}K_1 = \sup_{N\in\mathbb{N}} \int_0^T \mathbb{E}\Big[\Big|&\partial_s f(s, X_{\tau^N(s)}^N, Y_s) \\ &+ A_s \partial_y f(s, X_{\tau^N(s)}^N, Y_s) + \frac{B_s^2}{2}\;\partial_{yy} f(s, X_{\tau^N(s)}^N, Y_s)\Big|\Big]\;\mathrm{d}s < \infty\end{aligned} \tag{5.5}$$

and

$$K_2 = \sup_{N\in\mathbb{N}} \left(\int_0^T \mathbb{E}\left[\left|B_s \partial_y f(s, X_{\tau^N(s)}^N, Y_s)\right|^2\right]\;\mathrm{d}s\right)^{1/2} < \infty. \tag{5.6}$$

Then,

$$\int_0^{t_j} \mathbb{E}\Big[\Big| f(s, X_{\tau^N(s)}^N, Y_s) - f(\tau^N(s), X_{\tau^N(s)}^N, Y_{\tau^N(s)})\Big|\Big]\;\mathrm{d}s \leq C\Delta_N \tag{5.7}$$

for each $j = 0, 1, \ldots, N$, and a suitable constant $C > 0$.

Proof. Taking the expectation of the absolute value of (5.4), we see we need to estimate $\mathbb{E}[|I^1|]$ and $\mathbb{E}[|I^2|]$, defined in Lemma 5.1. The estimate of the first term is straightfoward and leads to

$$\begin{aligned}
\mathbb{E}[|I^1|] \leq \int_0^{t_j} & |\tau^N(s) + \Delta t_N - s| \\
& \left|\partial_s f(s, X^N_{\tau^N(s)}, Y_s) + A_s \partial_y f(s, X^N_{\tau^N(s)}, Y_s) + \frac{B_s^2}{2}\partial_{yy} f(s, X^N_{\tau^N(s)}, Y_s)\right| \,\mathrm{d}s \\
\leq \Delta t_N & \int_0^T \left|\partial_s f(s, X^N_{\tau^N(s)}, Y_s) + A_s \partial_y f(s, X^N_{\tau^N(s)}, Y_s) + \frac{B_s^2}{2}\partial_{yy} f(s, X^N_{\tau^N(s)}, Y_s)\right| \,\mathrm{d}s \\
\leq K_1 & \Delta t_N,
\end{aligned}$$

where K_1 is given by (5.5). For the second term, we use the Lyapunov inequality and the Itô isometry to find that

$$\begin{aligned}
\mathbb{E}[|I^2|] &\leq \sqrt{\mathbb{E}[|I^2|^2]} \\
&= \left(\int_{t_j}^{t_{j+1}} \mathbb{E}\left[\left((\tau^N(s) + \Delta t_N - \xi) B_\xi \partial_y f(\xi, X^N_{t_j}, Y_\xi)\right)^2\right] \,\mathrm{d}\xi\right)^{1/2} \\
&\leq \Delta t_N \left(\int_{t_j}^{t_{j+1}} \mathbb{E}\left[\left(B_\xi \partial_y f(\xi, X^N_{t_j}, Y_\xi)\right)^2\right] \,\mathrm{d}\xi\right)^{1/2} \\
&\leq K_2 \Delta t_N,
\end{aligned}$$

for K_2 given by (5.6). Putting the two estimates together, we find (5.7). □

Gathering the conditions, we obtain the following result.

Theorem 5.1. *Under the Standing Hypotheses 2.2, suppose, moreover, that* (4.8), (4.10), (5.5) *and* (5.6) *hold. Then, the Euler-Maruyama scheme* (1.2)-(1.3) *is of strong order 1, i.e.*

$$\max_{j=0,\ldots,N} \mathbb{E}\left[\left|X_{t_j} - X^N_{t_j}\right|\right] \leq C\Delta t_N, \qquad \forall N \in \mathbb{N}, \tag{5.8}$$

for a suitable constant $C \geq 0$

The required conditions in Theorem 5.1, especially conditions (5.5) and (5.6) of Lemma 5.2, are not readily checked. We present, below, some particular cases with more explicit bounds that imply those.

Corollary 5.1. *Yeah, corollaries are good.*

Remark 5.1. When the noise $\{Y_t\}_{t\in I}$ is a Wiener process, then $A_t = 0$ and $B_t = 1$ and the bounds (5.5) and (5.6) become

$$K_1 = \int_0^T \mathbb{E}\left[\left|\partial_s f(s, X^N_{\tau^N(s)}, Y_s) + \frac{1}{2}\partial_{yy} f(s, X^N_{\tau^N(s)}, Y_s)\right|\right] \,\mathrm{d}s < \infty \tag{5.9}$$

and

$$K_2 = \left(\int_0^{t_j} \mathbb{E}\left[\left| \partial_y f(s, X^N_{\tau^N(s)}, Y_s) \right|^2 \right] \,\mathrm{d}s \right)^{1/2} < \infty. \tag{5.10}$$

Remark 5.2. When the noise $\{Y_t\}_{t\in I}$ is a Ornstein-Uhlenbeck process, then $A_t = -\nu Y_t$ and $B_t = \sigma$, for $\nu \in \mathbb{R}$, $\sigma > 0$, and the bounds (5.5) and (5.6) become

$$K_1 = \int_0^T \mathbb{E}\left[\left| \partial_s f(s, X^N_{\tau^N(s)}, Y_s) - \nu \partial_y f(s, X^N_{\tau^N(s)}, Y_s) + \frac{\sigma^2}{2} \partial_{yy} f(s, X^N_{\tau^N(s)}, Y_s) \right| \right] \,\mathrm{d}s < \infty \tag{5.11}$$

and

$$K_2 = \left(\int_0^{t_j} \mathbb{E}\left[\left| \sigma \partial_y f(s, X^N_{\tau^N(s)}, Y_s) \right|^2 \right] \,\mathrm{d}s \right)^{1/2} < \infty. \tag{5.12}$$

Remark 5.3. When the noise $\{Y_t\}_{t\in I}$ is a geometric Brownian motion process, then $A_t = \mu Y_t$ and $B_t = \sigma Y_t$, for $\mu \in \mathbb{R}$, $\sigma > 0$, and the bounds (5.5) and (5.6) become

$$K_1 = \int_0^T \mathbb{E}\left[\left| \partial_s f(s, X^N_{\tau^N(s)}, Y_s) + \mu Y_s \partial_y f(s, X^N_{\tau^N(s)}, Y_s) + \frac{\sigma^2 Y_s^2}{2} \partial_{yy} f(s, X^N_{\tau^N(s)}, Y_s) \right| \right] \,\mathrm{d}s < \infty \tag{5.13}$$

and

$$K_2 = \left(\int_0^{t_j} \mathbb{E}\left[\left| \sigma Y_s \partial_y f(s, X^N_{\tau^N(s)}, Y_s) \right|^2 \right] \,\mathrm{d}s \right)^{1/2} < \infty. \tag{5.14}$$

6. The case of monotonic sample path bounds

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 for typical point processes, such as renewal-reward processes, Hawkes process, and such.

More precisely, we have the following result:

Lemma 6.1. *Besides the Standing Hypotheses 2.1, suppose that, for all $0 \leq \tau \leq s \leq T$,*

$$|f(s, X_\tau, Y_s) - f(\tau, X_\tau, Y_\tau)| \leq G_s - G_\tau, \tag{6.1}$$

where $\{G_t\}_{t\in I}$ is a real random process with monotonically non-decreasing sample paths. Then,

$$\left| \int_0^t \left(f(s, X^N_{\tau^N(s)}, Y_s) - f(\tau^N(s), X^N_{\tau^N(s)}, Y_{\tau^N(s)}) \right) \,\mathrm{d}s \right| \leq (G_t - G_0)\Delta t_N, \tag{6.2}$$

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

Proof. Let $N \in \mathbb{R}$. From the assumption (6.1) we have

$$|f(s, X^N_{\tau^N(s)}, Y_s) - f(\tau^N(s), X^N_{\tau^N(s)}, Y_{\tau^N(s)})| \leq G_s - G_{\tau^N(s)},$$

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

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

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,

$$\int_0^t (G_s - G_{\tau^N(s)}) \, \mathrm{d}s = \int_0^t (G_s - G_0) \, \mathrm{d}s.$$

Using the monotonicity,

$$\int_0^t (G_s - G_{\tau^N(s)}) \, \mathrm{d}s \leq \int_0^t (G_t - G_0) \, \mathrm{d}s = (G_t - G_0)t \leq (G_t - G_0)\Delta t_N.$$

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

$$\int_0^t (G_s - G_{\tau^N(s)}) \, \mathrm{d}s = \int_0^t G_s \, \mathrm{d}s - \int_0^{t_1} G_{\tau^N(s)} \, \mathrm{d}s - \int_{t_1}^t G_{\tau^N(s)} \, \mathrm{d}s$$

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

$$\begin{aligned}
\int_0^t (G_s - G_{\tau^N(s)}) \, \mathrm{d}s &\leq \int_0^t G_s \, \mathrm{d}s - \int_0^{\Delta t_N} G_0 \, \mathrm{d}s - \int_{\Delta t_N}^t G_{s-\Delta t_N} \, \mathrm{d}s \\
&= \int_0^t G_s \, \mathrm{d}s - \int_0^{\Delta t_N} G_0 \, \mathrm{d}s - \int_0^{T-\Delta t_N} G_s \, \mathrm{d}s \\
&= \int_{t-\Delta t_N}^t G_s \, \mathrm{d}s - G_0 \Delta t_N.
\end{aligned}$$

Using again the monotonicity yields

$$\int_0^t (G_s - G_{\tau^N(s)}) \, \mathrm{d}s \leq \int_{t-\Delta t_N}^t G_t \, \mathrm{d}s - G_0 \Delta t_N = (G_t - G_0)\Delta t_N.$$

Putting the estimates together proves (6.2). □

Theorem 6.1. *Under the Standing Hypotheses 2.1, suppose further that, for all* $0 \leq \tau \leq s \leq T$, *we have*

$$|f(s, X_s, Y_s)| \leq F_t, \tag{6.3}$$

and

$$|f(s, X_\tau, Y_s) - f(\tau, X_\tau, Y_\tau)| \leq G_s - G_\tau, \tag{6.4}$$

where $\{F_t\}_{t\in I}$ and $\{G_t\}_{t\in I}$ are real random process with $\{G_t\}_{t\in I}$ having monotonically non-decreasing sample paths. Assume, finally, that

$$\sup_{0\leq\tau<t\leq T}\frac{1}{t-\tau}\int_\tau^t \mathbb{E}[F_s]\;\mathrm{d}s<\infty, \qquad \mathbb{E}[(G_T-G_0)]<\infty.$$

Then, the Euler-Maruyama scheme (1.2)-(1.3) is of strong order 1, i.e.

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

for a constant $C\geq 0$ given by

$$C=\left(\frac{T}{2}\sup_{0\leq\tau<t\leq T}\frac{1}{t-\tau}\int_\tau^t\mathbb{E}[F_s]\;\mathrm{d}s+\mathbb{E}[(G_T-G_0)]\right)e^{L_X T}. \tag{6.6}$$

Proof. Under the supplied hypotheses, Lemma 4.1 applies and the global error estimate (4.1) holds. Since $X_0^N=X_0$, the first term on the right hand side vanishes and we have two terms left to estimate. The second term is handled via Lemma 4.3. For the third term, we apply Lemma 6.1 and use the estimate (6.2). Putting the estimates together, we bound the global error by

$$|X_{t_j}-X_{t_j}^N|\leq\left(\int_0^{t_j}\int_{\tau^N(s)}^s F_\sigma\;\mathrm{d}\sigma\;\mathrm{d}s+(G_t-G_0)\Delta t_N\right)e^{L_X t_j}.$$

Taking the expectation and using again Lemma 4.3, we find that

$$\begin{aligned}\mathbb{E}\left[|X_{t_j}-X_{t_j}^N|\right]&\leq\left(\frac{t_j}{2}K\Delta t_N+\mathbb{E}[(G_{t_j}-G_0)]\Delta t_N\right)e^{L_X t_j}\\&\leq\left(\frac{T}{2}K+\mathbb{E}[(G_T-G_0)]\right)e^{L_X T}\Delta t_N,\end{aligned}$$

where K is given in (4.12). This proves (6.5) with the constant C given by (6.6). $\square$

One typical case in which a bound such as that in Theorem 6.1 is possible is that of a linear equation, or, when f is *semi-separable*. More precisely, we have the following result.

Corollary 6.1. *Suppose that $f=f(t,x,y)$ is of the form*

$$f(t,x,y)=a(t,y)h(x)+b(t,y), \tag{6.7}$$

where $a=a(t,y)$, $h=h(x)$, and $b=b(t,y)$ are continous on $[0,T]\times\mathbb{R}$ and h is globally Lipschitz continous in $x\in\mathbb{R}$, uniformly in $t\in I$. Assume, further, that

$$|a(s,Y_s)-a(\tau,Y_\tau)|\leq A_s-A_\tau, \quad |b(s,Y_s)-b(\tau,Y_\tau)|\leq B_s-B_\tau, \quad |h(X_t)|\leq H_t,$$

where $\{A_t\}_{t\in I}$, $\{B_t\}_{t\in I}$ and $\{H_t\}_{t\in I}$ are nonnegative stochastic processes with monotonic non-decreasing sample paths. Suppose that

$$\mathbb{E}[(A_T-A_0)]<\infty, \quad \mathbb{E}[(B_T-B_0)]<\infty, \quad \mathbb{E}[H_t]<\infty.$$

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

$$\max_{j=0,\ldots,N} \mathbb{E}\left[\left|X_{t_j} - X_{t_j}^N\right|\right] \leq C\Delta t_N, \qquad \forall N \in \mathbb{N}, \tag{6.8}$$

for a suitable constant $C \geq 0$.

Proof. We have

$$\begin{aligned} |f(s, X_\tau, Y_s) - f(\tau, X_\tau, Y_\tau)| &\leq |a(s, Y_s) - a(\tau, Y_\tau)||h(X_\tau)| + |b(s, Y_s) - b(\tau, Y_\tau)| \\ &\leq (A_s - A_\tau)H_\tau + B_s - B_\tau \\ &\leq A_s H_s - A_\tau H_\tau + B_s - B_\tau \\ &= G_s - G_\tau, \end{aligned}$$

for $G_t = A_t H_t$. Notice $\{G_t\}_{t\geq 0}$ also has nondecreasing monotonic sample paths, and with

$$\mathbb{E}[G_T - G_0] < \infty.$$

Thus, Theorem 6.1 applies and the strong order 1 convergence holds. □

Remark 6.1. In many applications, it is possible to bound

$$f(t, x, y) \leq C(1 + |x|^a + |y|^b),$$

for suitable $a, b \geq 1$, in which case

$$f(t, X_\tau, Y_t) \leq C(1 + |X_\tau|^a + G_t^b)$$

where $G_t = \sup_{0\leq s\leq t} |Y_t|$ is monotonically nondecreasing, and we just need the bounds

$$\mathbb{E}[(|X_t|)^a] < \infty, \qquad \mathbb{E}[(\sup_{0\leq t\leq T} |Y_t|)^b] < \infty.$$

7. Applications

In this section, we describe a few explicit examples that fall into one of the cases considered above and, hence, the Euler-Maruyama method exhibits a strong order one convergence.

7.1. **Population dynamics.** Our first example is a population dynamics model modified from [2, Section 15.2],

$$\frac{\mathrm{d}X_t}{\mathrm{d}t} = Y_t X_t (r - X_t) \tag{7.1}$$

where $r > 0$ is constant and $\{Y_t\}_{t\geq}$ is a stochastic process playing the role of a random growth parameter and given by

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

where $0 < \varepsilon < 1$ and $\{O_t\}_{t\geq 0}$ is an Ornstein-Uhlenbeck process given by

$$\mathrm{d}O_t = (\theta_1 - \theta_2 O_t)\,\mathrm{d}t + \theta_3\,\mathrm{d}W_t,$$

with $\theta_1, \theta_2, \theta_3 > 0$ constant and $\{W_t\}_{t\geq 0}$ being a standard Wiener process. We do not need to approximate the coupled stochastic differential equation system for (X_t, O_t) since the Ornstein-Uhlenbeck process has an analytic solution, given by

$$O_t = \frac{\theta_1}{\theta_2} + e^{-\theta_2(t-s)}\left(O_s - \frac{\theta_1}{\theta_2}\right) + \theta_3 \int_s^t e^{-\theta_2(t-\sigma)}\,\mathrm{d}W_\sigma.$$

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 $\{\Lambda_t\}_{t\geq 0}$ itself satisfies

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

Define

$$f(t, x, y) = yx(r - x)$$

and notice that $f(t, x, y)x \geq 0$, for $x \geq 0$ and $y \geq 0$, and $f(t, x, y)x \leq 0$, for $x \geq r$ and $y \geq 0$. Hence the interval $[0, R]$ in x is positively invariant and the pathwise solutions of (7.1) are almost surely bounded as well, with

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

for all $t \geq 0$.

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

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

for $|x| \leq R$ and $0 \leq y \leq 2\lambda$.

The right hand side of (7.1) 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 (7.1) with $0 \leq X_0 \leq R$ almost surely, for some $R > r$, is equivalent to the RODE

$$\frac{\mathrm{d}X_t}{\mathrm{d}t} = \tilde{f}(t, X_t, Y_t). \tag{7.2}$$

With $\tilde{f} = \tilde{f}(t, x, y)$, the Standing Hypotheses 2.2 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.

Condition (4.8) is satisfied with

$$F_t?!?!?!$$

The Itô formula applied to $Y_t = g(O_t)$, where $g(\eta) = \lambda(1+\varepsilon\sin(\eta))$ implies $\{Y_t\}_{t\geq 0}$ is an Itô process with

$$\mathrm{d}Y_t = \left((\theta_1 - \theta_2 O_t)g'(O_t) + \frac{\theta_3^2}{2}g''(O_t)\right)\,\mathrm{d}t + \theta_3 g'(O_t)\,\mathrm{d}W_t.$$

We have

$$g'(\eta) = \lambda\varepsilon\cos(\eta), \quad g''(\eta) = -\lambda\varepsilon\sin(\eta)$$

hence both are uniformly bounded. Therefore, condition (5.5) is satisfied with

$$K_1 \leq \lambda\varepsilon\int_0^T \left((\theta_1 - \theta_2\mathbb{E}[|O_t|] + \frac{\theta_3^2}{2})\right)\,\mathrm{d}s < \infty,$$

while (5.6) is satisfied with

$$K_2 \leq \lambda\varepsilon\theta_3 T^{1/2} < \infty.$$

Hence, all the conditions of Theorem 5.1 hold and the Euler-Maruyama method is of strong order 1.

7.2. **Drug delivery.**

7.3. **Earthquake model.**

7.4. **Point-process noise.**

8. Strong order of convergence

We assume $f = f(t,x,y)$ is twice continuously differentiable with

$$L_T = \sup_{t,x,y}|\partial_t f(t,x,y)| < \infty \tag{8.1}$$

$$L_X = \sup_{t,x,y}|\partial_x f(t,x,y)| < \infty \tag{8.2}$$

$$L_Y = \sup_{t,x,y}|\partial_y f(t,x,y)| < \infty \tag{8.3}$$

$$L_{YY} = \sup_{t,x,y}|\partial_y^2 f(t,x,y)| < \infty, \tag{8.4}$$

where the suprema are taken for $(t,x,y) \in [0,T]\times\mathbb{R}\times\mathbb{R}$. The first three condition (8.1), (8.2), and (8.3) imply that f has an at most linear growth:

$$\sup_{t,x,y}|f(t,x,y)| \leq M_0 + L(|t| + |x| + |y|), \tag{8.5}$$

for suitable nonnegative constants M_0, L.

We also assume the drift and diffusion of the Itô process $\{Y_t\}_t$ are uniformly bounded,

$$M_A = \sup_\omega \sup_{t,x,y}|A_t(\omega)| < \infty, \tag{8.6}$$

$$M_B = \sup_\omega \sup_{t,x,y}|B_t(\omega)| < \infty, \tag{8.7}$$

where the suprema are taken for $t \in [0, T]$ and for samples in all sample space $\omega \in \Omega$.

8.1. **A single step.** Here we obtain an expression for a single time step which will be suitable for a proper estimate later on. For the sake of notational simplicity, we consider a single time step from a time t to a time $t + \tau$. Later on we take $t = t_{j-1}$ and $\tau = \Delta t_N$, with $t_j = t_{j-1} + \Delta t_N$.

The exact solution satisfies, for any $t, \tau \geq 0$,

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

The Euler-Maruyama 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

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

For the integral of the last pair of terms, we use the Itô formula on $Z_s = f(s, X_t^N, Y_s)$ and write

$$\begin{aligned} \int_t^{t+\tau} & \big(f(s, X_t^N, Y_s) - f(t, X_t^N, Y_t)\big) \,\mathrm{d}s = \int_t^{t+\tau} \int_t^s \mathrm{d}Z_\xi \,\mathrm{d}s \\ & = \int_t^{t+\tau} \int_t^s \left(\partial_\xi f(\xi, X_t^N, Y_\xi) + A_\xi \partial_y f(\xi, X_t^N, Y_\xi) + \frac{B_\xi^2}{2} \partial_{yy} f(\xi, X_t^N, Y_\xi)\right) \,\mathrm{d}s \,\mathrm{d}t \\ & \quad + \int_t^{t+\tau} \int_t^s B_\xi \partial_y f(\xi, X_t^N, Y_\xi) \,\mathrm{d}W_\xi \,\mathrm{d}s. \end{aligned}$$

Using Fubini's Theorem, the last integral is rewritten as

$$
\begin{aligned}
\int_t^{t+\tau}\int_t^s B_\xi \partial_y f(\xi, X_t^N, Y_\xi)\,\mathrm{d}W_\xi\,\mathrm{d}s &= \int_t^{t+\tau}\int_\xi^{t+\tau} B_\xi \partial_y f(\xi, X_t^N, Y_\xi)\,\mathrm{d}s\,\mathrm{d}W_\xi \\
&= \int_t^{t+\tau}(t+\tau-\xi) B_\xi \partial_y f(\xi, X_t^N, Y_\xi)\,\mathrm{d}W_\xi. \quad (8.8)
\end{aligned}
$$

We rearrange these terms and write, for $\tau = \Delta t_N$ and $t = t_{j-1} = (j-1)\Delta t_N$,

$$
X_{t_j} - X_{t_j}^N = X_{t_{j-1}} - X_{t_{j-1}}^N + I_{j-1}^1 + I_{j-1}^2 + I_{j-1}^3, \quad (8.9)
$$

where

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

$$
\begin{aligned}
I_j^2 = &\int_{t_j}^{t_{j+1}} \left(f(s, X_s, Y_s) - f(s, X_{t_j}, Y_s) \right)\,\mathrm{d}s \\
&+ \int_{t_j}^{t_{j+1}}\int_{t_j}^s \left(\partial_\xi f(\xi, X_{t_j}^N, Y_\xi) + A_\xi \partial_y f(\xi, X_{t_j}^N, Y_\xi) + \frac{B_\xi^2}{2}\partial_{yy} f(\xi, X_{t_j}^N, Y_\xi) \right)\,\mathrm{d}t,
\end{aligned}
$$

and

$$
I_j^3 = \int_{t_j}^{t_{j+1}} (t_{j+1} - \xi) B_\xi \partial_y f(\xi, X_{t_j}^N, Y_\xi)\,\mathrm{d}W_\xi.
$$

8.2. **Local estimates.** The term I_j^1 is estimated using that $f = f(t, x, y)$ is globally Lipschitz in x, so that

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

Hence,

$$
\begin{aligned}
\left| \int_{t_j}^{t_{j+1}} \left(f(s, X_{t_j}, Y_s) - f(s, X_{t_j}^N, Y_s) \right)\,\mathrm{d}s \right| &\leq \int_{t_j}^{t_{j+1}} \left| f(s, X_{t_j}, Y_s) - f(s, X_{t_j}^N, Y_s) \right|\,\mathrm{d}s \\
&\leq L_X |X_{t_j} - X_{t_j}^N| \Delta t_N.
\end{aligned}
$$

This means

$$
\left| I_j^1 \right| \leq L_X |X_{t_j} - X_{t_j}^N| \Delta t_N. \quad (8.10)
$$

For I_j^2, the first term is estimated as

$$
|f(s, X_s, Y_s) - f(s, X_t, Y_s)| \leq L_X |X_s - X_t| \leq L_X \int_t^s |f(\sigma, X_\sigma, Y_\sigma)|\,\mathrm{d}\sigma \leq L_X M_f (s-t).
$$

This yields, upon integration,

$$\begin{aligned}\left|\int_{t_j}^{t_{j+1}} \left(f(s,X_s,Y_s) - f(s,X_{t_j},Y_s)\right)\;\mathrm{d}s\right| &\leq \int_{t_j}^{t_{j+1}} \left|f(s,X_s,Y_s) - f(s,X_{t_j},Y_s)\right|\;\mathrm{d}s \\ &\leq \frac{L_X M_f}{2}\Delta t_N^2.\end{aligned}$$

The double integral is estimated as

$$\begin{aligned}&\left|\int_{t_j}^{t_{j+1}}\int_{\xi}^{t_{j+1}} \left(\partial_\xi f(\xi, X_{t_j}^N, Y_\xi) + A_\xi \partial_y f(\xi, X_{t_j}^N, Y_\xi) + \frac{B_\xi^2}{2}\partial_{yy} f(\xi, X_{t_j}^N, Y_\xi)\right)\;\mathrm{d}t\right| \\ &\qquad \leq \int_{t_j}^{t_{j+1}}\int_{\xi}^{t_{j+1}} \left(L_T + M_A L_Y + \frac{M_B^2}{2}L_{YY}\right)\;\mathrm{d}t \\ &\qquad = \frac{1}{2}\tau^2\left(L_T + M_A L_Y + \frac{M_B^2}{2}L_{YY}\right). \end{aligned} \tag{8.11}$$

Hence,

$$\left|I_j^2\right| \leq M\Delta t_N^2, \tag{8.12}$$

where

$$M = \frac{1}{2}\left(L_X M_f + L_T + M_A L_Y + \frac{M_B^2}{2}L_{YY}\right).$$

Remark 8.1. Notice that, at this point, we did not estimate the last integral, otherwise we are not able to obtain the strong order 1 estimate, only 1/2. Indeed, if we use Fubini and the Itô isometry in the last integral, we find

$$\begin{aligned}&\mathbb{E}\left[\left(\int_t^{t+\tau}\int_t^s B_\xi \partial_y f(\xi, X_t^N, Y_\xi)\;\mathrm{d}W_\xi\;\mathrm{d}s\right)^2\right] = \mathbb{E}\left[\left(\int_t^{t+\tau}\int_\xi^{t+\tau} B_\xi \partial_y f(\xi, X_t^N, Y_\xi)\;\mathrm{d}s\;\mathrm{d}W_\xi\right)^2\right] \\ &= \int_t^{t+\tau}\mathbb{E}\left[\left(\int_\xi^{t+\tau} B_\xi \partial_y f(\xi, X_t^N, Y_\xi)\;\mathrm{d}s\right)^2\right]\;\mathrm{d}\xi \leq \int_t^{t+\tau}\left(\int_\xi^{t+\tau} M_B^2 L_Y\;\mathrm{d}s\right)^2\;\mathrm{d}\xi \\ &\leq \int_t^{t+\tau} M_B^2 L_Y (t+\tau-\xi)^2\;\mathrm{d}\xi = -\frac{1}{3}M_B^2 L_Y^2 (t+\tau-\xi)^3\Big]_t^{t+\tau} = \frac{1}{3}M_B^2 L_Y^2\tau^3,\end{aligned}$$

so that

$$\sqrt{\mathbb{E}\left[\left(\int_t^{t+\tau}\int_t^s B_\xi \partial_y f(\xi, X_t^N, Y_\xi)\;\mathrm{d}W_\xi\;\mathrm{d}s\right)^2\right]} \leq \frac{\sqrt{3}}{3}M_B L_Y \tau^{3/2}. \tag{8.13}$$

After adding up n times, we end up with a $\tau^{1/2}$ estimate, which is not sufficient.

8.3. **Iterating the steps.** Iterating (8.9) and assuming that $X_0^N = X_0$, we find

$$X_{t_j} - X_{t_j}^N = \sum_{i=0}^{j-1} I_j^1 + \sum_{i=0}^{j-1} I_j^2 + \sum_{i=0}^{j-1} I_j^3. \tag{8.14}$$

We estimate the first moment as

$$\mathbb{E}\left[|X_{t_j} - X_{t_j}^N|\right] \leq \sum_{i=0}^{j-1}\mathbb{E}\left[|I_j^1|\right] + \sum_{i=0}^{j-1}\mathbb{E}\left[|I_j^2|\right] + \mathbb{E}\left[\left|\sum_{i=0}^{j-1} I_j^3\right|\right]. \tag{8.15}$$

Using (8.10), (8.12), and (**??**), we obtain

$$\mathbb{E}\left[|X_{t_j} - X_{t_j}^N|\right] \leq L_X \sum_{i=0}^{j-1} \mathbb{E}\left[|X_{t_j} - X_{t_j}^N|\right] \Delta t_N + \sum_{i=0}^{j-1} C\Delta t_N^2 + M_B L_Y t_j \Delta t_N$$
$$\leq L_X \sum_{i=0}^{j-1} \mathbb{E}\left[|X_{t_i} - X_{t_i}^N|\right] \Delta t_N + C_T \Delta t_N, \quad (8.16)$$

where

$$C_T = M + M_B L_y T^{1/2}.$$

Now, we show by induction that

$$\mathbb{E}\left[|X_{t_j} - X_{t_j}^N|\right] \leq C_T e^{L_X t_j} \Delta t_N.$$

This is trivially true for $j = 0$. Now suppose it is true up to $j-1$. It follows from (8.16) that

$$\mathbb{E}\left[|X_{t_j} - X_{t_j}^N|\right] \leq L_X \sum_{i=0}^{j-1} C_T \Delta t_N e^{L_X t_i} \Delta t_N + C_T \Delta t_N = C_T \Delta t_N \left(1 + L_X \Delta t_N \sum_{i=0}^{j-1} e^{L_X t_i}\right).$$

Using that $1 + r \leq e^r$, with $r = L_X \Delta t_N$ and $t_i + \Delta t_N = t_{i+1}$, we see that

$$L_X \Delta t_N \leq e^{L_X \Delta t_N} - 1,$$

which telescopes the sum and yields

$$\mathbb{E}\left[|X_{t_j} - X_{t_j}^N|\right] \leq C_T \Delta t_N \left(1 + (e^{L_X \Delta t_N} - 1) \sum_{i=0}^{j-1} e^{L_X t_i}\right) = C_T \Delta t_N \left(1 + (e^{L_X j \Delta t_N} - 1)\right).$$

Hence,

$$\mathbb{E}\left[|X_{t_j} - X_{t_j}^N|\right] \leq C_T e^{L_X t_j} \Delta t_N,$$

which completes the induction. Hence, we have proved the following result.

Theorem 8.1. *Consider the initial value problem (1.1), on a time interval $[0, T]$, with $T > 0$, and assume the noise is given by (5.1), with (8.6) and (8.7). Suppose $f = f(t, x, y)$ is twice continuously differentiable, with (8.5)-(8.4). Let $\{X_t\}_{t \geq 0}$ be the solution of (1.1). Let $N \in \mathbb{N}$ and let $\{X_{t_j}^N\}_{j=0,\ldots,N}$ be the solution of the Euler-Maruyama method (1.2)-(1.3). Then,*

$$\mathbb{E}\left[\left|X_{t_j} - X_{t_j}^N\right|\right] \leq C_T e^{L_X t_j} \Delta t_N, \qquad j = 0, \ldots, N, \ \forall N \in \mathbb{N}, \Delta t_N = \frac{T}{N}, \quad (8.17)$$

where

$$C_T = \frac{1}{2}\left(L_X M_f + L_T + M_A L_Y + \frac{M_B^2}{2} L_{YY}\right) + M_B L_y T^{1/2}. \quad (8.18)$$

We end this section by abstracting away the Gronwall type inequality we use (this is probably written somewhere, and I need to find the source):

Lemma 8.1. *Let $(e_j)_j$ be a (finite or infinite) sequence of positive numbers satisfying*

$$e_j \leq a \sum_{i=0}^{j-1} e_i + b, \tag{8.19}$$

with $e_0 = 0$, where $a, b > 0$. Then,

$$e_j \leq be^{aj}, \qquad \forall j. \tag{8.20}$$

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

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

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

$$e_j \leq be^{ja},$$

which completes the induction. □

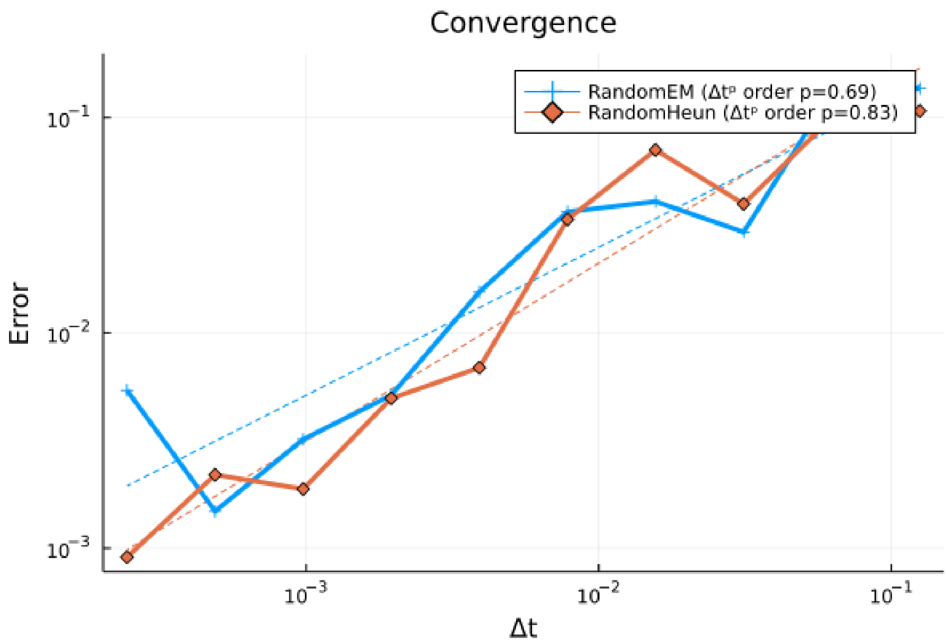
9. Numerical examples

9.1. **Lower-order converge.** For a lower order convergence, below order 1, we take the noise $\{Y_t\}_t$ to be the transport process defined by

$$Y_t = \sin(t/Z)^{1/3},$$

where Z is a beta random variable $Z \sim B(\alpha, \beta)$. Notice Z takes values strictly within $(0, 1)$ and, hence, $\sin(t/Z)$ can have arbitrarily high frequencies and, hence, go through the critic value $y = 0$ extremely often.

(Need to remove the Heun method and do more tests).



10. Estimate on the solution

We assume $f = f(t, x, y)$ is continuous on all variables and is Lipschitz continuous on each variable, i.e. there exist constants $L_T, L_X, L_Y \geq 0$ such that

$$|f(t_1, x, y) - f(t_2, x, y)| \leq L_T|t_1 - t_2|, \tag{10.1}$$

$$|f(t, x_1, y) - f(t, x_2, y)| \leq L_X|x_1 - x_2|, \tag{10.2}$$

$$|f(t, x, y_1) - f(t, x, y_2)| \leq L_Y|y_1 - y_2|, \tag{10.3}$$

for all $t, t_1, t_2 \in I$, $x, x_1, x_2 \in \mathbb{R}$ and $y, y_1, y_2 \in \mathbb{R}$. By the continuity of $f = f(t, x, y)$, we also have

$$M_0 = \sup_{t\in I} |f(t, 0, 0)| < \infty.$$

These conditions imply that f has an at most linear growth in x and y:

$$|f(t, x, y)| \leq M_0 + L_X|x| + L_Y|y|, \tag{10.4}$$

for every $(t, x, y) \in [0, T] \times \mathbb{R} \times \mathbb{R}$.

We assume the initial condition has a bounded first moment:

$$\mathbb{E}[|X_0|] \leq C_0 < \infty. \tag{10.5}$$

As for the noise, we assume, for now, that

$$\mathbb{E}[|Y_t|] \leq M_Y, \qquad \forall t \in [0, T]. \tag{10.6}$$

With the assumed regularity on $f = f(t, x, y)$, the solutions of (1.1) are pathwise solutions, so that

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

Using (10.4), we estimate each solution with

$$|X_t| \leq |X_0| + \int_0^t (M_0 + L_X|X_s| + L_Y|Y_s|)\;\mathrm{d}s.$$

Using Gronwall's lemma, we find

$$|X_t| \leq \left(|X_0| + M_0 t + L_Y\int_0^t |Y_s|\;\mathrm{d}s\right)e^{L_X t}, \quad t\in[0,T]. \tag{10.7}$$

In particular, taking the expectation,

$$\mathbb{E}[|X_t|] \leq \left(\mathbb{E}[|X_0|] + M_0 t + L_Y\int_0^t \mathbb{E}[|Y_s|]\;\mathrm{d}s\right)e^{L_X t}, \quad t\in[0,T].$$

Using hypotheses (10.5) and (10.6), we find that

$$\mathbb{E}[|X_t|] \leq (C_0 + (M_0 + L_Y M_Y)t)\,e^{L_X t}, \quad t\in[0,T].$$

hence,

$$\mathbb{E}[|X_t|] \leq M_X, \qquad t\in[0,T], \tag{10.8}$$

with

$$M_X = (C_0 + (M_0 + L_Y M_Y)T)e^{L_X T}. \tag{10.9}$$

Similarly, we write, for $t \geq t_0 > 0$,

$$X_t - X_{t_0} = \int_{t_0}^t f(s, X_s, Y_s)\;\mathrm{d}s.$$

Using (10.4), we estimate

$$\begin{aligned} |X_t - X_{t_0}| &\leq \int_{t_0}^t (M_0 + L_X|X_s| + L_Y|Y_s|)\;\mathrm{d}s \\ &\leq L_X\int_{t_0}^t |X_s|\;\mathrm{d}s + L_Y\int_{t_0}^t |Y_s|\;\mathrm{d}s + M_0(t-t_0). \end{aligned}$$

Using (10.7), we obtain

$$|X_t - X_{t_0}| \leq L_X\int_{t_0}^t \left(|X_0| + M_0 s + L_Y\int_0^s |Y_\sigma|\;\mathrm{d}\sigma\right)e^{L_X s}\;\mathrm{d}s + L_Y\int_{t_0}^t |Y_s|\;\mathrm{d}s + M_0(t-t_0) \tag{10.10}$$

Appendix

The heart of the matter is the following. Think of τ as the time-step Δt_N, but we use τ for simplicity. Let $g = g(t)$ be a θ-Hölder continuous function, with Hölder constant C. Then, we can do the usual "local"-type estimate

$$\begin{aligned}
\left|\int_0^T (g(t+\tau) - g(t))\ \mathrm{d}t\right| &\leq \int_0^T |g(t+\tau) - g(t)|\ \mathrm{d}t \\
&\leq C\int_0^T \tau^\theta\ \mathrm{d}t \\
&= C\tau^\theta T,
\end{aligned}$$

which yields an order θ approximation, with respect to the "time step" τ. However, we can also integrate first, so that

$$\begin{aligned}
\left|\int_0^T (g(t+\tau) - g(t))\ \mathrm{d}t\right| &= \left|\int_0^T g(t+\tau)\ \mathrm{d}t - \int_0^T g(t)\ \mathrm{d}t\right| \\
&= \left|\int_\tau^{T+\tau} g(t)\ \mathrm{d}t - \int_0^T g(t)\ \mathrm{d}t\right| \\
&= \left|\int_T^{T+\tau} g(t)\ \mathrm{d}t - \int_0^\tau g(t)\ \mathrm{d}t\right| \\
&\leq 2\max_t |g(t)|\tau,
\end{aligned}$$

which reveals the order 1 convergence, even without assuming that g is Hölder.

For the discretization, however, we don't have $g(t+\tau) - g(t)$, but actually steps $g(t) - g(\tau^N(t))$, where $\tau^N(t)$ picks the largest $j\tau$ smaller than or equal to t, i.e. $\tau^N(t) = \max\{j\tau;\ j\tau \leq t, j\}$. And there is also the dependency on the solution X_t itself, leading to the steps $f(t, X_{\tau^N(t)}, Y_t) - f(\tau^N(t), X_{\tau^N(t)}, Y_{\tau^N(t)})$. The idea, then, is to assume that these steps can be bound by

$$\begin{aligned}
&|f(t, X_{\tau^N(t)}, Y_t) - f(\tau^N(t), X_{\tau^N(t)}, Y_{\tau^N(t)})| \\
&\qquad\qquad\qquad \leq (G_t - G_{\tau^N(t)})h(X_{\tau^N(t)}) + G_t^0 - G_{\tau^N(t)}^0,
\end{aligned}$$

where the bounding process G_t (usually $G_t = g(t, Y_t)$ for some $g = g(t, y)$, but not necessarily) is assumed to have monotone nondecreasing sample paths. In this case, an estimate similar to the above can be obtained, and the strong order 1 convergence, achieved.

Keep in mind that assuming that $g(t)$ is the difference of monotone functions, then g is differentiable almost everywhere, but that is not quite the same as saying that it is Lipschitz, not even absolutely continuous nor of bounded variation. Think of that classical example that g is constant almost everywhere, hence $g' = 0$ almost everywhere, and $\int_0^1 g'(s)\ \mathrm{d}s = 0$, but $g(1) > g(0)$. In fact, there is an important case that falls into this category which is the renewal-reward process, that has jump discontinuities and each sample path can be written as the difference between two monotonically nondecreasing jump functions. More general point-process such as the Hawkes process used, e.g. in earthquake models should also work. These are great examples!

Acknowledgments

References

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

[2] X. Han & P. E. Kloeden (2017), *Random Ordinary Differential Equations and Their Numerical Solution,* Probability Theory and Stochastic Modelling, vol. 85, Springer Singapore. DOI: 10.1007/978-981-10-6265-0.

(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