Received: 30 September 2024 | Revised: 20 June 2025 | Accepted: 19 August 2025

DOI: 10.1112/jlms.70297

RESEARCH ARTICLE

Journal of the London Mathematical Society

Pathwise convergence of the Euler scheme for rough and stochastic differential equations

Andrew L. Allan[1] | **Anna P. Kwossek**[2] | **Chong Liu**[3] | **David J. Prömel**[4]

[1]Durham University, Durham, UK

[2]University of Vienna, Vienna, Austria

[3]ShanghaiTech University, Shanghai, China

[4]University of Mannheim, Mannheim, Germany

Correspondence
Andrew L. Allan, Department of Mathematical Sciences, Durham University, Mathematical Sciences and Computer Science Building, Upper Mountjoy Campus, Stockton Road, Durham, DH1 3LE, UK.
Email: andrew.l.allan@durham.ac.uk

Funding information
Baden-Württemberg Stiftung, Germany

Abstract

The convergence of the first-order Euler scheme and an approximative variant thereof, along with convergence rates, are established for rough differential equations driven by càdlàg paths satisfying a suitable criterion, namely the so-called Property (RIE), along time discretizations with vanishing mesh size. This property is then verified for almost all sample paths of Brownian motion, Itô processes, Lévy processes, and general càdlàg semimartingales, as well as the driving signals of both mixed and rough stochastic differential equations, relative to various time discretizations. Consequently, we obtain pathwise convergence in p-variation of the Euler–Maruyama scheme for stochastic differential equations driven by these processes.

MSC 2020

65C30, 65L20, 60L20 (primary), 60H35 (secondary)

1 | INTRODUCTION

Stochastic differential equations (SDEs) serve as models for dynamical systems that evolve randomly in time, and are fundamental mathematical objects, essential to numerous applications in finance, engineering, biology, and beyond. In a fairly general form, an SDE is given by

$$Y_t = y_0 + \int_0^t b(s, Y_s)\,\mathrm{d}s + \int_0^t \sigma(s, Y_s)\,\mathrm{d}X_s, \qquad t \in [0,T], \tag{1.1}$$

where $y_0 \in \mathbb{R}^k$ is the initial condition, $b: [0,T] \times \mathbb{R}^k \to \mathbb{R}^k$ and $\sigma: [0,T] \times \mathbb{R}^k \to \mathbb{R}^{k\times d}$ are coefficients, and the driving signal $X = (X_t)_{t\in[0,T]}$ is a d-dimensional stochastic process that models the random noise affecting the system.

Assuming that X is a càdlàg semimartingale, such as a Brownian motion or a Lévy process, and the coefficients b, σ are suitably regular, it is well known that (1.1) is well posed as an Itô SDE. That is, $\int_0^t \sigma(s, Y_s)\,\mathrm{d}X_s$ can be defined as a stochastic Itô integral, and the equation admits a unique adapted solution $Y = (Y_t)_{t\in[0,T]}$; see, for example, [36]. Unfortunately, such SDEs, including many of those which appear in practical applications, can rarely be solved explicitly, which has led to a vast literature on various numerical approximations of the solutions to SDEs; see, for example, [29].

One of the most common approaches to numerically approximate the solution of an SDE is to rely on a time-discretized modification of the equation. This type of discretization is implemented in particular by the Euler scheme (also called the Euler–Maruyama scheme) and its higher order variants. For the SDE (1.1), the (first-order) Euler approximation is defined by

$$Y_t^n = y_0 + \sum_{i\,:\,t_{i+1}^n \leqslant t} b(t_i^n, Y_{t_i^n}^n)(t_{i+1}^n - t_i^n) + \sum_{i\,:\,t_{i+1}^n \leqslant t} \sigma(t_i^n, Y_{t_i^n}^n)(X_{t_{i+1}^n} - X_{t_i^n}), \tag{1.2}$$

for $t \in [0,T]$, along a sequence of partitions $\mathcal{P}^n = \{0 = t_0^n < t_1^n < \cdots < t_{N_n}^n = T\}$. Higher order Euler approximations, such as the Milstein scheme, introduce additional higher order correction terms in the approximation (1.2), which often involve iterated integrals of the driving signal X. In general, the numerical calculation of the approximation Y^n is carried out path by path, motivating a pathwise convergence analysis of the Euler scheme and its higher order variants. Indeed, it is well known that, for SDEs driven by Brownian motion, the (higher-order) Euler approximations converge pathwise; see, for example, [7, 23, 27, 28, 39].

A fully pathwise solution theory for SDEs like (1.1) is provided by the theory of rough paths; see, for example, [17, 21]. Loosely speaking, in our context, a rough path is pair $\mathbf{X} = (X, \mathbb{X})$, consisting of a deterministic càdlàg $\mathbb{R}^d$-valued path X, and a two-parameter càdlàg $\mathbb{R}^{d\times d}$-valued function $\mathbb{X}$, which satisfy certain analytic and algebraic conditions. We will work with càdlàg rough paths with finite p-variation, in the regime with $p \in (2,3)$, which includes in particular almost any sample path of a general semimartingale X, in which case the corresponding rough path $\mathbf{X} = (X, \mathbb{X})$ is given by $\mathbb{X}_{s,t} = \int_s^t (X_{r-} - X_s) \otimes \mathrm{d}X_r$ via stochastic integration.

Replacing the stochastic driving signal X in (1.1) by a (deterministic) rough path $\mathbf{X} = (X, \mathbb{X})$, we obtain a so-called rough differential equation (RDE). Assuming sufficient regularity of the coefficients b, σ, the RDE (1.1) driven by a given càdlàg rough path $\mathbf{X} = (X, \mathbb{X})$ is well posed, in the sense that $\int_0^t \sigma(s, Y_s)\,\mathrm{d}\mathbf{X}_s$ is defined as a rough integral, and the equation admits a unique solution $Y = (Y_t)_{t\in[0,T]}$; see [22]. Moreover, if the rough path is, say, the Itô lift of a semimartingale X, then the solution of the resulting random RDE is consistent with the solution of the corresponding SDE driven by X. Both interpretations of the equation are thus essentially equivalent. Furthermore, in contrast to classical SDE theory, rough path theory is not limited to the semimartingale setting, and it comes with powerful pathwise stability estimates.

Rough path theory is intrinsically linked to the numerical approximation of SDEs — see, for example, [4, 12] — and provides a transparent explanation for the pathwise convergence of higher order Euler approximations and their modifications; see, for example, [14, 20–22, 32]. More precisely, the existence of a rough path lift of the driving signal is a sufficient condition for the pathwise convergence of higher order Euler schemes for RDEs, thus implying pathwise convergence for the corresponding SDEs driven by, for example, semimartingales. However, the pathwise convergence of the first-order Euler scheme — the most prominent numerical scheme for differential equations — cannot be explained by the rough path lift of the driving signal. Moreover, in general, an Euler approximation cannot converge to the solution of an RDE driven by an arbitrary rough path, for at least two reasons: First, the Euler approximation for an SDE driven by a fractional Brownian motion with Hurst index $H < \frac{1}{2}$ fails to converge (see, e.g., [14]), and second, while the rough path lift $\mathbf{X} = (X, \mathbb{X})$ of a path X is not unique, leading to potentially multiple solutions of the RDE, the Euler approximation Y^n as defined in (1.2) is independent of the choice of rough path, and can thus only converge to at most one such solution.

In the present paper, we clarify the gap between rough and SDEs from the perspective of numerical approximation, by establishing the convergence of the first-order Euler scheme for RDEs driven by Itô-type rough path lifts. More precisely, in Theorem 2.3, we obtain convergence in p-variation of the Euler scheme for RDEs driven by càdlàg paths satisfying a suitable criterion — namely the so-called Property (RIE) — relative to a sequence of partitions with vanishing mesh size.

Property (RIE) was first introduced in [35] and [2], motivated by applications in mathematical finance under model uncertainty. While, strictly speaking, it is a condition on a càdlàg path $X : [0,T] \to \mathbb{R}^d$, it always ensures the existence of an Itô-type rough path lift $\mathbf{X} = (X, \mathbb{X})$, allowing one to treat (1.1) as an RDE. Using this fact, we will show that Property (RIE) is a sufficient condition on the sample paths of a stochastic driving signal to guarantee the convergence of the first-order Euler scheme for the corresponding SDE. We note, in particular, that the Euler scheme converges surely on the set where the stochastic driving signal satisfies Property (RIE), which is a stronger statement compared to the earlier results in [7, 23, 27, 28, 39], in which the set on which the Euler scheme converges can depend on the coefficients b, σ. A criterion for Hölder continuous rough paths, related to Property (RIE), was previously introduced by Davie [12], which also allows one to obtain convergence of the Euler scheme for RDEs, and will be discussed in more detail in Remark 2.4.

Exploiting the continuity results of rough path theory, in Theorem 2.3, we derive a precise error estimate in p-variation for the Euler approximation of RDEs with respect to the discretization error of the driving signal. The convergence rate is expressed transparently, in terms of the mesh size of the approximating partition, and the approximation error of the discretized signal and of its rough path lift. We also obtain an error estimate for the Euler approximation with respect to pathwise perturbations of the driving signal; see Proposition 2.13. This latter perturbation is motivated by so-called approximate Euler schemes for SDEs driven by jump processes; see, for example, [13, 25, 38]. For instance, approximate Euler schemes are used for Lévy-driven SDEs, since the increments of Lévy processes cannot always be simulated, and thus the increments of the driving Lévy process need to be approximated by random variables with known distributions.

To obtain pathwise convergence of the Euler scheme in p-variation for an SDE, it is then sufficient to verify that the associated stochastic driving signal of the equation satisfies Property (RIE), almost surely, relative to a sequence of partitions; see Sections 3 and 4. Unsurprisingly, we find that the more regular the driving signal is, the more general the sequence of partitions may be chosen. Indeed, while the sample paths of a Brownian motion satisfy Property (RIE), almost surely, relative to sequences of partitions whose mesh size can converge to zero very slowly, the sample paths of more general Itô processes satisfy Property (RIE), almost surely, relative to sequences of partitions whose mesh size is of order 2^{-n}. For stochastic processes with jumps, such as Lévy processes or general càdlàg semimartingales, one needs to ensure that the jump times are exhausted by the sequence of partitions, which is a necessary condition, for both the Euler scheme to converge pathwise, and for Property (RIE) to be satisfied by the driving signal.

The presented pathwise analysis of the first-order Euler approximation is not limited to SDEs in a semimartingale setting. As examples, we consider mixed SDEs driven by both Brownian motion and fractional Brownian motion with Hurst index $H > \frac{1}{2}$, as in, for example, [33, 41], as well as rough SDEs, which are differential equations driven by both a rough path and a Brownian motion; see [18]. The latter equations are of interest, for example, in the context of robust stochastic filtering; see [11, 15].

Organization of the paper: In Section 2, we prove the convergence of the Euler scheme for RDEs assuming that the driving paths satisfy Property (RIE). In Sections 3 and 4, we provide various examples of stochastic processes that satisfy Property (RIE) along suitable sequences of partitions, making the established convergence analysis applicable to the corresponding SDEs, and derive associated convergence rates.

2 | THE EULER SCHEME FOR ROUGH DIFFERENTIAL EQUATIONS

In this section, we study convergence of the (first-order) Euler scheme for RDEs, which does not rely on the Lévy area of the path, and is known to converge pathwise for certain classes of SDEs. Before treating the Euler scheme, we will first recall some essentials from the theory of càdlàg rough paths, as introduced in [19, 22].

2.1 | Essentials on rough path theory

A *partition* $\mathcal{P}$ of an interval $[s,t]$ is a finite set of points between and including the points s and t, that is, $\mathcal{P} = \{s = u_0 < u_1 < \cdots < u_N = t\}$ for some $N \in \mathbb{N}$, and its mesh size is denoted by $|\mathcal{P}| := \max\{|u_{i+1} - u_i| \,:\, i = 0, \dots, N-1\}$. A sequence $(\mathcal{P}^n)_{n\in\mathbb{N}}$ of partitions is said to be *nested*, if $\mathcal{P}^n \subset \mathcal{P}^{n+1}$ for all $n \in \mathbb{N}$.

Throughout, we let $T > 0$ be a fixed finite time horizon. We let $\Delta_T := \{(s,t) \in [0,T]^2 \,:\, s \leqslant t\}$ denote the standard 2-simplex. A function $w : \Delta_T \to [0,\infty)$ is called a *control function* if it is superadditive, in the sense that $w(s,u) + w(u,t) \leqslant w(s,t)$ for all $0 \leqslant s \leqslant u \leqslant t \leqslant T$. For two vectors $x = (x^1, \dots, x^d), y = (y^1, \dots, y^d) \in \mathbb{R}^d$, we use the usual tensor product

$$x \otimes y := (x^i y^j)_{i,j=1,\dots,d} \in \mathbb{R}^{d\times d}.$$

Whenever $(B, \|\cdot\|)$ is a normed space and $f, g : B \to \mathbb{R}$ are two functions on B, we shall write $f \lesssim g$ or $f \leqslant Cg$ to mean that there exists a constant $C > 0$ such that $f(x) \leqslant Cg(x)$ for all $x \in B$. The constant C may depend on the normed space, for example, through its dimension or regularity parameters.

The space of linear maps from $\mathbb{R}^d \to \mathbb{R}^n$ is denoted by $\mathcal{L}(\mathbb{R}^d;\mathbb{R}^n)$, and we write, for example, $C_b^k(\mathbb{R}^m;\mathcal{L}(\mathbb{R}^d;\mathbb{R}^n))$ for the space of k-times differentiable (in the Fréchet sense) functions $f : \mathbb{R}^m \to \mathcal{L}(\mathbb{R}^d;\mathbb{R}^n)$ such that f and all its derivatives up to order k are continuous and bounded. We equip this space with the norm

$$\|f\|_{C_b^k} := \|f\|_\infty + \|\mathrm{D}f\|_\infty + \cdots + \|\mathrm{D}^k f\|_\infty,$$

where $\mathrm{D}^r f$ denotes the rth order derivative of f, and $\|\cdot\|_\infty$ denotes the supremum norm on the corresponding spaces of operators.

For a normed space $(E, |\cdot|)$, we let $D([0,T];E)$ denote the set of càdlàg (right continuous with left limits) paths from $[0,T]$ to E. For $X \in D([0,T];E)$, the supremum norm of the path X is given by

$$\|X\|_\infty := \sup_{t\in[0,T]} |X_t|,$$

and, for $p \geqslant 1$, the p-variation of the path X is given by

$$\|X\|_p := \|X\|_{p,[0,T]} \qquad \text{with} \qquad \|X\|_{p,[s,t]} := \left(\sup_{\mathcal{P}\subset[s,t]} \sum_{[u,v]\in\mathcal{P}} |X_v - X_u|^p \right)^{\frac{1}{p}}, \quad (s,t) \in \Delta_T,$$

where the supremum is taken over all possible partitions $\mathcal{P}$ of the interval $[s,t]$. We recall that, given a path X, we have that $\|X\|_p < \infty$ if and only if there exists a control function w such that†

$$\sup_{(u,v)\in\Delta_T} \frac{|X_v - X_u|^p}{w(u,v)} < \infty.$$

We write $D^p = D^p([0,T];E)$ for the space of paths $X \in D([0,T];E)$ that satisfy $\|X\|_p < \infty$. Moreover, for a path $X \in D([0,T];\mathbb{R}^d)$, we will often use the shorthand notation:

$$X_{s,t} := X_t - X_s \quad \text{and} \quad X_{t-} := \lim_{u\nearrow t} X_u, \qquad \text{for} \quad (s,t) \in \Delta_T.$$

For $r \geqslant 1$ and a two-parameter function $\mathbb{X} : \Delta_T \to E$, we similarly define

$$\|\mathbb{X}\|_r := \|\mathbb{X}\|_{r,[0,T]} \qquad \text{with} \qquad \|\mathbb{X}\|_{r,[s,t]} := \left(\sup_{\mathcal{P}\subset[s,t]} \sum_{[u,v]\in\mathcal{P}} |\mathbb{X}_{u,v}|^r \right)^{\frac{1}{r}}, \quad (s,t) \in \Delta_T.$$

We write $D_2^r = D_2^r(\Delta_T;E)$ for the space of all functions $\mathbb{X} : \Delta_T \to E$ that satisfy $\|\mathbb{X}\|_r < \infty$, and are such that the maps $s \mapsto \mathbb{X}_{s,t}$ for fixed t, and $t \mapsto \mathbb{X}_{s,t}$ for fixed s, are both càdlàg.

For $p \in [2,3)$, a pair $\mathbf{X} = (X, \mathbb{X})$ is called a *càdlàg p-rough path* over $\mathbb{R}^d$ if

(i) $X \in D^p([0,T];\mathbb{R}^d)$ and $\mathbb{X} \in D_2^{\frac{p}{2}}(\Delta_T;\mathbb{R}^{d\times d})$, and
(ii) Chen's relation: $\mathbb{X}_{s,t} = \mathbb{X}_{s,u} + \mathbb{X}_{u,t} + X_{s,u} \otimes X_{u,t}$ holds for all $0 \leqslant s \leqslant u \leqslant t \leqslant T$.

In component form, condition (ii) states that $\mathbb{X}^{ij}_{s,t} = \mathbb{X}^{ij}_{s,u} + \mathbb{X}^{ij}_{u,t} + X^i_{s,u}X^j_{u,t}$ for every i and j. We will denote the space of càdlàg p-rough paths by $\mathcal{D}^p = \mathcal{D}^p([0,T];\mathbb{R}^d)$. On the space $\mathcal{D}^p([0,T];\mathbb{R}^d)$, we use the natural seminorm

$$\|\mathbf{X}\|_p := \|\mathbf{X}\|_{p,[0,T]} \qquad \text{with} \qquad \|\mathbf{X}\|_{p,[s,t]} := \|X\|_{p,[s,t]} + \|\mathbb{X}\|_{\frac{p}{2},[s,t]}$$

for $(s,t) \in \Delta_T$, and the induced distance

$$\|\mathbf{X};\widetilde{\mathbf{X}}\|_p := \|\mathbf{X};\widetilde{\mathbf{X}}\|_{p,[0,T]} \qquad \text{with} \qquad \|\mathbf{X};\widetilde{\mathbf{X}}\|_{p,[s,t]} := \|X - \widetilde{X}\|_{p,[s,t]} + \|\mathbb{X} - \widetilde{\mathbb{X}}\|_{\frac{p}{2},[s,t]}, \tag{2.1}$$

whenever $\mathbf{X} = (X, \mathbb{X}), \widetilde{\mathbf{X}} = (\widetilde{X}, \widetilde{\mathbb{X}}) \in \mathcal{D}^p([0,T];\mathbb{R}^d)$.

† Here and throughout, we adopt the convention that $\frac{0}{0} := 0$.

Let $p \in [2,3)$, $q \in [p,\infty)$ and $r \in [\frac{p}{2},2)$ such that $\frac{1}{p}+\frac{1}{r}>1$ and $\frac{1}{p}+\frac{1}{q}=\frac{1}{r}$. Let $X \in D^p([0,T];\mathbb{R}^d)$. We say that a pair (Y,Y') is a *controlled path* (with respect to X), if

$$Y \in D^p([0,T];E), \quad Y' \in D^q([0,T];\mathcal{L}(\mathbb{R}^d;E)), \quad \text{and} \quad R^Y \in D_2^r(\Delta_T;E),$$

where R^Y is defined by

$$Y_{s,t} = Y'_s X_{s,t} + R^Y_{s,t} \qquad \text{for all} \quad (s,t) \in \Delta_T.$$

We write $\mathcal{V}_X^{q,r} = \mathcal{V}_X^{q,r}([0,T];E)$ for the space of E-valued controlled paths, which becomes a Banach space when equipped with the norm

$$(Y,Y') \mapsto |Y_0| + |Y'_0| + \|Y'\|_{q,[0,T]} + \|R^Y\|_{r,[0,T]}.$$

Remark 2.1. The definition of a controlled path adopted here is slightly more general than the classical definition in, for example, [22], in which one takes $q=p$ and $r=\frac{p}{2}$. Allowing these regularity parameters to take larger values allows us to consider slightly more general integrands in rough integrals. In particular, this is convenient in Theorem 2.2 below, as otherwise we would require further restrictions on the regularity of the paths A and H therein.

For paths $A \in D^{q_1}$, $H \in D^{q_2}$ for $q_1,q_2 \in [1,2)$, and a rough path $\mathbf{X} \in \mathcal{D}^p$ for $p \in [2,3)$, we consider the RDE:

$$Y_t = y_0 + \int_0^t b(H_s,Y_s)\,\mathrm{d}A_s + \int_0^t \sigma(H_s,Y_s)\,\mathrm{d}\mathbf{X}_s, \qquad t \in [0,T]. \tag{2.2}$$

Provided that $\frac{1}{p}+\frac{1}{q_1}>1$ and $\frac{1}{p}+\frac{1}{q_2}>1$, the first integral in this equation can be defined as a Young integral, while the second integral is defined as a rough integral. For precise definitions, constructions, and properties of these integrals, we refer to the comprehensive exposition in [22].

Theorem 2.2. *Let $p \in [2,3)$ and $q_1,q_2 \in [1,2)$ such that $\frac{1}{p}+\frac{1}{q_1}>1$ and $\frac{1}{p}+\frac{1}{q_2}>1$. Let $b \in C_b^2(\mathbb{R}^{m+k};\mathcal{L}(\mathbb{R}^n;\mathbb{R}^k))$, $\sigma \in C_b^3(\mathbb{R}^{m+k};\mathcal{L}(\mathbb{R}^d;\mathbb{R}^k))$, $y_0 \in \mathbb{R}^k$, $A \in D^{q_1}([0,T];\mathbb{R}^n)$, $H \in D^{q_2}([0,T];\mathbb{R}^m)$, and $\mathbf{X}=(X,\mathbb{X}) \in \mathcal{D}^p([0,T];\mathbb{R}^d)$. Let $r \in [\frac{p}{2} \vee q_1 \vee q_2, 2)$ such that $\frac{1}{p}+\frac{1}{r}>1$, and let $q \in [p,\infty)$ such that $\frac{1}{p}+\frac{1}{q}=\frac{1}{r}$. Then there exists a unique path $Y \in D^p([0,T];\mathbb{R}^k)$ such that the controlled path $(Y,\sigma(H,Y)) \in \mathcal{V}_X^{q,r}$ satisfies the RDE (2.2).*

Moreover, if $\widetilde{y}_0 \in \mathbb{R}^k$, $\widetilde{A} \in D^{q_1}$, $\widetilde{H} \in D^{q_2}$ and $\widetilde{\mathbf{X}}=(\widetilde{X},\widetilde{\mathbb{X}}) \in \mathcal{D}^p$ with corresponding solution $\widetilde{Y}$, and if $\|A\|_r, \|\widetilde{A}\|_r, \|H\|_r, \|\widetilde{H}\|_r, \|\mathbf{X}\|_p, \|\widetilde{\mathbf{X}}\|_p \leqslant L$ for some $L>0$, then

$$\begin{aligned}&\|Y-\widetilde{Y}\|_p + \|Y'-\widetilde{Y}'\|_q + \|R^Y - R^{\widetilde{Y}}\|_r \\ &\quad \lesssim |y_0-\widetilde{y}_0| + |H_0-\widetilde{H}_0| + \|H-\widetilde{H}\|_r + \|A-\widetilde{A}\|_r + \|\mathbf{X};\widetilde{\mathbf{X}}\|_p,\end{aligned} \tag{2.3}$$

where the implicit multiplicative constant depends only on p,q,r, $\|b\|_{C_b^2}$, $\|\sigma\|_{C_b^3}$, and L.

The result of Theorem 2.2 may be considered classical, and will be unsurprising to readers familiar with RDEs. However, to the best of our knowledge, a proof of the precise statement of the theorem does not appear in the existing literature. A sketch of the proof, based on the proof of [1, Theorem 2.3], is therefore given in Appendix A.

2.2 | Convergence of the Euler scheme

Let us consider the RDE

$$Y_t = y_0 + \int_0^t b(s, Y_s)\,\mathrm{d}s + \int_0^t \sigma(s, Y_s)\,\mathrm{d}\mathbf{X}_s, \qquad t \in [0,T], \tag{2.4}$$

where $y_0 \in \mathbb{R}^k$, $b \in C_b^2(\mathbb{R}^{k+1};\mathbb{R}^k)$, $\sigma \in C_b^3(\mathbb{R}^{k+1};\mathcal{L}(\mathbb{R}^d;\mathbb{R}^k))$ and $\mathbf{X} = (X, \mathbb{X}) \in \mathcal{D}^p([0,T];\mathbb{R}^d)$ is the driving càdlàg p-rough path for $p \in [2,3)$. Given a sequence of partitions $\mathcal{P}^n = \{0 = t_0^n < t_1^n < \cdots < t_{N_n}^n = T\}$, $n \in \mathbb{N}$, the Euler approximation Y^n corresponding to the RDE (2.4) along the partition $\mathcal{P}^n$ is given by

$$Y_t^n = y_0 + \sum_{i\,:\,t_{i+1}^n \leqslant t} b(t_i^n, Y_{t_i^n}^n)(t_{i+1}^n - t_i^n) + \sum_{i\,:\,t_{i+1}^n \leqslant t} \sigma(t_i^n, Y_{t_i^n}^n)(X_{t_{i+1}^n} - X_{t_i^n}), \tag{2.5}$$

for $t \in [0,T]$.

It is a classical result in the numerical analysis of SDEs that, if the driving signal is, for example, a Brownian motion, then the Euler scheme (often also called the Euler–Maruyama scheme) converges pathwise; see, for example, [28]. On the other hand, it is known that in general the Euler scheme cannot converge if the driving signal is an arbitrary rough path, since the corresponding Euler scheme for SDEs driven by fractional Brownian motion fails to converge; see [14] for a more detailed discussion on this observation.

Moreover, since the extension of a path X to a rough path $\mathbf{X} = (X, \mathbb{X})$ is not unique, and the Euler approximation Y^n defined in (2.5) is independent of $\mathbb{X}$, the sequence $(Y^n)_{n\in\mathbb{N}}$ cannot converge to the solution of a general RDE. Thus, in order to ensure the convergence of the Euler scheme, it is necessary to identify the "correct" rough path lift $\mathbf{X}$ as the driving signal for the RDE (2.4). A suitable resolution to this is provided by the so-called Property (RIE), as introduced in [35] and [2].

Property (RIE). Let $p \in (2,3)$ and let $\mathcal{P}^n = \{0 = t_0^n < t_1^n < \cdots < t_{N_n}^n = T\}$, $n \in \mathbb{N}$, be a sequence of partitions of the interval $[0,T]$ such that $|\mathcal{P}^n| \to 0$ as $n \to \infty$. For $X \in D([0,T];\mathbb{R}^d)$, and each $n \in \mathbb{N}$, we define $X^n : [0,T] \to \mathbb{R}^d$ by

$$X_t^n = X_T \mathbf{1}_{\{T\}}(t) + \sum_{k=0}^{N_n - 1} X_{t_k^n} \mathbf{1}_{[t_k^n, t_{k+1}^n)}(t), \qquad t \in [0,T].$$

We assume that

(i) the sequence of paths $(X^n)_{n\in\mathbb{N}}$ converges uniformly to X as $n \to \infty$,
(ii) the Riemann sums

$$\int_0^t X_u^n \otimes \mathrm{d}X_u := \sum_{k=0}^{N_n - 1} X_{t_k^n} \otimes X_{t_k^n \wedge t, t_{k+1}^n \wedge t}$$

converge uniformly as $n \to \infty$ to a limit, which we denote by $\int_0^t X_u \otimes \mathrm{d}X_u$, $t \in [0,T]$, and

(iii) there exists a control function w such that

$$\sup_{(s,t)\in\Delta_T}\frac{|X_{s,t}|^p}{w(s,t)}+\sup_{n\in\mathbb{N}}\sup_{0\leqslant k<\ell\leqslant N_n}\frac{|\int_{t_k^n}^{t_\ell^n}X_u^n\otimes \mathrm{d}X_u - X_{t_k^n}\otimes X_{t_k^n,t_\ell^n}|^{\frac{p}{2}}}{w(t_k^n,t_\ell^n)}\leqslant 1. \tag{2.6}$$

We say that a path $X\in D([0,T];\mathbb{R}^d)$ satisfies Property (RIE) relative to p and $(\mathcal{P}^n)_{n\in\mathbb{N}}$, if p, $(\mathcal{P}^n)_{n\in\mathbb{N}}$ and X together satisfy Property (RIE).

It is known that, if a path $X\in D([0,T];\mathbb{R}^d)$ satisfies Property (RIE), then X extends canonically to a rough path $\mathbf{X}=(X,\mathbb{X})\in\mathcal{D}^p([0,T];\mathbb{R}^d)$, where the lift $\mathbb{X}$ is defined by

$$\mathbb{X}_{s,t}:=\int_s^t X_u\otimes \mathrm{d}X_u - X_s\otimes(X_t-X_s),\qquad (s,t)\in\Delta_T, \tag{2.7}$$

with $\int_s^t X_u\otimes \mathrm{d}X_u:=\int_0^t X_u\otimes \mathrm{d}X_u-\int_0^s X_u\otimes \mathrm{d}X_u$, and the existence of the integral $\int_0^t X_u\otimes \mathrm{d}X_u$ is ensured by condition (ii) of Property (RIE); see [2, Lemma 2.13]. When assuming Property (RIE) for a path X, we will always work with the rough path $\mathbf{X}=(X,\mathbb{X})$ defined via (2.7), and note that $\mathbf{X}=(X,\mathbb{X})$ corresponds to the Itô rough path lift of a stochastic process, since the "iterated integral" $\mathbb{X}$ is given as a limit of left-point Riemann sums, analogously to the stochastic Itô integral.

Postulating Property (RIE) for the driving signal of an RDE ensures that the (first-order) Euler approximation converges to the solution of the equation, as stated precisely in the next theorem.

Theorem 2.3. *Suppose that $X:[0,T]\to\mathbb{R}^d$ satisfies Property (RIE) relative to some $p\in(2,3)$ and a sequence of partitions $(\mathcal{P}^n)_{n\in\mathbb{N}}$, and let $\mathbf{X}$ be the canonical rough path lift of X, as defined in (2.7). Let Y be the solution to the RDE (2.4) driven by $\mathbf{X}$, and let Y^n be the Euler approximation defined in (2.5). Then,*

$$\|Y^n-Y\|_{p'}\longrightarrow 0\qquad \text{as}\quad n\longrightarrow\infty,$$

for any $p'\in(p,3)$, and the rate of convergence is determined by the estimate

$$\|Y^n-Y\|_{p'}\lesssim|\mathcal{P}^n|^{1-\frac{1}{q}}+\|X^n-X\|_\infty^{1-\frac{p}{p'}}+\left\|\int_0^\cdot X_u^n\otimes \mathrm{d}X_u-\int_0^\cdot X_u\otimes \mathrm{d}X_u\right\|_\infty^{1-\frac{p}{p'}}, \tag{2.8}$$

which holds for any $q\in(1,2)$ such that $\frac{1}{p'}+\frac{1}{q}>1$, where the implicit multiplicative constant depends only on $p,p',q,\|b\|_{C_b^2},\|\sigma\|_{C_b^3},T,|X_0|$ and $w(0,T)$, where w is the control function for which (2.6) holds.

Note that Property (RIE) implies that each of the terms on the right-hand side of (2.8) tends to zero as $n\to\infty$.

Remark 2.4. In [12], A. M. Davie observed that, under suitable conditions, the first-order Euler scheme along equidistant partitions converges to the solution of a given RDE. More precisely, for $p\in(2,3)$ and $\alpha:=\frac{1}{p}$, let $\mathbf{X}=(X,\mathbb{X})$ be an α-Hölder continuous rough path, so that $|X_{s,t}|\lesssim|t-s|^\alpha$ and $|\mathbb{X}_{s,t}|\lesssim|t-s|^{2\alpha}$ for $(s,t)\in\Delta_T$, such that, for some $\beta\in(1-\alpha,2\alpha)$, there exists a constant $C > 0$ such that

$$\left| \sum_{j=k}^{\ell-1} \mathbb{X}_{jh,(j+1)h} \right| \leqslant C(\ell - k)^{\beta} h^{2\alpha}$$

whenever $h > 0$ and $0 \leqslant k < \ell$ are integers such that $\ell h \leqslant T$. Under this condition on the driving signal $\mathbf{X}$, [12, Theorem 7.1] states that the Euler approximations Y^n, as defined in (2.5), converge uniformly to the solution Y of the RDE (2.4) along the equidistant partitions $(\mathcal{P}^n_{\mathrm{U}})_{n\in\mathbb{N}}$, where $\mathcal{P}^n_U = \{\frac{iT}{n} : i = 0, 1, \dots, n\}$. Note that Davie's condition implies Property (RIE) — see [35, Appendix B] — and is thus less general, even in the case of Hölder continuous rough paths.

Remark 2.5. Since the "iterated integrals" appearing in the definition of a rough path (and in, for example, higher order Euler schemes) are often numerically difficult to simulate, various approaches have been developed to avoid the direct involvement of iterated integrals in the approximation of stochastic and RDEs. For instance, [14] introduced a simplified Milstein scheme for SDEs driven by fractional Brownian motion, where the iterated integrals are replaced by products of the increments of the driving process. Using this idea, simplified Runge–Kutta methods for differential equations driven by general (continuous) rough paths were investigated in [37]; see also [24].

The rest of this subsection is devoted to the proof of Theorem 2.3, which first requires us to establish some auxiliary results.

In the following, we will always assume that $X : [0, T] \to \mathbb{R}^d$ satisfies Property (RIE) relative to some $p \in (2, 3)$ and a sequence of partitions $(\mathcal{P}^n)_{n\in\mathbb{N}}$. As the piecewise constant approximation X^n (as defined in Property (RIE)) has finite 1-variation, it possesses a canonical rough path lift $\mathbf{X}^n = (X^n, \mathbb{X}^n) \in \mathcal{D}^p([0, T]; \mathbb{R}^d)$, with $\mathbb{X}^n$ given by

$$\mathbb{X}^n_{s,t} := \int_s^t X^n_{s,u} \otimes \mathrm{d}X^n_u, \qquad (s, t) \in \Delta_T, \tag{2.9}$$

where the integral is defined as a classical limit of left-point Riemann sums. Note that, while [22, Section 5.3] discretizes the rough path $\mathbf{X} = (X, \mathbb{X})$ in a piecewise constant manner, here we instead discretize the path X and then extend it to a rough path $\mathbf{X}^n = (X^n, \mathbb{X}^n)$ via (2.9).

As a first step toward the proof of Theorem 2.3, we establish the convergence of the rough paths $(\mathbf{X}^n)_{n\in\mathbb{N}}$ to the rough path $\mathbf{X}$ in a suitable rough path distance. For this purpose, we need two auxiliary lemmas.

Lemma 2.6. *Suppose that $X : [0, T] \to \mathbb{R}^d$ satisfies Property (RIE) relative to some $p \in (2, 3)$ and a sequence of partitions $(\mathcal{P}^n)_{n\in\mathbb{N}}$. Then, we have the estimate*

$$\sup_{(s,t)\in\Delta_T} |\mathbb{X}^n_{s,t} - \mathbb{X}_{s,t}| \leqslant 2\|X\|_\infty \|X^n - X\|_\infty + \sup_{(s,t)\in\Delta_T} \left| \int_s^t X^n_{s,u} \otimes \mathrm{d}X_u - \mathbb{X}_{s,t} \right|,$$

where $\mathbb{X}^n$ and $\mathbb{X}$ were defined in (2.9) and (2.7), respectively. In particular, we have that

$$\mathbb{X}^n \longrightarrow \mathbb{X} \quad \textit{uniformly as} \quad n \longrightarrow \infty.$$

Proof. Since

$$|\mathbb{X}^n_{s,t} - \mathbb{X}_{s,t}| \leqslant \left|\mathbb{X}^n_{s,t} - \int_s^t X^n_{s,u} \otimes \mathrm{d}X_u\right| + \left|\int_s^t X^n_{s,u} \otimes \mathrm{d}X_u - \mathbb{X}_{s,t}\right|,$$

and the limit in condition (ii) of Property (RIE) holds uniformly, it is enough to prove that the function given by

$$\Lambda^n_{s,t} := \mathbb{X}^n_{s,t} - \int_s^t X^n_{s,u} \otimes \mathrm{d}X_u = \int_s^t X^n_{s,u} \otimes \mathrm{d}(X^n - X)_u$$

satisfies

$$\sup_{(s,t)\in\Delta_T} |\Lambda^n_{s,t}| \leqslant 2\|X\|_\infty \|X^n - X\|_\infty. \tag{2.10}$$

If $t^n_k \leqslant s < t \leqslant t^n_{k+1}$ for some k, then $X^n_{s,u} = X_{t^n_k, t^n_k} = 0$ for every $u \in [s,t)$, so that $\Lambda^n_{s,t} = 0$. Otherwise, let k_0 be the smallest k such that $t^n_k \in (s,t)$, and let k_1 be the largest such k. It is straightforward to see that the triplet $(X^n - X, X^n, \Lambda^n)$ satisfies Chen's relation:

$$\Lambda^n_{s,t} = \Lambda^n_{s,u} + \Lambda^n_{u,t} + X^n_{s,u} \otimes (X^n - X)_{u,t}$$

for all $s \leqslant u \leqslant t$, from which it follows that

$$\Lambda^n_{s,t} = \Lambda^n_{s,t^n_{k_0}} + \Lambda^n_{t^n_{k_0},t^n_{k_1}} + \Lambda^n_{t^n_{k_1},t} + X^n_{s,t^n_{k_0}} \otimes (X^n - X)_{t^n_{k_0},t^n_{k_1}} + X^n_{s,t^n_{k_1}} \otimes (X^n - X)_{t^n_{k_1},t}.$$

As we already observed, we have that $\Lambda^n_{s,t^n_{k_0}} = \Lambda^n_{t^n_{k_1},t} = 0$. In fact, we also have that

$$\begin{aligned}\Lambda^n_{t^n_{k_0},t^n_{k_1}} &= \int_{t^n_{k_0}}^{t^n_{k_1}} X^n_{t^n_{k_0},u} \otimes \mathrm{d}(X^n - X)_u = \sum_{i=k_0}^{k_1-1} \int_{t^n_i}^{t^n_{i+1}} X^n_{t^n_{k_0},u} \otimes \mathrm{d}(X^n - X)_u \\ &= \sum_{i=k_0}^{k_1-1} \int_{t^n_i}^{t^n_{i+1}} X_{t^n_{k_0},t^n_i} \otimes \mathrm{d}(X^n - X)_u = \sum_{i=k_0}^{k_1-1} X_{t^n_{k_0},t^n_i} \otimes (X^n - X)_{t^n_i,t^n_{i+1}} = 0.\end{aligned} \tag{2.11}$$

Since $(X^n - X)_{t^n_{k_0}} = (X^n - X)_{t^n_{k_1}} = 0$, we simply obtain $\Lambda^n_{s,t} = X^n_{s,t^n_{k_1}} \otimes (X^n_t - X_t)$, from which (2.10) follows. □

Lemma 2.7. *Suppose that $X : [0,T] \to \mathbb{R}^d$ satisfies Property (RIE) relative to some $p \in (2,3)$ and a sequence of partitions $(\mathcal{P}^n)_{n\in\mathbb{N}}$. Let w be the control function with respect to which X satisfies the inequality (2.6). Then, there exists a constant C, which depends only on p, such that*

$$\|\mathbb{X}^n\|_{\frac{p}{2}} \leqslant C w(0,T)^{\frac{2}{p}} \tag{2.12}$$

for every $n \in \mathbb{N}$, where $\mathbb{X}^n$ was defined in (2.9).

Proof. Let $n \in \mathbb{N}$, and let $(s,t) \in \Delta_T$. If $t^n_k \leqslant s < t \leqslant t^n_{k+1}$ for some k, then $X^n_{s,u} = X_{t^n_k,t^n_k} = 0$ for every $u \in [s,t)$, so that $\mathbb{X}^n_{s,t} = 0$. Otherwise, let k_0 be the smallest k such that $t^n_k \in (s,t)$, and let k_1 be the largest such k. It is straightforward to see that $(X^n, \mathbb{X}^n)$ satisfies Chen's relation:

$$\mathbb{X}^n_{s,t} = \mathbb{X}^n_{s,u} + \mathbb{X}^n_{u,t} + X^n_{s,u} \otimes X^n_{u,t}$$

for all $s \leqslant u \leqslant t$, from which it follows that

$$\mathbb{X}^n_{s,t} = \mathbb{X}^n_{s,t^n_{k_0}} + \mathbb{X}^n_{t^n_{k_0},t^n_{k_1}} + \mathbb{X}^n_{t^n_{k_1},t} + X^n_{s,t^n_{k_0}} \otimes X^n_{t^n_{k_0},t^n_{k_1}} + X^n_{s,t^n_{k_1}} \otimes X^n_{t^n_{k_1},t}.$$

As we have already seen, we have that $\mathbb{X}^n_{s,t^n_{k_0}} = \mathbb{X}^n_{t^n_{k_1},t} = 0$. Recalling the calculation in (2.11), we note that

$$\mathbb{X}^n_{t^n_{k_0},t^n_{k_1}} = \int_{t^n_{k_0}}^{t^n_{k_1}} X^n_{t^n_{k_0},u} \otimes \mathrm{d}X^n_u = \int_{t^n_{k_0}}^{t^n_{k_1}} X^n_{t^n_{k_0},u} \otimes \mathrm{d}X_u,$$

and hence, by the inequality in (2.6), that

$$\left|\mathbb{X}^n_{t^n_{k_0},t^n_{k_1}}\right|^{\frac{p}{2}} = \left|\int_{t^n_{k_0}}^{t^n_{k_1}} X^n_{t^n_{k_0},u} \otimes \mathrm{d}X_u\right|^{\frac{p}{2}} \leqslant w\left(t^n_{k_0}, t^n_{k_1}\right) \leqslant w\left(t^n_{k_0-1}, t^n_{k_1+1}\right).$$

We estimate the remaining terms as

$$\begin{aligned}
&\left|X^n_{s,t^n_{k_0}} \otimes X^n_{t^n_{k_0},t^n_{k_1}}\right|^{\frac{p}{2}} + \left|X^n_{s,t^n_{k_1}} \otimes X^n_{t^n_{k_1},t}\right|^{\frac{p}{2}} \lesssim \left|X^n_{s,t^n_{k_0}}\right|^p + \left|X^n_{t^n_{k_0},t^n_{k_1}}\right|^p + \left|X^n_{s,t^n_{k_1}}\right|^p + \left|X^n_{t^n_{k_1},t}\right|^p \\
&\leqslant \left|X_{t^n_{k_0-1},t^n_{k_0}}\right|^p + \left|X_{t^n_{k_0},t^n_{k_1}}\right|^p + \left|X_{t^n_{k_0-1},t^n_{k_1}}\right|^p + \left|X_{t^n_{k_1},t^n_{k_1+1}}\right|^p \\
&\leqslant w\left(t^n_{k_0-1}, t^n_{k_0}\right) + w\left(t^n_{k_0}, t^n_{k_1}\right) + w\left(t^n_{k_0-1}, t^n_{k_1}\right) + w\left(t^n_{k_1}, t^n_{k_1+1}\right) \\
&\leqslant 2w\left(t^n_{k_0-1}, t^n_{k_1+1}\right).
\end{aligned}$$

Putting this together, we have that

$$|\mathbb{X}^n_{s,t}|^{\frac{p}{2}} \leqslant \widetilde{C} w\left(t^n_{k_0-1}, t^n_{k_1+1}\right)$$

for some constant $\widetilde{C}$. It follows that, for an arbitrary partition $\mathcal{P}$ of the interval $[0,T]$, we have the bound

$$\sum_{[s,t]\in\mathcal{P}} |\mathbb{X}^n_{s,t}|^{\frac{p}{2}} \leqslant 3\widetilde{C} w(0,T),$$

and hence that (2.12) holds with $C = (3\widetilde{C})^{\frac{2}{p}}$. ☐

Using the previous two lemmas, we can now infer the convergence of the rough paths $(\mathbf{X}^n)_{n\in\mathbb{N}}$ to the rough path $\mathbf{X}$.

Lemma 2.8. *Suppose that $X : [0,T] \to \mathbb{R}^d$ satisfies Property (RIE) relative to some $p \in (2,3)$ and a sequence of partitions $(\mathcal{P}^n)_{n\in\mathbb{N}}$. Let $\mathbf{X} = (X, \mathbb{X})$ and $\mathbf{X}^n = (X^n, \mathbb{X}^n)$ be the càdlàg rough paths defined via (2.7) and (2.9), respectively. Then, for any $p' > p$, we have that*

$$\|\mathbf{X}^n; \mathbf{X}\|_{p'} \longrightarrow 0 \qquad \textit{as} \quad n \longrightarrow \infty, \tag{2.13}$$

with a rate of convergence given by

$$\|\mathbf{X}^n; \mathbf{X}\|_{p'} \lesssim \|X^n - X\|_\infty^{1-\frac{p}{p'}} + \sup_{(s,t)\in\Delta_T} \left| \int_s^t X^n_{s,u} \otimes \mathrm{d}X_u - \mathbb{X}_{s,t} \right|^{1-\frac{p}{p'}}, \tag{2.14}$$

where the implicit multiplicative constant depends only on $p, p', |X_0|$ and $w(0,T)$, where w is the control function for which (2.6) holds.

Proof. By a standard interpolation estimate (e.g., [21, Proposition 5.5]), it follows, for any $p' > p$, that

$$\|X^n - X\|_{p'} \leqslant \|X^n - X\|_p^{\frac{p}{p'}} \|X^n - X\|_\infty^{1-\frac{p}{p'}}.$$

We similarly have that

$$\|\mathbb{X}^n - \mathbb{X}\|_{\frac{p'}{2}} \leqslant \|\mathbb{X}^n - \mathbb{X}\|_{\frac{p}{2}}^{\frac{p}{p'}} \sup_{(s,t)\in\Delta_T} |\mathbb{X}^n_{s,t} - \mathbb{X}_{s,t}|^{1-\frac{p}{p'}}.$$

We recall from Lemma 2.6 that

$$\sup_{(s,t)\in\Delta_T} |\mathbb{X}^n_{s,t} - \mathbb{X}_{s,t}| \leqslant 2\|X\|_\infty \|X^n - X\|_\infty + \sup_{(s,t)\in\Delta_T} \left| \int_s^t X^n_{s,u} \otimes \mathrm{d}X_u - \mathbb{X}_{s,t} \right|.$$

We have that $\sup_{n\in\mathbb{N}} \|X^n\|_p \leqslant \|X\|_p$ and $\|X\|_\infty \leqslant |X_0| + \|X\|_p \leqslant |X_0| + w(0,T)^{\frac{1}{p}}$, and, by the lower semicontinuity of the $\frac{p}{2}$-variation norm and Lemma 2.7, $\|\mathbb{X}\|_{\frac{p}{2}} \leqslant \liminf_{n\to\infty} \|\mathbb{X}^n\|_{\frac{p}{2}} \leqslant \sup_{n\in\mathbb{N}} \|\mathbb{X}^n\|_{\frac{p}{2}} \leqslant Cw(0,T)^{\frac{2}{p}}$. Putting this together, we conclude that (2.14) holds. By conditions (i) and (ii) in Property (RIE), the convergence in (2.13) then also follows. □

As a next step toward the proof of Theorem 2.3, we introduce a discretized version of the RDE (2.4). For this purpose, we define a time discretization path along $\mathcal{P}^n$ by

$$\gamma^n_t := T\mathbf{1}_{\{T\}}(t) + \sum_{k=0}^{N_n-1} t^n_k \mathbf{1}_{[t^n_k, t^n_{k+1})}(t), \qquad t \in [0,T], \tag{2.15}$$

and consider the RDE

$$\widetilde{Y}^n_t = y_0 + \int_0^t b(\gamma^n_s, \widetilde{Y}^n_s)\,\mathrm{d}\gamma^n_s + \int_0^t \sigma(\gamma^n_s, \widetilde{Y}^n_s)\,\mathrm{d}\mathbf{X}^n_s, \qquad t \in [0,T]. \tag{2.16}$$

Thanks to Lemma 2.8 and the local Lipschitz continuity of the Itô–Lyons map, we obtain the following proposition.

Proposition 2.9. *Suppose that $X : [0,T] \to \mathbb{R}^d$ satisfies Property (RIE) relative to some $p \in (2,3)$ and a sequence of partitions $(\mathcal{P}^n)_{n\in\mathbb{N}}$. Let Y be the solution of the RDE (2.4), and let $\widetilde{Y}^n$ be the solution of the RDE (2.16). Then,*

$$\|\widetilde{Y}^n - Y\|_{p'} \longrightarrow 0 \qquad \text{as} \quad n \longrightarrow \infty, \tag{2.17}$$

for any $p' \in (p,3)$, with a rate of convergence given by

$$\|\widetilde{Y}^n - Y\|_{p'} \lesssim |\mathcal{P}^n|^{1-\frac{1}{q}} + \|X^n - X\|_\infty^{1-\frac{p}{p'}} + \left\| \int_0^\cdot X^n_u \otimes \mathrm{d}X_u - \int_0^\cdot X_u \otimes \mathrm{d}X_u \right\|_\infty^{1-\frac{p}{p'}},$$

for any $q \in (1,2)$ such that $\frac{1}{p'} + \frac{1}{q} > 1$, where the implicit multiplicative constant depends only on $p, p', q, \|b\|_{C^2_b}, \|\sigma\|_{C^3_b}, T, |X_0|$ and $w(0,T)$, where w is the control function for which (2.6) holds.

Proof. Setting $\gamma_t := t$ for $t \in [0,T]$, the RDE (2.4) may be rewritten as

$$Y_t = y_0 + \int_0^t b(\gamma_s, Y_s)\,\mathrm{d}\gamma_s + \int_0^t \sigma(\gamma_s, Y_s)\,\mathrm{d}\mathbf{X}_s, \qquad t \in [0,T].$$

Hence, by Theorem 2.2, we know that

$$\|\widetilde{Y}^n - Y\|_{p'} \lesssim \|\gamma^n - \gamma\|_q + \|\mathbf{X}^n; \mathbf{X}\|_{p'} \tag{2.18}$$

for any $p' \in (p,3)$ and any $q \in [1,2)$ such that $\frac{1}{p'} + \frac{1}{q} > 1$.

Note that γ^n and γ have finite 1-variation, with $\|\gamma^n\|_1 = \|\gamma\|_1 = T$, and $\|\gamma^n - \gamma\|_1 = 2T$. Although γ^n does not converge to γ in 1-variation, it is straightforward to see by interpolation that

$$\|\gamma^n - \gamma\|_q \leqslant \|\gamma^n - \gamma\|_1^{\frac{1}{q}} \|\gamma^n - \gamma\|_\infty^{1-\frac{1}{q}} = (2T)^{\frac{1}{q}} |\mathcal{P}^n|^{1-\frac{1}{q}}$$

for any $q > 1$. Combining this with the estimate in (2.18) and the result of Lemma 2.8, we infer the convergence in (2.17), and the estimate

$$\begin{aligned}\|\widetilde{Y}^n - Y\|_{p'} &\lesssim \|\gamma^n - \gamma\|_q + \|X^n - X\|_\infty^{1-\frac{p}{p'}} + \sup_{(s,t)\in\Delta_T} \left| \int_s^t X^n_{s,u} \otimes \mathrm{d}X_u - \mathbb{X}_{s,t} \right|^{1-\frac{p}{p'}} \\ &\lesssim |\mathcal{P}^n|^{1-\frac{1}{q}} + \|X^n - X\|_\infty^{1-\frac{p}{p'}} + \left\| \int_0^\cdot X^n_u \otimes \mathrm{d}X_u - \int_0^\cdot X_u \otimes \mathrm{d}X_u \right\|_\infty^{1-\frac{p}{p'}}. \end{aligned}$$

□

Remark 2.10. For a path $A \in D^1([0,T];\mathbb{R}^d)$ of finite 1-variation, let us consider the controlled ordinary differential equation (ODE)

$$Z_t = z_0 + \int_0^t \sigma(Z_s)\,\mathrm{d}A_s, \qquad t \in [0,T], \tag{2.19}$$

where the integral is interpreted in the Riemann–Stieltjes sense. It is a classical result that, provided that σ is sufficiently regular, the ODE in (2.19) is well posed, and that the solution map $\Phi\colon A \mapsto Z$ is continuous with respect to the 1-variation norm $\|\cdot\|_1$. A major insight of the theory of rough paths is that the solution map Φ can be extended from the space of smooth paths to the space $\mathscr{C}^{0,p\text{-var}}([0,T];\mathbb{R}^d)$ of continuous geometric rough paths for $p \in (2,3)$; see, for example, [21]. Of course, the closure of a set containing only continuous paths with respect to p-variation norms will again only contain continuous paths.

In the current framework of càdlàg rough paths, Lemma 2.8 and Proposition 2.9 motivate us to consider instead the closure of càdlàg paths of finite 1-variation. For $p \in (2,3)$, let $\mathcal{D}^{0,p}([0,T];\mathbb{R}^d)$ denote the closure of the set

$$\left\{ \mathbf{A} = (A,\mathbb{A}) \ :\ A \in D^1([0,T];\mathbb{R}^d) \text{ and } \mathbb{A}_{s,t} := \int_s^t A_{s,u} \otimes \mathrm{d}A_u \text{ for all } (s,t) \in \Delta_T \right\}$$

with respect to the rough path distance $\|\cdot\,;\cdot\|_p$ (as defined in (2.1)), where $\int_s^t A_{s,u} \otimes \mathrm{d}A_u$ is defined as a left-point Riemann–Stieltjes integral. Then, the solution map $\Phi\colon A \mapsto Z$ extends continuously to the space $\mathcal{D}^{0,p}([0,T];\mathbb{R}^d)$ by Theorem 2.2, and every path satisfying Property (RIE) is in $\mathcal{D}^{0,p'}([0,T];\mathbb{R}^d)$ for $p' \in (p,3)$ by Lemma 2.8.

Next, we shall verify that the piecewise constant approximation X^n of X, as defined in Property (RIE), itself satisfies Property (RIE) relative to any sequence of partitions $(\widetilde{\mathcal{P}}^m)_{m\in\mathbb{N}}$ that are coarser than $\mathcal{P}^n$ and have vanishing mesh size.

Lemma 2.11. *Suppose that a path X satisfies Property (RIE) relative to $p \in (2,3)$ and a sequence of partitions $(\mathcal{P}^n)_{n\in\mathbb{N}}$, and let X^n be the usual piecewise constant approximation of X along $\mathcal{P}^n$. Then the path X^n satisfies Property (RIE) relative to p and any sequence of partitions $(\widetilde{\mathcal{P}}^m)_{m\in\mathbb{N}}$ such that $\mathcal{P}^n \subseteq \widetilde{\mathcal{P}}^m$ for every $m \in \mathbb{N}$, and $|\widetilde{\mathcal{P}}^m| \to 0$ as $m \to \infty$.*

Proof. We need to verify each of the conditions (i)–(iii) of Property (RIE) along the sequence of partitions $(\widetilde{\mathcal{P}}^m)_{m\in\mathbb{N}}$. Since $\mathcal{P}^n \subseteq \widetilde{\mathcal{P}}^m$ for every $m \in \mathbb{N}$, the piecewise constant approximation of X^n along the partition $\widetilde{\mathcal{P}}^m$ is simply the path X^n itself. Conditions (i) and (ii) thus hold trivially.

Let $w_{1,n}$ be the control function given by $w_{1,n}(s,t) := \|X^n\|_{p,[s,t]}^p$, so that $|X^n_{s,t}|^p \leqslant w_{1,n}(s,t)$ for all $(s,t) \in \Delta_T$, and similarly, let $w_{2,n}$ be the control function given by $w_{2,n}(s,t) := \|\mathbb{X}^n\|_{\frac{p}{2},[s,t]}^{\frac{p}{2}}$. Let us also write $\widetilde{\mathcal{P}}^m = \{0 = r_0^m < r_1^m < \cdots < r_{\widetilde{N}_m}^m = T\}$ for each $m \in \mathbb{N}$. Then, for any $m \in \mathbb{N}$ and any $0 \leqslant k < \ell \leqslant \widetilde{N}_m$, using the standard estimate for Young integration (see, e.g., [22, Proposition 2.4]), we have that

$$\left|\int_{r_k^m}^{r_\ell^m} X_u^n \otimes \mathrm{d}X_u^n - X_{r_k^m}^n \otimes X_{r_k^m, r_\ell^m}^n\right|^{\frac{p}{2}} \lesssim \|X^n\|_{p,[r_k^m,r_\ell^m]}^{\frac{p}{2}} \|X^n\|_{\frac{p}{2},[r_k^m,r_\ell^m]}^{\frac{p}{2}}$$

$$\leqslant \|X^n\|_p^{\frac{p}{2}} \|X^n\|_{\frac{p}{2},[r_k^m,r_\ell^m]}^{\frac{p}{2}} \leqslant \|X\|_p^{\frac{p}{2}} w_{2,n}(r_k^m, r_\ell^m).$$

Thus, condition (iii) holds for X^n with the control function $w_{3,n}$, given by

$$w_{3,n}(s,t) := w_{1,n}(s,t) + \|X\|_p^{\frac{p}{2}} w_{2,n}(s,t), \qquad (s,t) \in \Delta_T. \qquad \square$$

We are now in a position to complete the proof of Theorem 2.3. For this, we will apply in particular the result of Theorem B.2, which states that, under Property (RIE), the rough integral can be obtained as a limit of classical left-point Riemann sums.

Proof of Theorem 2.3. Note that the Euler scheme in (2.5) may be expressed as the solution of the controlled ODE

$$Y_t^n = y_0 + \int_0^t b(\gamma_s^n, Y_s^n)\,\mathrm{d}\gamma_s^n + \int_0^t \sigma(\gamma_s^n, Y_s^n)\,\mathrm{d}X_s^n, \qquad t \in [0,T], \tag{2.20}$$

where γ^n denotes the time discretization path along $\mathcal{P}^n$ defined in (2.15), and the integrals are defined as limits of left-point Riemann sums. Recall that $\widetilde{Y}^n$ denotes the solution of the RDE in (2.16), that is,

$$\widetilde{Y}_t^n = y_0 + \int_0^t b(\gamma_s^n, \widetilde{Y}_s^n)\,\mathrm{d}\gamma_s^n + \int_0^t \sigma(\gamma_s^n, \widetilde{Y}_s^n)\,\mathrm{d}\mathbf{X}_s^n, \qquad t \in [0,T], \tag{2.21}$$

where $\mathbf{X}^n$ is the canonical rough path lift of X^n, as constructed in (2.9).

Since X^n is piecewise constant, it is clear from the definition of $\mathbb{X}^n$ that $\mathbb{X}_{s,t}^n = 0$ for any times $s \leqslant t$ that lie in the same subinterval $[t_k^n, t_{k+1}^n)$ of the partition $\mathcal{P}^n$. Since γ^n is also constant on each such subinterval, it follows from the definitions of Young and rough integrals that the solution $\widetilde{Y}^n$ of (2.21) is itself also piecewise constant along the partition $\mathcal{P}^n$.

Let $\widetilde{\mathcal{P}}^m = \{0 = r_0^m < r_1^m < \cdots < r_{\widetilde{N}_m}^m = T\}$, $m \in \mathbb{N}$, be any sequence of partitions with mesh size converging to 0, such that $\mathcal{P}^n \subseteq \widetilde{\mathcal{P}}^m$ for every $m \in \mathbb{N}$. By Lemma 2.11, we have that the path X^n satisfies Property (RIE) relative to p and the sequence $(\widetilde{\mathcal{P}}^m)_{m\in\mathbb{N}}$. Since γ^n and $\widetilde{Y}^n$ are piecewise constant along the partition $\mathcal{P}^n$, it is clear that the jump times of the integrand $s \mapsto \sigma(\gamma_s^n, \widetilde{Y}_s^n)$ all belong to $\mathcal{P}^n$, and thus also belong to the set $\liminf_{m\to\infty} \widetilde{\mathcal{P}}^m$. It thus follows from Theorem B.2 that the rough integral $\int_0^t \sigma(\gamma_s^n, \widetilde{Y}_s^n)\,\mathrm{d}\mathbf{X}_s^n$ is equal to a limit of left-point Riemann sums along the sequence $(\widetilde{\mathcal{P}}^m)_{m\in\mathbb{N}}$. That is, for any $t \in [0,T]$, we have that

$$\int_0^t \sigma(\gamma_s^n, \widetilde{Y}_s^n)\,\mathrm{d}\mathbf{X}_s^n = \lim_{m\to\infty} \sum_{k=0}^{\widetilde{N}_m - 1} \sigma(\gamma_{r_k^m}^n, \widetilde{Y}_{r_k^m}^n) X_{r_k^m\wedge t, r_{k+1}^m\wedge t}^n$$

$$= \sum_{k=0}^{N_n-1} \sigma(\gamma_{t_k^n}^n, \widetilde{Y}_{t_k^n}^n) X_{t_k^n\wedge t, t_{k+1}^n\wedge t}^n = \int_0^t \sigma(\gamma_s^n, \widetilde{Y}_s^n)\,\mathrm{d}X_s^n.$$

Since these integrals are equal, it follows that the ODE in (2.20) and the RDE in (2.21) are consistent, so that $Y^n = \widetilde{Y}^n$. The result then follows from Proposition 2.9. □

2.3 Error bound for an approximate Euler scheme

In general, the Euler scheme (2.5) is not applicable to numerically approximate the solution of an SDE driven by a general Lévy process — as we will consider in Section 3.3 below — since the increments of Lévy processes cannot always be simulated. Therefore, to obtain a numerical approximation of the solution of such a Lévy-driven SDE, one needs to consider approximate Euler schemes — see, for example, [13, 25, 38] — where the increments of the driving Lévy process are approximated by random variables with known distributions.

As a pathwise counterpart, we introduce the approximate Euler scheme $\widehat{Y}^n$ of the RDE (2.4) along the partition $\mathcal{P}^n$, given by

$$\widehat{Y}^n_t = y_0 + \sum_{i\,:\,t^n_{i+1} \leqslant t} b(t^n_i, \widehat{Y}^n_{t^n_i})(t^n_{i+1} - t^n_i) + \sum_{i\,:\,t^n_{i+1} \leqslant t} \sigma(t^n_i, \widehat{Y}^n_{t^n_i})(\widehat{X}_{t^n_{i+1}} - \widehat{X}_{t^n_i}), \tag{2.22}$$

for $t \in [0,T]$, with the modified driving signal

$$\widehat{X} := X + \varphi,$$

where $\varphi \in D^q([0,T];\mathbb{R}^d)$, for some $q \in [1,2)$ such that $\frac{1}{p} + \frac{1}{q} > 1$, and, as usual, we write $\mathcal{P}^n = \{0 = t^n_0 < t^n_1 < \cdots < t^n_{N_n} = T\}$.

While the approximation error of the Euler scheme (2.5) was only caused by discretizing the time interval $[0,T]$, the approximate Euler scheme (2.22) produces an additional approximation error due to taking the modified driving signal $\widehat{X}$ as an input, instead of the actual driving signal X of the RDE (2.4).

To ensure the convergence of the approximate Euler scheme, we first need to verify that, if the actual driving signal satisfies Property (RIE), then the same is true for the modified driving signal.

Proposition 2.12. *Suppose that $X \in D([0,T];\mathbb{R}^d)$ satisfies Property (RIE) relative to some $p \in (2,3)$ and a sequence of partitions $\mathcal{P}^n = \{0 = t^n_0 < t^n_1 < \cdots < t^n_{N_n} = T\}$, $n \in \mathbb{N}$. Let $\varphi \in D^q([0,T];\mathbb{R}^d)$ for some $q \in [1,2)$ such that $\frac{1}{p} + \frac{1}{q} > 1$. For each $n \in \mathbb{N}$, we define $\varphi^n : [0,T] \to \mathbb{R}^d$ by*

$$\varphi^n_t = \varphi_T \mathbf{1}_{\{T\}}(t) + \sum_{k=0}^{N_n - 1} \varphi_{t^n_k} \mathbf{1}_{[t^n_k, t^n_{k+1})}(t), \qquad t \in [0,T], \tag{2.23}$$

as the discretization of φ along $\mathcal{P}^n$. Suppose that $\|\varphi^n - \varphi\|_q \to 0$ as $n \to \infty$. Then the path $\widehat{X} = X + \varphi$ satisfies Property (RIE) relative to p and $(\mathcal{P}^n)_{n\in\mathbb{N}}$.

Proof. We need to verify the conditions (i)–(iii) of Property (RIE).

(i) Letting $\widehat{X}^n$ denote the piecewise constant approximation of $\widehat{X}$ along the partition $\mathcal{P}^n$, it is clear that $\widehat{X}^n = X^n + \varphi^n$ for each $n \in \mathbb{N}$. Since X^n converges uniformly to X by Property

(RIE), and $\|\varphi^n - \varphi\|_q \to 0$ by assumption, it is clear that $\widehat{X}^n$ converges uniformly to $\widehat{X}$ as $n \to \infty$.

(ii) We need to verify that the integral

$$\int_0^t \widehat{X}_u^n \otimes \mathrm{d}\widehat{X}_u = \int_0^t X_u^n \otimes \mathrm{d}X_u + \int_0^t X_u^n \otimes \mathrm{d}\varphi_u + \int_0^t \varphi_u^n \otimes \mathrm{d}X_u + \int_0^t \varphi_u^n \otimes \mathrm{d}\varphi_u,$$

converges as $n \to \infty$ to the limit

$$\int_0^t \widehat{X}_u \otimes \mathrm{d}\widehat{X}_u := \int_0^t X_u \otimes \mathrm{d}X_u + \int_0^t X_u \otimes \mathrm{d}\varphi_u + \int_0^t \varphi_u \otimes \mathrm{d}X_u + \int_0^t \varphi_u \otimes \mathrm{d}\varphi_u,$$

uniformly in $t \in [0, T]$, where the latter three integrals are defined as Young integrals.

Since X satisfies Property (RIE), we have that

$$\left\| \int_0^{\cdot} X_u^n \otimes \mathrm{d}X_u - \int_0^{\cdot} X_u \otimes \mathrm{d}X_u \right\|_\infty \longrightarrow 0 \qquad \text{as} \quad n \longrightarrow \infty.$$

Let $p' > p$ such that $\frac{1}{p'} + \frac{1}{q} > 1$. By the standard estimate for Young integrals — see, for example, [22, Proposition 2.4] — we have, for all $t \in [0, T]$, that

$$\left| \int_0^t X_u^n \otimes \mathrm{d}\varphi_u - \int_0^t X_u \otimes \mathrm{d}\varphi_u \right| \lesssim \|X^n - X\|_{p'} \|\varphi\|_q.$$

It follows by interpolation (see, e.g., [21, Proposition 5.5]) that

$$\|X^n - X\|_{p'} \leqslant \|X^n - X\|_\infty^{1-\frac{p}{p'}} \|X^n - X\|_p^{\frac{p}{p'}}.$$

Since X^n converges uniformly to X as $n \to \infty$, and $\sup_{n\in\mathbb{N}} \|X^n\|_p \leqslant \|X\|_p < \infty$, we deduce that

$$\left\| \int_0^{\cdot} X_u^n \otimes \mathrm{d}\varphi_u - \int_0^{\cdot} X_u \otimes \mathrm{d}\varphi_u \right\|_\infty \longrightarrow 0 \qquad \text{as} \quad n \longrightarrow \infty.$$

Similarly, for each $t \in [0, T]$, it holds that

$$\left| \int_0^t \varphi_u^n \otimes \mathrm{d}X_u - \int_0^t \varphi_u \otimes \mathrm{d}X_u \right| \lesssim \|\varphi^n - \varphi\|_q \|X\|_p,$$

and

$$\left| \int_0^t \varphi_u^n \otimes \mathrm{d}\varphi_u - \int_0^t \varphi_u \otimes \mathrm{d}\varphi_u \right| \lesssim \|\varphi^n - \varphi\|_q \|\varphi\|_q,$$

and, since $\|\varphi^n - \varphi\|_q \to 0$ as $n \to \infty$, we infer the required convergence.

(iii) We aim to find a control function w such that

$$\sup_{(s,t)\in\Delta_T}\frac{|\widehat{X}_{s,t}|^p}{w(s,t)}+\sup_{n\in\mathbb{N}}\sup_{0\leqslant k<\ell\leqslant N_n}\frac{|\int_{t_k^n}^{t_\ell^n}\widehat{X}^n_{t_k^n,u}\otimes \mathrm{d}\widehat{X}_u|^{\frac{p}{2}}}{w(t_k^n,t_\ell^n)}\leqslant 1, \tag{2.24}$$

where

$$\int_{t_k^n}^{t_\ell^n}\widehat{X}^n_{t_k^n,u}\otimes \mathrm{d}\widehat{X}_u=\int_{t_k^n}^{t_\ell^n}X^n_{t_k^n,u}\otimes \mathrm{d}X_u+\int_{t_k^n}^{t_\ell^n}X^n_{t_k^n,u}\otimes \mathrm{d}\varphi_u$$
$$+\int_{t_k^n}^{t_\ell^n}\varphi^n_{t_k^n,u}\otimes \mathrm{d}X_u+\int_{t_k^n}^{t_\ell^n}\varphi^n_{t_k^n,u}\otimes \mathrm{d}\varphi_u.$$

Let w_X be the control function with respect to which X satisfies Property (RIE), and define moreover the control function w_φ, given by $w_\varphi(s,t)=\|\varphi\|^q_{q,[s,t]}$ for $(s,t)\in\Delta_T$.

We have from Property (RIE) that

$$\sup_{(s,t)\in\Delta_T}\frac{|\widehat{X}_{s,t}|^p}{w_X(s,t)+w_\varphi(s,t)}\lesssim\sup_{(s,t)\in\Delta_T}\frac{|X_{s,t}|^p}{w_X(s,t)}+\sup_{(s,t)\in\Delta_T}\frac{|\varphi_{s,t}|^p}{w_\varphi(s,t)}\leqslant 2,$$

and that

$$\sup_{n\in\mathbb{N}}\sup_{0\leqslant k<\ell\leqslant N_n}\frac{\left|\int_{t_k^n}^{t_\ell^n}X^n_{t_k^n,u}\otimes \mathrm{d}X_u\right|^{\frac{p}{2}}}{w_X(t_k^n,t_\ell^n)}\leqslant 1.$$

By the standard estimate for Young integrals (see, e.g., [22, Proposition 2.4]), for every $n\in\mathbb{N}$ and $0\leqslant k<\ell\leqslant N_n$, we have

$$\left|\int_{t_k^n}^{t_\ell^n}X^n_{t_k^n,u}\otimes \mathrm{d}\varphi_u\right|^{\frac{p}{2}}\lesssim\|X^n\|^{\frac{p}{2}}_{p,[t_k^n,t_\ell^n]}\|\varphi\|^{\frac{p}{2}}_{q,[t_k^n,t_\ell^n]}$$
$$\leqslant\|X\|^{\frac{p}{2}}_{p,[t_k^n,t_\ell^n]}\|\varphi\|^{\frac{p}{2}}_{q,[t_k^n,t_\ell^n]}\leqslant w_X(t_k^n,t_\ell^n)^{\frac{1}{2}}w_\varphi(t_k^n,t_\ell^n)^{\frac{p}{2q}},$$

and we can similarly obtain

$$\left|\int_{t_k^n}^{t_\ell^n}\varphi^n_{t_k^n,u}\otimes \mathrm{d}X_u\right|^{\frac{p}{2}}\lesssim w_X(t_k^n,t_\ell^n)^{\frac{1}{2}}w_\varphi(t_k^n,t_\ell^n)^{\frac{p}{2q}}$$

and

$$\left|\int_{t_k^n}^{t_\ell^n}\varphi^n_{t_k^n,u}\otimes \mathrm{d}\varphi_u\right|^{\frac{p}{2}}\lesssim w_\varphi(t_k^n,t_\ell^n)^{\frac{p}{q}}.$$

Since $p \in (2,3)$ and $q \in [1,2)$, we have that $\frac{1}{2} + \frac{p}{2q} > 1$ and $\frac{p}{q} > 1$, and it follows that the maps $(s,t) \mapsto w_X(s,t)^{\frac{1}{2}} w_\varphi(s,t)^{\frac{p}{2q}}$ and $(s,t) \mapsto w_\varphi(s,t)^{\frac{p}{q}}$ are superadditive and thus control functions. We deduce that (2.24) holds with a control function w of the form

$$w(s,t) = C\left(w_X(s,t) + w_\varphi(s,t) + w_X(s,t)^{\frac{1}{2}} w_\varphi(s,t)^{\frac{p}{2q}} + w_\varphi(s,t)^{\frac{p}{q}}\right), \qquad (s,t) \in \Delta_T,$$

where $C > 0$ is a suitable constant which depends only on p and q. □

By Proposition 2.12, the modified driving signal $\widehat{X}$ satisfies Property (RIE), and can thus be canonically lifted to a rough path $\widehat{\mathbf{X}} = (\widehat{X}, \widehat{\mathbb{X}}) \in \mathcal{D}^p([0,T];\mathbb{R}^d)$ via (2.7). By Theorem 2.2, the RDE (2.4) driven by $\widehat{\mathbf{X}}$ has a unique solution $\widehat{Y}$, and the approximate Euler scheme $\widehat{Y}^n$ in (2.22) converges to $\widehat{Y}$ by Theorem 2.3. We will see an application of this to SDEs driven by Lévy processes in Section 3.3.

The next proposition provides an error and convergence analysis for the approximate Euler scheme (2.22) with respect to the solution Y of the RDE (2.4) driven by the rough path $\mathbf{X} = (X, \mathbb{X})$ under Property (RIE).

Proposition 2.13. *Suppose that $X \in D([0,T];\mathbb{R}^d)$ satisfies Property (RIE) relative to $p \in (2,3)$ and a sequence of partitions $(\mathcal{P}^n)_{n\in\mathbb{N}}$, and let $\mathbf{X}$ be its canonical rough path lift. Let $\varphi \in D^q([0,T];\mathbb{R}^d)$ for some $q \in (1,2)$ such that $\frac{1}{p} + \frac{1}{q} > 1$, let φ^n be the piecewise constant approximation of φ, as defined in (2.23), and assume that $\|\varphi^n - \varphi\|_q \to 0$ as $n \to \infty$. Let Y be the solution of the RDE (2.4) driven by $\mathbf{X}$, and let $\widehat{Y}^n$ be the approximate Euler scheme defined in (2.22). We have the error estimate*

$$\|\widehat{Y}^n - Y\|_{p'} \lesssim (1 + \|X\|_p + \|\varphi\|_q)\|\varphi\|_q + |\mathcal{P}^n|^{1-\frac{1}{q}} + (\|X^n - X\|_\infty + \|\varphi^n - \varphi\|_\infty)^{1-\frac{p}{p'}}$$
$$+ \left(\left\|\int_0^\cdot X_u^n \otimes \mathrm{d}X_u - \int_0^\cdot X_u \otimes \mathrm{d}X_u\right\|_\infty + \|X^n - X\|_{p'} + \|\varphi^n - \varphi\|_q\right)^{1-\frac{p}{p'}}$$

for any $p' \in (p,3)$ such that $\frac{1}{p'} + \frac{1}{q} > 1$, where the implicit multiplicative constant depends on $p, p', q, \|b\|_{C_b^2}, \|\sigma\|_{C_b^3}, T, \|X\|_\infty, \|\mathbf{X}\|_p, \|\varphi\|_\infty, \|\varphi\|_q$ and $w(0,T)$, where w is the control function for which (2.6) holds. In particular, we have that

$$\limsup_{n\to\infty} \|\widehat{Y}^n - Y\|_{p'} \lesssim (1 + \|X\|_p + \|\varphi\|_q)\|\varphi\|_q. \tag{2.25}$$

Proof. By Proposition 2.12, we know that the path $\widehat{X} = X + \varphi$ satisfies Property (RIE) relative to p and $(\mathcal{P}^n)_{n\in\mathbb{N}}$. Let $\widehat{\mathbf{X}}$ be the canonical rough path lift of $\widehat{X}$, and let Y and $\widehat{Y}$ be the solutions of the RDE (2.4) driven by $\mathbf{X}$ and $\widehat{\mathbf{X}}$, respectively. It is clear that

$$\|\widehat{Y}^n - Y\|_{p'} \leqslant \|\widehat{Y}^n - \widehat{Y}\|_{p'} + \|\widehat{Y} - Y\|_{p'}.$$

By Theorem 2.2, we have the estimate

$$\|\widehat{Y} - Y\|_{p'} \lesssim \|\widehat{\mathbf{X}}; \mathbf{X}\|_{p'},$$

and, by Theorem 2.3, we have that

$$\|\widehat{Y}^n - \widehat{Y}\|_{p'} \lesssim |\mathcal{P}^n|^{1-\frac{1}{q}} + \|\widehat{X}^n - \widehat{X}\|_\infty^{1-\frac{p}{p'}} + \left\|\int_0^\cdot \widehat{X}^n_u \otimes \mathrm{d}\widehat{X}_u - \int_0^\cdot \widehat{X}_u \otimes \mathrm{d}\widehat{X}_u\right\|_\infty^{1-\frac{p}{p'}},$$

where $\widehat{X}^n$ is the piecewise constant approximation of $\widehat{X}$ along $\mathcal{P}^n$. Since $\widehat{X}^n = X^n + \varphi^n$, we can bound

$$\|\widehat{X}^n - \widehat{X}\|_\infty \leqslant \|X^n - X\|_\infty + \|\varphi^n - \varphi\|_\infty.$$

As shown in the proof of Proposition 2.12,

$$\begin{aligned}
&\left\|\int_0^\cdot \widehat{X}^n_u \otimes \mathrm{d}\widehat{X}_u - \int_0^\cdot \widehat{X}_u \otimes \mathrm{d}\widehat{X}_u\right\|_\infty \\
&\quad\lesssim \left\|\int_0^\cdot X^n_u \otimes \mathrm{d}X_u - \int_0^\cdot X_u \otimes \mathrm{d}X_u\right\|_\infty + \|X^n - X\|_{p'}\|\varphi\|_q + \|\varphi^n - \varphi\|_q(\|X\|_p + \|\varphi\|_q).
\end{aligned}$$

We also note that

$$\widehat{\mathbb{X}}_{s,t} - \mathbb{X}_{s,t} = \int_s^t X_{s,u} \otimes \mathrm{d}\varphi_u + \int_s^t \varphi_{s,u} \otimes \mathrm{d}X_u + \int_s^t \varphi_{s,u} \otimes \mathrm{d}\varphi_u$$

for $(s,t) \in \Delta_T$, so that, by the standard estimate for Young integrals (see, e.g., [22, Proposition 2.4]), we obtain

$$|\widehat{\mathbb{X}}_{s,t} - \mathbb{X}_{s,t}| \lesssim \|X\|_{p,[s,t]}\|\varphi\|_{q,[s,t]} + \|\varphi\|^2_{q,[s,t]}.$$

This implies that, for any partition $\mathcal{P}$ of the interval $[0,T]$,

$$\begin{aligned}
\sum_{[s,t]\in\mathcal{P}} |\widehat{\mathbb{X}}_{s,t} - \mathbb{X}_{s,t}|^{\frac{p}{2}} &\lesssim \sum_{[s,t]\in\mathcal{P}} \left(\|X\|^{\frac{p}{2}}_{p,[s,t]}\|\varphi\|^{\frac{p}{2}}_{q,[s,t]} + \|\varphi\|^p_{q,[s,t]}\right) \\
&\leqslant \left(\sum_{[s,t]\in\mathcal{P}} \|X\|^p_{p,[s,t]}\right)^{\frac{1}{2}}\left(\sum_{[s,t]\in\mathcal{P}} \|\varphi\|^p_{q,[s,t]}\right)^{\frac{1}{2}} + \sum_{[s,t]\in\mathcal{P}} \|\varphi\|^p_{q,[s,t]} \\
&\leqslant \left(\sum_{[s,t]\in\mathcal{P}} \|X\|^p_{p,[s,t]}\right)^{\frac{1}{2}}\left(\sum_{[s,t]\in\mathcal{P}} \|\varphi\|^q_{q,[s,t]}\right)^{\frac{p}{2q}} + \left(\sum_{[s,t]\in\mathcal{P}} \|\varphi\|^q_{q,[s,t]}\right)^{\frac{p}{q}} \leqslant \|X\|_p^{\frac{p}{2}}\|\varphi\|_q^{\frac{p}{2}} + \|\varphi\|_q^p,
\end{aligned}$$

so that $\|\widehat{\mathbb{X}} - \mathbb{X}\|_{\frac{p}{2}} \lesssim \|X\|_p\|\varphi\|_q + \|\varphi\|_q^2$. We thus deduce that

$$\|\widehat{\mathbf{X}};\mathbf{X}\|_{p'} \leqslant \|\widehat{X} - X\|_p + \|\widehat{\mathbb{X}} - \mathbb{X}\|_{\frac{p}{2}} \lesssim (1 + \|X\|_p + \|\varphi\|_q)\|\varphi\|_q,$$

and combining the estimates above, we obtain the desired error estimate. □

As an immediate consequence of Proposition 2.13, if the modified driving signal $\widehat{X}$ converges to the driving signal X, then the approximate Euler scheme converges to the solution Y of the RDE (2.4). This is made precise in the following corollary, which follows from (2.25).

Corollary 2.14. *Recall the setting of Proposition 2.13, and now let $\check{Y}^n$ be the approximate Euler scheme of the RDE (2.4) along the partition $\mathcal{P}^n$, given by*

$$\check{Y}^n_t = y_0 + \sum_{i\,:\,t^n_{i+1} \leqslant t} b(t^n_i, \check{Y}^n_{t^n_i})(t^n_{i+1} - t^n_i) + \sum_{i\,:\,t^n_{i+1} \leqslant t} \sigma(t^n_i, \check{Y}^n_{t^n_i})(\check{X}^n_{t^n_{i+1}} - \check{X}^n_{t^n_i})$$

for $t \in [0, T]$, with the modified driving signal

$$\check{X}^n := X + \psi^n,$$

where $\psi^n \in D^q([0, T]; \mathbb{R}^d)$ for some $q \in (1, 2)$ such that $\frac{1}{p} + \frac{1}{q} > 1$. If $\|\psi^n\|_q \to 0$ as $n \to \infty$, then

$$\|\check{Y}^n - Y\|_{p'} \longrightarrow 0 \qquad as \quad n \longrightarrow \infty$$

for any $p' \in (p, 3)$ such that $\frac{1}{p'} + \frac{1}{q} > 1$.

Remark 2.15. In this section, we handled the modified driving signal $X + \varphi$ by considering the rough path lift $\widehat{\mathbf{X}}$ of $\widehat{X} = X + \varphi$, and considering the solution $\widehat{Y}$ of the RDE (2.4) driven by $\widehat{\mathbf{X}}$. An alternative, equally valid approach would be to instead absorb φ into the drift of the RDE. The resulting equation would not strictly speaking be of the form in (2.4), but it would still fall into the regime of the more general RDE in (2.2), and an error estimate could be obtained using the stability estimate in Theorem 2.2.

3 | APPLICATIONS TO STOCHASTIC DIFFERENTIAL EQUATIONS

In this section, we apply the deterministic theory developed in Section 2, regarding the Euler scheme for RDEs, to SDEs. For this purpose, we now let X be a d-dimensional càdlàg semimartingale, defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with a filtration $(\mathcal{F}_t)_{t \in [0,T]}$ satisfying the usual conditions, that is, completeness and right continuity. We consider the SDE

$$Y_t = y_0 + \int_0^t b(s, Y_{s-})\,\mathrm{d}s + \int_0^t \sigma(s, Y_{s-})\,\mathrm{d}X_s, \qquad t \in [0, T], \tag{3.1}$$

where $y_0 \in \mathbb{R}^k$, $b \in C^2_b(\mathbb{R}^{k+1}; \mathbb{R}^k)$, and $\sigma \in C^3_b(\mathbb{R}^{k+1}; \mathcal{L}(\mathbb{R}^d; \mathbb{R}^k))$, and $\int_0^t \sigma(s, Y_{s-})\,\mathrm{d}X_s$ is defined as an Itô integral. For a comprehensive introduction to stochastic Itô integration and SDEs, we refer, for example, to the textbook [36]. It is well known that the SDE (3.1) possesses a unique (strong) solution (see, e.g., [36, Chapter V, Theorem 6]), and that the semimartingale X can be lifted to a random rough path via Itô integration, by defining $\mathbf{X} = (X, \mathbb{X}) \in \mathcal{D}^p([0, T]; \mathbb{R}^d)$, $\mathbb{P}$-a.s., for any $p \in (2, 3)$, where

$$\mathbb{X}_{s,t} := \int_s^t (X_{r-} - X_s) \otimes \mathrm{d}X_r = \int_s^t X_{r-} \otimes \mathrm{d}X_r - X_s \otimes X_{s,t}, \qquad (s, t) \in \Delta_T; \tag{3.2}$$

see [31, Proposition 3.4] or [22, Theorem 6.5]. It turns out that, if the semimartingale X satisfies Property (RIE) relative to $p \in (2, 3)$ and a suitable sequence of partitions $(\mathcal{P}^n)_{n \in \mathbb{N}}$, then the solutions to the SDE (3.1) and to the RDE (2.4) driven by the random rough path $\mathbf{X} = (X, \mathbb{X})$ coincide $\mathbb{P}$-almost surely.

Lemma 3.1. *Let $p \in (2,3)$ and let $\mathcal{P}^n = \{\tau^n_k\}$, $n \in \mathbb{N}$, be a sequence of adapted partitions (so that each τ^n_k is a stopping time), such that, for almost every $\omega \in \Omega$, $(\mathcal{P}^n(\omega))_{n\in\mathbb{N}}$ is a sequence of (finite) partitions of $[0,T]$ with vanishing mesh size. Let X be a càdlàg semimartingale, and suppose that, for almost every $\omega \in \Omega$, the sample path $X(\omega)$ satisfies Property (RIE) relative to p and $(\mathcal{P}^n(\omega))_{n\in\mathbb{N}}$.*

(i) *The random rough paths $\mathbf{X} = (X, \mathbb{X})$, with $\mathbb{X}$ defined pathwise via (2.7), and with $\mathbb{X}$ defined by stochastic integration as in (3.2), coincide $\mathbb{P}$-almost surely.*
(ii) *The solution of the SDE (3.1) driven by X, and the solution of the RDE (2.4) driven by the random rough path $\mathbf{X} = (X, \mathbb{X})$, coincide $\mathbb{P}$-almost surely.*

Proof.

(i) By construction, the pathwise rough integral $\int_0^t X_u(\omega) \otimes \mathrm{d}X_u(\omega)$ constructed via Property (RIE) is given by the limit as $n \to \infty$ of left-point Riemann sums:

$$\sum_{k=0}^{N_n-1} X_{\tau^n_k(\omega)}(\omega) \otimes X_{\tau^n_k(\omega)\wedge t, \tau^n_{k+1}(\omega)\wedge t}(\omega). \tag{3.3}$$

It is known that these Riemann sums also converge uniformly in probability to the Itô integral $\int_0^t X_{u-} \otimes \mathrm{d}X_u$ (see, e.g., [36, Chapter II, Theorem 21]), and the result thus follows from the (almost sure) uniqueness of limits.

(ii) In the following, we adopt the notation J_F for the set of jump times of a path F, and we write $\liminf_{n\to\infty} \mathcal{P}^n := \bigcup_{m\in\mathbb{N}} \bigcap_{n\geqslant m} \mathcal{P}^n$.

Let Y be the solution to the RDE (2.4) driven by the random rough path $\mathbf{X} = (X, \mathbb{X})$. By the definition of $\mathbb{X}$ in (2.7), it is straightforward to see that $\mathbb{X}_{t-,t} = 0$ for every $t \in (0, T]$. It then follows from the definition of rough integration that the integral $t \mapsto \int_0^t \sigma(s, Y_s)\,\mathrm{d}\mathbf{X}_s$ can only have a jump at the jump times of X, and it follows that the same is true of the solution Y to the RDE (2.4), that is, $J_Y \subseteq J_X$.

Since the piecewise constant approximation X^n of X along $\mathcal{P}^n$ converges uniformly to X (by condition (i) of Property (RIE)), we have from Proposition B.1 that $J_X \subseteq \liminf_{n\to\infty} \mathcal{P}^n$. Since $J_Y \subseteq J_X$, we have that $J_Y \subseteq \liminf_{n\to\infty} \mathcal{P}^n$. It then follows from Theorem B.2 that

$$\int_0^t \sigma(s, Y_s)\,\mathrm{d}\mathbf{X}_s = \lim_{n\to\infty} \sum_{k=0}^{N_n-1} \sigma(\tau^n_k, Y_{\tau^n_k}) X_{\tau^n_k\wedge t, \tau^n_{k+1}\wedge t}.$$

Since these Riemann sums also converge in probability to the Itô integral $\int_0^t \sigma(s, Y_{s-})\,\mathrm{d}X_s$ (see, e.g., [36, Chapter II, Theorem 21]), these integrals coincide almost surely. We infer that Y is also a solution of the SDE (3.1), which has a unique solution (by, e.g., [36, Chapter V, Theorem 6]). □

As a consequence of Theorem 2.3 and Lemma 3.1, for semimartingales that satisfy Property (RIE) relative to a sequence of adapted partitions, the Euler scheme (2.5) converges pathwise to the solution of the SDE (3.1). In the following subsections, we verify Property (RIE) for various semimartingales relative to suitable sequences of partitions, and derive the pathwise convergence rate of the associated Euler scheme with respect to the p-variation norm.

3.1 | Brownian motion

We start with the most prominent example of a semimartingale, by taking $X = W$ to be a d-dimensional Brownian motion $W = (W_t)_{t\in[0,T]}$ with respect to the underlying filtration $(\mathcal{F}_t)_{t\in[0,T]}$.

Proposition 3.2. *Let $p \in (2,3)$ and let $\mathcal{P}^n = \{0 = t_0^n < t_1^n < \cdots < t_{N_n}^n = T\}$, $n \in \mathbb{N}$, be a sequence of equidistant partitions of the interval $[0,T]$, so that, for each $n \in \mathbb{N}$, there exists some $\pi_n > 0$ such that $t_{i+1}^n - t_i^n = \pi_n$ for each $0 \leqslant i < N_n$. If $\pi_n^{2-\frac{4}{p}} \log(n) \to 0$ as $n \to \infty$, then, for almost every $\omega \in \Omega$, the sample path $W(\omega)$ satisfies Property (RIE) relative to p and $(\mathcal{P}^n)_{n\in\mathbb{N}}$.*

Proof. As stated in Remark 2.4, Davie's condition implies Property (RIE). While [35, Appendix B] shows this for the sequence of partitions $(\mathcal{P}_U^n)_{n\in\mathbb{N}}$, where $\mathcal{P}_U^n = \{\frac{iT}{n} : i = 0,1,\dots,n\}$, that is, $\pi_n = \frac{T}{n}$, their proof actually holds for any sequence of equidistant partitions of the interval $[0,T]$. We therefore show the necessary condition proposed in [12], under the assumption that $\pi_n^{2-\frac{4}{p}} \log(n) \to 0$ as $n \to \infty$.

More precisely, let $\mathbf{W} = (W, \mathbb{W})$ be the Itô Brownian rough path lift of W. Write $\alpha := \frac{1}{p}$ and let $\beta \in (1-\alpha, 2\alpha)$. We show that, almost surely, there exists a constant $C > 0$ such that

$$\left| \sum_{m=k}^{\ell-1} \mathbb{W}^{ij}_{t_m^n, t_{m+1}^n} \right| \leqslant C(\ell - k)^\beta \pi_n^{2\alpha},$$

for every $i,j = 1,\dots,d$ and $n \in \mathbb{N}$, whenever $0 < k < \ell$ are integers such that $\ell\pi_n \leqslant T$.

Step 1. We recall that a (zero mean) random variable Z is said to be *sub-Gaussian* if its sub-Gaussian norm $\|Z\|_{\psi_2} := \inf\{z > 0 : \mathbb{E}[\exp(Z^2/z^2)] \leqslant 2\}$ is finite. It is well known that the sub-Gaussian property admits an equivalent formulation; namely, Z is sub-Gaussian if and only if $\mathbb{E}[\exp(\lambda^2 Z^2)] \leqslant \exp(\lambda^2 K^2)$ holds for all λ such that $|\lambda| \leqslant \frac{1}{K}$, for some $K > 0$. In this case, we have $\|Z\|_{\psi_2} = K$ up to a multiplicative constant.

We will prove that $\mathbb{W}^{ij}_{t_m^n, t_{m+1}^n}$, $m = k,\dots,\ell-1$, are independent sub-Gaussian random variables with sub-Gaussian norm $\|\mathbb{W}^{ij}_{t_m^n, t_{m+1}^n}\|_{\psi_2} = C\pi_n$ for some $C > 0$.

First, we note that, by [21, Proposition 13.4], for all $m \in \mathbb{N}$, the random variables

$$\frac{\mathbb{W}^{ij}_{t_m^n, t_{m+1}^n}}{t_{m+1}^n - t_m^n}$$

are independent and identically distributed, with the same distribution as $\mathbb{W}^{ij}_{0,1}$, and that the latter satisfies $\mathbb{E}[\exp(\eta \mathbb{W}^{ij}_{0,1})] < \infty$ for some sufficiently small $\eta > 0$, which is equivalent to the Gaussian tail property, that is, that $\|\mathbb{W}^{ij}_{0,1}\|_{L^q} \leqslant c\sqrt{q}$ for all $q \geqslant 1$, where the constant c is independent of q;

see [21, Lemma A.17]. As a consequence, using the fact that $t^n_{m+1} - t^n_m = \pi_n$ for all m, and setting $q = 2\nu$, we deduce that

$$\mathbb{E}[|\mathbb{W}^{ij}_{t^n_m,t^n_{m+1}}|^{2\nu}] \leqslant c^\nu \nu^\nu \pi_n^{2\nu}, \qquad \nu \in \mathbb{N}, \tag{3.4}$$

for a new constant $c > 0$ which does not depend on ν.

We now aim to show that there exists a constant $C > 0$ such that

$$\mathbb{E}[\exp(\lambda^2(\mathbb{W}^{ij}_{t^n_m,t^n_{m+1}})^2)] \leqslant \exp(C^2\pi_n^2\lambda^2), \tag{3.5}$$

for all λ such that $|\lambda| \leqslant \frac{1}{C\pi_n}$, which then implies that $\mathbb{W}^{ij}_{t^n_m,t^n_{m+1}}$ is sub-Gaussian with norm $\|\mathbb{W}^{ij}_{t^n_m,t^n_{m+1}}\|_{\psi_2} = C\pi_n$, up to a multiplicative constant which we may then absorb into C. Using the Taylor expansion for the exponential function, we get, for $\lambda \in \mathbb{R}$, that

$$\mathbb{E}[\exp(\lambda^2(\mathbb{W}^{ij}_{t^n_m,t^n_{m+1}})^2)] = \mathbb{E}\left[1 + \sum_{\nu=1}^{\infty} \frac{\lambda^{2\nu}(\mathbb{W}^{ij}_{t^n_m,t^n_{m+1}})^{2\nu}}{\nu!}\right] = 1 + \sum_{\nu=1}^{\infty} \frac{\lambda^{2\nu}\mathbb{E}[(\mathbb{W}^{ij}_{t^n_m,t^n_{m+1}})^{2\nu}]}{\nu!}.$$

By the bound in (3.4) and Stirling's a pproximation (which implies in particular that $\nu! \geqslant (\frac{\nu}{e})^\nu$ for all $\nu \geqslant 1$), we obtain

$$\mathbb{E}[\exp(\lambda^2(\mathbb{W}^{ij}_{t^n_m,t^n_{m+1}})^2)] \leqslant 1 + \sum_{\nu=1}^{\infty}(ec\lambda^2\pi_n^2)^\nu = \frac{1}{1 - ec\lambda^2\pi_n^2} \leqslant \exp(2ec\lambda^2\pi_n^2),$$

which is valid provided that

$$ec\lambda^2\pi_n^2 \leqslant \frac{1}{2}, \tag{3.6}$$

since $\frac{1}{1-x} \leqslant \exp(2x)$ for $x \in [0, \frac{1}{2}]$. We then obtain (3.5) by choosing $C = \sqrt{2ec}$, and note that then (3.6) does indeed hold when $|\lambda| \leqslant \frac{1}{C\pi_n}$.

Step 2. Let $C > 0$ be the constant found above, so that $\|\mathbb{W}^{ij}_{t^n_m,t^n_{m+1}}\|_{\psi_2} = C\pi_n$. Then Hoeffding's inequality (see, e.g., [40, Theorem 2.6.2]) gives

$$\begin{aligned}\mathbb{P}\left(\left|\sum_{m=k}^{\ell-1} \mathbb{W}^{ij}_{t^n_m,t^n_{m+1}}\right| \geqslant C(\ell-k)^\beta \pi_n^{2\alpha}\right) &\leqslant 2\exp\left(-\frac{C^2(\ell-k)^{2\beta}\pi_n^{4\alpha}}{\sum_{m=k}^{\ell-1}\|\mathbb{W}^{ij}_{t^n_m,t^n_{m+1}}\|^2_{\psi_2}}\right)\\ &= 2\exp\left(-\frac{(\ell-k)^{2\beta-1}}{\pi_n{}^{2-4\alpha}}\right).\end{aligned}$$

Since $\beta > 1 - \alpha > \frac{1}{2}$, we can bound this further by

$$\mathbb{P}\left(\left|\sum_{m=k}^{\ell-1} \mathbb{W}^{ij}_{t^n_m,t^n_{m+1}}\right| \geqslant C(\ell-k)^\beta \pi_n^{2\alpha}\right) \leqslant 2\exp\left(-\frac{1}{\pi_n{}^{2-4\alpha}}\right) = 2n^{-\frac{1}{\gamma_n}},$$

where we denote $\gamma_n = {\pi_n}^{2-4\alpha}\log(n)$. Since, by assumption, $\gamma_n \to 0$ as $n \to \infty$, we have that $\frac{1}{\gamma_n} > 1$ for all sufficiently large $n \in \mathbb{N}$, and hence that the series $\sum_{n\in\mathbb{N}} n^{-\frac{1}{\gamma_n}}$ is absolutely convergent. The desired statement then follows from the Borel–Cantelli lemma. □

Remark 3.3. Proposition 3.2 can be generalized to any sequence of partitions $(\mathcal{P}^n)_{n\in\mathbb{N}}$, which possibly consists of nonequidistant partitions, such that $|\mathcal{P}^n|^{2-\frac{4}{p}}\log(n) \to 0$ as $n \to \infty$, provided that there exists a positive number $\eta > 0$ such that

$$\frac{|\mathcal{P}^n|}{\min_{0\leqslant k<N_n}|t^n_{k+1} - t^n_k|} \leqslant \eta$$

for every $n \in \mathbb{N}$. This additional condition requires that the sequence $(\mathcal{P}^n)_{n\in\mathbb{N}}$ is a “balanced partition sequence” in the sense of [9].

Remark 3.4. Combining Proposition 3.2 with Lemma 2.8, we infer that the piecewise constant approximations of a Brownian motion along equidistant partitions converge to its Itô rough path lift, which, as far as we are aware, is a novel construction of this lift. Existing approximations of Brownian rough path are all continuous approximations, such as piecewise linear or mollifier approximations — cf. [21] — which play a crucial role, for example, in the rough path-based proofs of Wong–Zakai results, support theorems and large deviation principles.

Corollary 3.5. *Let $p \in (2,3)$ and let $\mathcal{P}^n_U = \{0 = t^n_0 < t^n_1 < \cdots < t^n_n = T\}$, $n \in \mathbb{N}$, with $t^n_i = \frac{iT}{n}$, be the sequence of equidistant partitions with width $\frac{T}{n}$ of the interval $[0,T]$. Let Y be the solution of the SDE (3.1) driven by a Brownian motion W, and let Y^n be the corresponding Euler approximation along $\mathcal{P}^n_U$, as defined in (2.5). For any $p' \in (p,3)$, $q \in (1,2)$ and $\beta \in (1-\frac{1}{p}, \frac{2}{p})$ such that $\frac{1}{p'} + \frac{1}{q} > 1$, there exists a random variable C, which does not depend on n, such that*

$$\|Y^n - Y\|_{p'} \leqslant C\left(n^{-(1-\frac{1}{q})} + n^{-(\frac{2}{p}-\beta)(1-\frac{p}{p'})}\right), \qquad n \in \mathbb{N}. \tag{3.7}$$

Proof. Since $|\mathcal{P}^n_U| = \frac{T}{n}$, we have that $|\mathcal{P}^n_U|^{2-\frac{4}{p}}\log(n) \to 0$ as $n \to \infty$. Thus, by Proposition 3.2, for almost every $\omega \in \Omega$, the sample path $W(\omega)$ satisfies Property (RIE) relative to p and $(\mathcal{P}^n_U)_{n\in\mathbb{N}}$, which allows us to apply the result of Theorem 2.3.

Since the sample paths of W are almost surely $\frac{1}{p}$-Hölder continuous, it is easy to see that

$$\|W^n - W\|_\infty \lesssim n^{-\frac{1}{p}}, \qquad n \in \mathbb{N},$$

where the implicit multiplicative constant is a random variable that does not depend on n. Moreover, by [35, Appendix B], the left-point Riemann sums along $(\mathcal{P}^n_U)_{n\in\mathbb{N}}$ converge uniformly as $n \to \infty$, with rate $n^{-(\frac{2}{p}-\beta)}$ for $\beta \in (1-\frac{1}{p}, \frac{2}{p})$, that is,

$$\left\|\int_0^\cdot W^n_u \otimes \mathrm{d}W_u - \int_0^\cdot W_u \otimes \mathrm{d}W_u\right\|_\infty \lesssim n^{-(\frac{2}{p}-\beta)}, \qquad n \in \mathbb{N}.$$

Hence, by Theorem 2.3, we get that

$$\|Y^n - Y\|_{p'} \lesssim n^{-(1-\frac{1}{q})} + n^{-\frac{1}{p}(1-\frac{p}{p'})} + n^{-(\frac{2}{p}-\beta)(1-\frac{p}{p'})}.$$

Since $\frac{1}{p} < 1 - \frac{1}{p} < \beta$ for $p \in (2,3)$, this gives the rate of convergence in (3.7). □

3.2 | Itô processes

In this subsection, we let X be an Itô process. More precisely, we suppose that

$$X_t = x_0 + \int_0^t b_r \,\mathrm{d}r + \int_0^t H_r \,\mathrm{d}W_r, \qquad t \in [0,T], \tag{3.8}$$

for some $x_0 \in \mathbb{R}^d$, and some locally bounded predictable integrands $b\colon \Omega \times [0,T] \to \mathbb{R}^d$ and $H\colon \Omega \times [0,T] \to \mathcal{L}(\mathbb{R}^m;\mathbb{R}^d)$, where W is an $\mathbb{R}^m$-valued Brownian motion. We consider the sequence of dyadic partitions $(\mathcal{P}^n_D)_{n\in\mathbb{N}}$ of $[0,T]$, given by

$$\mathcal{P}^n_D := \{0 = t^n_0 < t^n_1 < \cdots < t^n_{2^n} = T\} \qquad \text{with} \quad t^n_k := k2^{-n}T \quad \text{for} \quad k = 0,1,\dots,2^n. \tag{3.9}$$

In the next proposition, we will show that X satisfies Property (RIE) along the sequence of partitions $(\mathcal{P}^n_D)_{n\in\mathbb{N}}$, and establish the rate of convergence of the associated Euler scheme. Note that, in contrast to the proof of Proposition 3.2, for general Itô processes, we cannot rely on the concentration of measure inequality for sub-Gaussian distributions.

Proposition 3.6. *Let $p \in (2,3)$ and let X be an Itô process of the form in (3.8). Let Y be the solution of the SDE (3.1) driven by X, and let Y^n denote the corresponding Euler approximation, as defined in (2.5), based on X and the sequence of dyadic partitions $(\mathcal{P}^n_D)_{n\in\mathbb{N}}$.*

(i) *For almost every $\omega \in \Omega$, the sample path $X(\omega)$ satisfies Property (RIE) relative to p and $(\mathcal{P}^n_D)_{n\in\mathbb{N}}$.*
(ii) *For any $p' \in (p,3)$ and $q \in (1,2)$ such that $\frac{1}{p'} + \frac{1}{q} > 1$, and any $\varepsilon \in (0,1)$, there exists a random variable C, which does not depend on n, such that*

$$\|Y^n - Y\|_{p'} \leqslant C\big(2^{-n(1-\frac{1}{q})} + 2^{-n(\frac{1}{p}-\frac{1}{p'})} + 2^{-\frac{n}{2}(1-\varepsilon)(1-\frac{p}{p'})}\big), \qquad n \in \mathbb{N}, \tag{3.10}$$

and

$$\|Y^n - Y\|_3 \leqslant C2^{-n(\frac{1}{6}-\varepsilon)}, \qquad n \in \mathbb{N}. \tag{3.11}$$

Proof.

(i) By a localization argument, we may assume that b and H are globally bounded. Let

$$A_t := \int_0^t b_r \,\mathrm{d}r \qquad \text{and} \qquad M_t := \int_0^t H_r \,\mathrm{d}W_r$$

for $t \in [0,T]$, so that $X = x_0 + A + M$, and recall that we denote the piecewise constant approximation of X along $\mathcal{P}^n_D$ by

$$X^n_t = X_T \mathbf{1}_T(t) + \sum_{k=0}^{2^n-1} X_{t^n_k} \mathbf{1}_{[t^n_k, t^n_{k+1})}(t), \qquad t \in [0,T],$$

with $t^n_k = k2^{-n}T$ for each $k = 0, 1, \dots, 2^n$ and $n \in \mathbb{N}$. Note that, by the uniform continuity of the sample paths of X, it is clear that X^n converges uniformly to X almost surely as $n \to \infty$.

Step 1. In this step, we verify that the sample paths of X are almost surely $\frac{1}{p}$-Hölder continuous. This is a standard application of the Burkholder–Davis–Gundy inequality. Indeed, for any $q \geqslant 1$, using the boundedness of H, and writing $[\cdot]$ for quadratic variation, we have that

$$\mathbb{E}[|M_t - M_s|^q] = \mathbb{E}\left[\left|\int_s^t H_u \,\mathrm{d}W_u\right|^q\right] \lesssim \mathbb{E}\left[\left[\int_0^\cdot H_u \,\mathrm{d}W_u\right]_{s,t}^{\frac{q}{2}}\right] \lesssim |t-s|^{\frac{q}{2}},$$

so that $\|M_t - M_s\|_{L^q} \lesssim |t-s|^{\frac{1}{2}}$. By the Kolmogorov continuity theorem (see, e.g., [21, Theorem A.10]), it follows that $\mathbb{E}[\|M\|_{\gamma\text{-Höl}}] < \infty$, where $\|\cdot\|_{\gamma\text{-Höl}}$ denotes the γ-Hölder norm, for any $\gamma \in [0, \frac{1}{2} - \frac{1}{q})$, which, taking q sufficiently large, implies that the sample paths of M are almost surely $\frac{1}{p}$-Hölder continuous. Since $A = \int_0^\cdot b_r \,\mathrm{d}r$ with the bounded integrand b, the sample paths of A are Lipschitz continuous, and thus also $\frac{1}{p}$-Hölder continuous.

Step 2. In this step, we show that, almost surely, $\int_0^\cdot X^n_u \otimes \mathrm{d}X_u$ converges uniformly to the Itô integral $\int_0^\cdot X_u \otimes \mathrm{d}X_u$ as $n \to \infty$. For this purpose, we write $X^n = x_0 + A^n + M^n$, where

$$A^n_t := A_T \mathbf{1}_{\{T\}}(t) + \sum_{k=0}^{2^n-1} A_{t^n_k} \mathbf{1}_{[t^n_k, t^n_{k+1})}(t) \quad \text{and} \quad M^n_t := M_T \mathbf{1}_{\{T\}}(t) + \sum_{k=0}^{2^n-1} M_{t^n_k} \mathbf{1}_{[t^n_k, t^n_{k+1})}(t),$$

for $t \in [0,T]$. Since $X = x_0 + A + M$, we obtain

$$\begin{aligned} &\mathbb{E}\left[\left\|\int_0^\cdot X^n_u \otimes \mathrm{d}X_u - \int_0^\cdot X_u \otimes \mathrm{d}X_u\right\|_\infty^2\right] \\ &\lesssim \mathbb{E}\left[\left\|\int_0^\cdot (A^n_u - A_u) \otimes \mathrm{d}A_u\right\|_\infty^2\right] + \mathbb{E}\left[\left\|\int_0^\cdot (M^n_u - M_u) \otimes \mathrm{d}A_u\right\|_\infty^2\right] \\ &\quad + \mathbb{E}\left[\left\|\int_0^\cdot (A^n_u - A_u) \otimes \mathrm{d}M_u\right\|_\infty^2\right] + \mathbb{E}\left[\left\|\int_0^\cdot (M^n_u - M_u) \otimes \mathrm{d}M_u\right\|_\infty^2\right]. \end{aligned} \tag{3.12}$$

Using the Burkholder–Davis–Gundy inequality, the fact that $[M] = [\int_0^\cdot H_t \,\mathrm{d}W_t] = \int_0^\cdot |H_t|^2 \,\mathrm{d}t$, and the boundedness of H, we can bound

$$\mathbb{E}\left[\left\|\int_0^\cdot (M^n_u - M_u) \otimes \mathrm{d}M_u\right\|_\infty^2\right] \lesssim \mathbb{E}\left[\int_0^T |M^n_t - M_t|^2 \,\mathrm{d}[M]_t\right]$$

$$\lesssim \int_0^T \mathbb{E}[|M_t^n - M_t|^2]\,\mathrm{d}t = \sum_{k=0}^{2^n-1} \int_{t_k^n}^{t_{k+1}^n} \mathbb{E}[|M_{t_k^n} - M_t|^2]\,\mathrm{d}t \lesssim \sum_{k=0}^{2^n-1} \int_{t_k^n}^{t_{k+1}^n} \mathbb{E}[|[M]_{t_k^n,t}|]\,\mathrm{d}t$$

$$= \sum_{k=0}^{2^n-1} \int_{t_k^n}^{t_{k+1}^n} \mathbb{E}\left[\int_{t_k^n}^{t} |H_r|^2\,\mathrm{d}r\right] \mathrm{d}t \lesssim \sum_{k=0}^{2^n-1} \int_{t_k^n}^{t_{k+1}^n} (t - t_k^n)\,\mathrm{d}t \leqslant \sum_{k=0}^{2^n-1} (t_{k+1}^n - t_k^n)^2 = 2^{-n}.$$

The other terms on the right-hand side of (3.12) can be bounded similarly by 2^{-n}, up to a constant that does not depend on n, and we thus have that

$$\mathbb{E}\left[\left\| \int_0^{\cdot} X_u^n \otimes \mathrm{d}X_u - \int_0^{\cdot} X_u \otimes \mathrm{d}X_u \right\|_\infty^2\right] \lesssim 2^{-n},$$

for every $n \in \mathbb{N}$. By Markov's inequality, for any $\varepsilon \in (0,1)$, we then have that

$$\mathbb{P}\left(\left\| \int_0^{\cdot} X_u^n \otimes \mathrm{d}X_u - \int_0^{\cdot} X_u \otimes \mathrm{d}X_u \right\|_\infty \geqslant 2^{-\frac{n}{2}(1-\varepsilon)} \right)$$

$$\leqslant 2^{n(1-\varepsilon)} \mathbb{E}\left[\left\| \int_0^{\cdot} X_u^n \otimes \mathrm{d}X_u - \int_0^{\cdot} X_u \otimes \mathrm{d}X_u \right\|_\infty^2\right] \lesssim 2^{n(1-\varepsilon)} 2^{-n} = 2^{-n\varepsilon}.$$

It then follows from the Borel–Cantelli lemma that, almost surely,

$$\left\| \int_0^{\cdot} X_u^n \otimes \mathrm{d}X_u - \int_0^{\cdot} X_u \otimes \mathrm{d}X_u \right\|_\infty < 2^{-\frac{n}{2}(1-\varepsilon)} \tag{3.13}$$

for all sufficiently large n, which implies the desired convergence.

Step 3. Let $\varepsilon \in (0,1)$ and $\rho = 2 + \frac{(1-\varepsilon)(p-2)}{4} \in (2,3)$. We infer from Step 1 above that the sample paths of X are almost surely $\frac{1}{\rho}$-Hölder continuous, from which it follows that

$$|X_{s,t}| \lesssim |t - s|^{\frac{1}{\rho}},$$

where the implicit multiplicative constant is a random variable that does not depend on s or t. Proceeding as in the proof of [31, Lemma 3.2], we can show, for any $0 \leqslant k < \ell \leqslant 2^n$, and writing $N = \ell - k = 2^n |t_\ell^n - t_k^n| T^{-1}$, that

$$\left| \int_{t_k^n}^{t_\ell^n} X_u^n \otimes \mathrm{d}X_u - X_{t_k^n} \otimes X_{t_k^n,t_\ell^n} \right| \lesssim N^{1-\frac{2}{\rho}} |t_\ell^n - t_k^n|^{\frac{2}{\rho}} \lesssim 2^{n(1-\frac{2}{\rho})} |t_\ell^n - t_k^n| \leqslant 2^{n(\rho-2)} |t_\ell^n - t_k^n|.$$

If $2^{-n} \geqslant |t_\ell^n - t_k^n|^{\frac{4}{p(1-\varepsilon)}}$, then it follows that

$$\left| \int_{t_k^n}^{t_\ell^n} X_u^n \otimes \mathrm{d}X_u - X_{t_k^n} \otimes X_{t_k^n,t_\ell^n} \right| \lesssim |t_\ell^n - t_k^n|^{1-\frac{4}{p(1-\varepsilon)}(\rho-2)} = |t_\ell^n - t_k^n|^{\frac{2}{p}}.$$

We will now aim to obtain the same estimate in the case that $2^{-n} < |t^n_\ell - t^n_k|^{\frac{4}{p(1-\varepsilon)}}$. To this end, let $\mathbb{X}$ denote the second-level component of the Itô rough path lift of X, as defined in (3.2). It follows from the Kolmogorov criterion for rough paths (see [17, Theorem 3.1]) that

$$|\mathbb{X}_{s,t}| \lesssim |t-s|^{\frac{2}{p}}, \tag{3.14}$$

where the implicit multiplicative constant is a random variable that does not depend on s or t. Using the bounds in (3.13) and (3.14), we then have, for all sufficiently large n, that

$$\begin{aligned}
&\left|\int_{t^n_k}^{t^n_\ell} X^n_u \otimes \mathrm{d}X_u - X_{t^n_k} \otimes X_{t^n_k,t^n_\ell}\right| \\
&= \left|\int_{t^n_k}^{t^n_\ell} X^n_u \otimes \mathrm{d}X_u - \int_{t^n_k}^{t^n_\ell} X_u \otimes \mathrm{d}X_u + \int_{t^n_k}^{t^n_\ell} X_u \otimes \mathrm{d}X_u - X_{t^n_k} \otimes X_{t^n_k,t^n_\ell}\right| \\
&\leqslant 2\left\|\int_0^\cdot X^n_u \otimes \mathrm{d}X_u - \int_0^\cdot X_u \otimes \mathrm{d}X_u\right\|_\infty + |\mathbb{X}_{t^n_k,t^n_\ell}| \\
&\lesssim 2^{-\frac{n}{2}(1-\varepsilon)} + |t^n_\ell - t^n_k|^{\frac{2}{p}} \\
&\lesssim |t^n_\ell - t^n_k|^{\frac{2}{p}}.
\end{aligned}$$

We have thus established that

$$\left|\int_{t^n_k}^{t^n_\ell} X^n_u \otimes \mathrm{d}X_u - X_{t^n_k} \otimes X_{t^n_k,t^n_\ell}\right|^{\frac{p}{2}} \lesssim |t^n_\ell - t^n_k|$$

holds for all $0 \leqslant k < \ell \leqslant 2^n$ and all sufficiently large n. It follows that there exists a random control function $w(s,t) := c|t-s|$, for some random variable c, such that

$$\sup_{(s,t)\in\Delta_T} \frac{|X_{s,t}|^p}{w(s,t)} + \sup_{n\in\mathbb{N}} \sup_{0\leqslant k<\ell\leqslant 2^n} \frac{|\int_{t^n_k}^{t^n_\ell} X^n_u \otimes \mathrm{d}X_u - X_{t^n_k} \otimes X_{t^n_k,t^n_\ell}|^{\frac{p}{2}}}{w(t^n_k, t^n_\ell)} \leqslant 1$$

holds almost surely. This means that, for almost every $\omega \in \Omega$, the sample path $X(\omega)$ satisfies Property (RIE) relative to any $p \in (2,3)$ and the sequence of dyadic partitions $(\mathcal{P}^n_D)_{n\in\mathbb{N}}$.

(ii) Since the sample paths of X are almost surely $\frac{1}{p}$-Hölder continuous (by Step 1 above), it is straightforward to see that

$$\|X^n - X\|_\infty \lesssim 2^{-\frac{n}{p}}, \qquad n \in \mathbb{N},$$

and, recalling (3.13), we have that

$$\left\|\int_0^\cdot X^n_u \otimes \mathrm{d}X_u - \int_0^\cdot X_u \otimes \mathrm{d}X_u\right\|_\infty \lesssim 2^{-\frac{n}{2}(1-\varepsilon)}, \qquad n \in \mathbb{N}.$$

Hence, by Theorem 2.3, we deduce that

$$\|Y^n - Y\|_3 \leqslant \|Y^n - Y\|_{p'} \lesssim 2^{-n(1-\frac{1}{q})} + 2^{-\frac{n}{p}(1-\frac{p}{p'})} + 2^{-\frac{n}{2}(1-\varepsilon)(1-\frac{p}{p'})},$$

for any $p' \in (p, 3)$ and $q \in (1, 2)$ such that $\frac{1}{p'} + \frac{1}{q} > 1$, which leads to (3.10). Choosing p sufficiently close to 2, p' to 3, and q to $\frac{3}{2}$, and replacing ε by 6ε, then reveals (3.11). □

3.3 | Lévy processes

Let $L = (L_t)_{t\in[0,T]}$ be a d-dimensional Lévy process with characteristics (λ, Σ, ν). In this section, we shall work under the assumption that $\int_{|x|<1} |x|^q \, \nu(\mathrm{d}x) < \infty$ for some $q \in [1, 2)$.

By the Lévy–Itô decomposition (see, e.g., [3, Theorem 2.4.16]), there exists a Brownian motion W with covariance matrix Σ, and an independent Poisson random measure μ on $[0, T] \times (\mathbb{R}^d \setminus \{0\})$ with compensator ν, such that $L = W + \varphi$, where

$$\varphi_t = \lambda t + \int_{|x|\geqslant 1} x \, \mu(t, \mathrm{d}x) + \int_{|x|<1} x \, (\mu(t, \mathrm{d}x) - t\, \nu(\mathrm{d}x)), \qquad t \in [0, T]. \tag{3.15}$$

Since $\int_{|x|<1} |x|^q \, \nu(\mathrm{d}x) < \infty$, we have that $\varphi(\omega) \in D^q([0, T]; \mathbb{R}^d)$ for almost every $\omega \in \Omega$; see [3, Theorem 2.4.25] and [8, Théorème IIIb].

Let $(\mathcal{P}_D^n)_{n\in\mathbb{N}}$ be the dyadic partitions of $[0, T]$, as defined in (3.9). For each $n \in \mathbb{N}$, we also let $J^n = \{t \in (0, T] \, : \, |\Delta\varphi_t| \geqslant 2^{-n}\}$, where $\Delta\varphi_t = \varphi_t - \varphi_{t-}$ denotes the jump of φ at time t, and we let

$$\mathcal{P}_L^n = \mathcal{P}_D^n \cup J^n. \tag{3.16}$$

We will consider $(\mathcal{P}_L^n)_{n\in\mathbb{N}}$ as our sequence of adapted partitions, noting in particular that, for almost every $\omega \in \Omega$, $(\mathcal{P}_L^n(\omega))_{n\in\mathbb{N}}$ is a nested sequence of (finite) partitions with vanishing mesh size, and that $\{t \in (0, T] \, : \, L_{t-}(\omega) \neq L_t(\omega)\} \subseteq \cup_{n\in\mathbb{N}} \mathcal{P}_L^n(\omega)$.

Remark 3.7. In order to obtain pointwise convergence of an Euler scheme, it is necessary that the jump times of the driving signal belong to the partitions used to construct the discretization, a fact that follows immediately from Proposition B.1, necessitating the inclusion of the jump times $(J^n)_{n\in\mathbb{N}}$ above.

Proposition 3.8. *Let L be a d-dimensional Lévy process with characteristics (λ, Σ, ν), and assume that $\int_{|x|<1} |x|^q \, \nu(\mathrm{d}x) < \infty$ for some $q \in [1, 2)$. Let $p \in (2, 3)$ such that $\frac{1}{p} + \frac{1}{q} > 1$. Let Y be the solution to the SDE (3.1) driven by L, and let Y^n be the corresponding Euler approximation along $\mathcal{P}_L^n$, as defined in (2.5).*

(i) *For almost every $\omega \in \Omega$, the sample path $L(\omega)$ satisfies Property (RIE) relative to p and $(\mathcal{P}_L^n(\omega))_{n\in\mathbb{N}}$.*
(ii) *For any $p' \in (p, 3)$ and $q' \in (q, 2)$ such that $\frac{1}{p'} + \frac{1}{q'} > 1$, any $\gamma \in (0, \frac{1}{p})$, and any $\delta \in (0, 1 - \frac{q}{2})$, there exists a random variable C, which does not depend on n, such that*

$$\|Y^n - Y\|_{p'} \leqslant C\left(2^{-n(1-\frac{1}{q'})} + (2^{-n(\frac{1}{p}-\gamma)} + 2^{-n(\frac{1}{p}-\frac{1}{p'})} + 2^{-n\delta(1-\frac{q}{q'})})^{1-\frac{p}{p'}}\right), \qquad n \in \mathbb{N}.$$

To prove this statement, we need the following lemma.

Lemma 3.9. *Let $p \in (2,3)$, let W be a d-dimensional Brownian motion with covariance matrix Σ, and let $(\mathcal{P}^n_L)_{n\in\mathbb{N}}$ be the sequence of adapted partitions defined in (3.16). For almost every $\omega \in \Omega$, the sample path $W(\omega)$ satisfies Property (RIE) relative to p and $(\mathcal{P}^n_L(\omega))_{n\in\mathbb{N}}$.*

Proof. We need to verify each of the conditions (i)–(iii) in Property (RIE).

(i) Since the sample paths of W are uniformly continuous on the compact interval $[0,T]$, it is straightforward to see that $W^n(\omega) \to W(\omega)$ uniformly as $n \to \infty$ for almost every $\omega \in \Omega$, where W^n denotes the piecewise constant approximation of W along $\mathcal{P}^n_L$.

(ii) It follows from the Kolmogorov continuity criterion that the sample paths of Brownian motion are almost surely $\frac{1}{p}$-Hölder continuous, and that the Hölder constant $\|W\|_{\frac{1}{p}\text{-Höl}}$ has finite moments of all orders (see, e.g., [6, Theorem A.1]). Applying the Burkholder–Davis–Gundy inequality, we then have that

$$\mathbb{E}\left[\left\|\int_0^\cdot W^n_u \otimes \mathrm{d}W_u - \int_0^\cdot W_u \otimes \mathrm{d}W_u\right\|_\infty^2\right] \lesssim \mathbb{E}\left[\int_0^T |W^n_t - W_t|^2 \,\mathrm{d}t\right]$$
$$\leqslant \mathbb{E}\left[\|W\|^2_{\frac{1}{p}\text{-Höl}} \int_0^T |\mathcal{P}^n_L|^{\frac{2}{p}} \,\mathrm{d}t\right] \lesssim \mathbb{E}[\|W\|^2_{\frac{1}{p}\text{-Höl}}]2^{-\frac{2n}{p}}.$$

Let $\gamma \in (0, \frac{1}{p})$ and $\varepsilon = 1 - \frac{2}{p} + 2\gamma \in (1 - \frac{2}{p}, 1)$. By Markov's inequality, we infer that

$$\mathbb{P}\left(\left\|\int_0^\cdot W^n_u \otimes \mathrm{d}W_u - \int_0^\cdot W_u \otimes \mathrm{d}W_u\right\|_\infty \geqslant 2^{-\frac{n}{2}(1-\varepsilon)}\right) \lesssim 2^{-\frac{2n}{p}+n(1-\varepsilon)} = 2^{-2n\gamma}.$$

By the Borel–Cantelli lemma, we then have that, almost surely,

$$\left\|\int_0^\cdot W^n_u \otimes \mathrm{d}W_u - \int_0^\cdot W_u \otimes \mathrm{d}W_u\right\|_\infty < 2^{-\frac{n}{2}(1-\varepsilon)} \tag{3.17}$$

for all sufficiently large n. It follows that $(\int_0^\cdot W^n_u \otimes \mathrm{d}W_u)(\omega)$ converges uniformly to $(\int_0^\cdot W_u \otimes \mathrm{d}W_u)(\omega)$ as $n \to \infty$ for almost every $\omega \in \Omega$.

(iii) Let $\rho = 2 + \frac{(1-\varepsilon)(p-2)}{4} \in (2,3)$. Since the sample paths of W are almost surely $\frac{1}{\rho}$-Hölder continuous, it follows that

$$|W_{s,t}|^\rho \lesssim |t-s|,$$

where the implicit multiplicative constant is a random variable that does not depend on s or t. Proceeding as in the proof of [31, Lemma 3.2], we can show, for any $0 \leqslant k < \ell$, and writing $N = \ell - k$, we can show that

$$\left|\int_{t_k^n}^{t_\ell^n} W_u^n \otimes \mathrm{d}W_u - W_{t_k^n} \otimes W_{t_k^n, t_\ell^n}\right| \lesssim N^{1-\frac{2}{\rho}} |t_\ell^n - t_k^n|^{\frac{2}{\rho}},$$

where $\{0 = t_0^n < t_1^n < \cdots\}$ are the partition points of $\mathcal{P}_L^n(\omega)$ for some (here fixed) $\omega \in \Omega$. Using $|\cdot|$ here to denote the cardinality of a set, we note that the number N can be bounded by

$$N \leqslant |\mathcal{P}_D^n(\omega) \cap (t_k^n, t_\ell^n]| + |J^n(\omega) \cap (t_k^n, t_\ell^n]| \leqslant 2^n T^{-1} |t_\ell^n - t_k^n| + 2^{nq} \sum_{t \in J^n(\omega) \cap (t_k^n, t_\ell^n]} |\Delta \varphi_t(\omega)|^q$$

$$\lesssim 2^n |t_\ell^n - t_k^n| + 2^{nq} \|\varphi(\omega)\|_{q,[t_k^n, t_\ell^n]}^q \leqslant 2^{n\rho} c(t_k^n, t_\ell^n),$$

where c is the control function defined by $c(s,t) := |t-s| + \|\varphi(\omega)\|_{q,[s,t]}^q$ for $(s,t) \in \Delta_T$. If $2^{-n} \geqslant c(t_k^n, t_\ell^n)^{\frac{4}{p(1-\varepsilon)}}$, this implies that

$$\left|\int_{t_k^n}^{t_\ell^n} W_u^n \otimes \mathrm{d}W_u - W_{t_k^n} \otimes W_{t_k^n, t_\ell^n}\right| \lesssim 2^{n(\rho-2)} c(t_k^n, t_\ell^n) \leqslant c(t_k^n, t_\ell^n)^{1-\frac{4}{p(1-\varepsilon)}(\rho-2)} = c(t_k^n, t_\ell^n)^{\frac{2}{p}}.$$

In the case that $2^{-n} < c(t_k^n, t_\ell^n)^{\frac{4}{p(1-\varepsilon)}}$, we can follow the same argument as in Step 3 of the proof of part (i) of Proposition 3.6 (using in particular the bound in (3.17)) to obtain again that

$$\left|\int_{t_k^n}^{t_\ell^n} W_u^n \otimes \mathrm{d}W_u - W_{t_k^n} \otimes W_{t_k^n, t_\ell^n}\right| \lesssim c(t_k^n, t_\ell^n)^{\frac{2}{p}},$$

where, as usual, the implicit multiplicative constant depends on ω, but not on n.

It follows that there exists a random control function w such that

$$\sup_{(s,t) \in \Delta_T} \frac{|W_{s,t}|^p}{w(s,t)} + \sup_{n \in \mathbb{N}} \sup_{0 \leqslant k < \ell} \frac{|\int_{t_k^n}^{t_\ell^n} W_u^n \otimes \mathrm{d}W_u - W_{t_k^n} \otimes W_{t_k^n, t_\ell^n}|^{\frac{p}{2}}}{w(s,t)} \leqslant 1$$

holds almost surely. □

Proof of Proposition 3.8. Let W be a Brownian motion with covariance matrix Σ, and let φ be the process defined in (3.15), so that $L = W + \varphi$. As usual, we let L^n, W^n, and φ^n denote the piecewise constant approximations of L, W, and φ, respectively, along the adapted partition $\mathcal{P}_L^n$.

Recalling (3.15), we see that we can write $\varphi = \eta + \xi$, where

$$\eta_t := \lambda t + \int_{|x| \geqslant 2^{-n}} x\,\mu(t, \mathrm{d}x) - t \int_{2^{-n} \leqslant |x| < 1} x\,\nu(\mathrm{d}x) \tag{3.18}$$

and

$$\xi_t := \int_{|x| < 2^{-n}} x(\mu(t, \mathrm{d}x) - t\,\nu(\mathrm{d}x)).$$

Let η^n and ξ^n denote the piecewise constant approximations of η and ξ along $\mathcal{P}^n_L$. Recalling how the adapted partition $\mathcal{P}^n_L$ was defined in (3.16), we note that, when estimating the difference $\eta^n - \eta$, we may ignore all jumps of size greater than 2^{-n}, and may thus ignore the first integral on the right-hand side of (3.18). We then have that

$$\begin{aligned}\|\eta^n - \eta\|_\infty &\leqslant 2^{-n}T|\lambda| + 2^{-n}T\int_{2^{-n}\leqslant|x|<1}|x|\,\nu(\mathrm{d}x)\\ &\leqslant 2^{-n}T|\lambda| + 2^{-n(2-q)}T\int_{2^{-n}\leqslant|x|<1}|x|^q\,\nu(\mathrm{d}x) \lesssim 2^{-n(2-q)}.\end{aligned} \tag{3.19}$$

Writing $\langle\cdot\rangle$ for the predictable quadratic variation, we have (see, e.g., [26, Chapter 2, Theorem 1.33]) that

$$\mathbb{E}[\langle\xi\rangle_T] \leqslant T\int_{|x|<2^{-n}}|x|^2\,\nu(\mathrm{d}x) \leqslant 2^{-n(2-q)}T\int_{|x|<2^{-n}}|x|^q\,\nu(\mathrm{d}x).$$

Since this quantity is finite, the process ξ is a square integrable martingale, and in particular $\mathbb{E}[[\xi]_T] = \mathbb{E}[\langle\xi\rangle_T]$, where $[\cdot]$ denotes the usual quadratic variation. By the Burkholder–Davis–Gundy inequality, we then have that

$$\mathbb{E}[\|\xi\|_\infty^2] \lesssim \mathbb{E}[[\xi]_T] = \mathbb{E}[\langle\xi\rangle_T] \lesssim 2^{-n(2-q)}. \tag{3.20}$$

Note that, for any $a > 0$, if $\|\xi\|_\infty < \frac{a}{2}$, then $\|\xi^n - \xi\|_\infty < a$. It follows that, for any $\delta \in (0, 1 - \frac{q}{2})$,

$$\mathbb{P}(\|\xi^n - \xi\|_\infty \geqslant 2^{-n\delta}) \leqslant \mathbb{P}(\|\xi\|_\infty \geqslant 2^{-1-n\delta}).$$

By Markov's inequality and the bound in (3.20), we see that

$$\mathbb{P}(\|\xi^n - \xi\|_\infty \geqslant 2^{-n\delta}) \lesssim 2^{2-n(2-q-2\delta)},$$

and the Borel–Cantelli lemma then implies that, almost surely,

$$\|\xi^n - \xi\|_\infty \lesssim 2^{-n\delta}, \tag{3.21}$$

where the implicit multiplicative constant is a random variable that does not depend on n. It follows from (3.19) and (3.21) that

$$\|\varphi^n - \varphi\|_\infty \lesssim 2^{-n\delta}. \tag{3.22}$$

Let $p' \in (p, 3)$ and $q' \in (q, 2)$ such that $\frac{1}{p'} + \frac{1}{q'} > 1$. Using interpolation, the fact that $\sup_{n\in\mathbb{N}}\|\varphi^n\|_q \leqslant \|\varphi\|_q$, and the bound in (3.22), we have that, almost surely,

$$\|\varphi^n - \varphi\|_{q'} \leqslant \|\varphi^n - \varphi\|_\infty^{1-\frac{q}{q'}}\|\varphi^n - \varphi\|_q^{\frac{q}{q'}} \lesssim \|\varphi^n - \varphi\|_\infty^{1-\frac{q}{q'}} \lesssim 2^{-n\delta(1-\frac{q}{q'})}. \tag{3.23}$$

We also have from Lemma 3.9 that, for almost every $\omega \in \Omega$, the sample path $W(\omega)$ satisfies Property (RIE) relative to p and $(\mathcal{P}^n_L(\omega))_{n\in\mathbb{N}}$. Thus, by Proposition 2.12, for almost every $\omega \in \Omega$, the sample path $L(\omega)$ satisfies Property (RIE) relative to p and $(\mathcal{P}^n(\omega))_{n\in\mathbb{N}}$, which establishes part (i).

Since the sample paths of W are almost surely $\frac{1}{p}$-Hölder continuous, it is straightforward to see that

$$\|W^n - W\|_\infty \lesssim 2^{-\frac{n}{p}},$$

where the implicit multiplicative constant depends on the (random) Hölder constant of the path. Since $L = W + \varphi$, we have that

$$\|L^n - L\|_\infty \leqslant \|W^n - W\|_\infty + \|\varphi^n - \varphi\|_\infty \lesssim 2^{-\frac{n}{p}} + 2^{-n\delta}.$$

We recall from (3.17) that

$$\left\|\int_0^\cdot W^n_u \otimes \mathrm{d}W_u - \int_0^\cdot W_u \otimes \mathrm{d}W_u\right\|_\infty \lesssim 2^{-\frac{n}{2}(1-\varepsilon)} = 2^{-n(\frac{1}{p}-\gamma)}$$

for any $\gamma \in (0, \frac{1}{p})$. We obtained a bound for $\|\varphi^n - \varphi\|_{q'}$ in (3.23), and an analogous argument also shows that

$$\|W^n - W\|_{p'} \leqslant \|W^n - W\|_\infty^{1-\frac{p}{p'}} \|W^n - W\|_p^{\frac{p}{p'}} \lesssim \|W^n - W\|_\infty^{1-\frac{p}{p'}} \lesssim 2^{-n(\frac{1}{p}-\frac{1}{p'})}.$$

Using the standard estimate for Young integrals (see, e.g., [22, Proposition 2.4]), similarly to the proof of Proposition 2.12, we then obtain

$$\begin{aligned}
&\left\|\int_0^\cdot L^n_u \otimes \mathrm{d}L_u - \int_0^\cdot L_u \otimes \mathrm{d}L_u\right\|_\infty \\
&\lesssim \left\|\int_0^\cdot W^n_u \otimes \mathrm{d}W_u - \int_0^\cdot W_u \otimes \mathrm{d}W_u\right\|_\infty + \|W^n - W\|_{p'}\|\varphi\|_q + \|\varphi^n - \varphi\|_{q'}(\|W\|_p + \|\varphi\|_q) \\
&\lesssim 2^{-n(\frac{1}{p}-\gamma)} + 2^{-n(\frac{1}{p}-\frac{1}{p'})} + 2^{-n\delta(1-\frac{q}{q'})}.
\end{aligned}$$

Hence, by Theorem 2.3, we establish the estimate in part (ii). □

In the following remark, we briefly discuss α-stable Lévy processes.

Remark 3.10. Suppose now that L is an α-stable Lévy process for some $\alpha \in (0, 2]$. That is, for all $a > 0$, there exists $c \in \mathbb{R}^d$ such that

$$(L_{at})_{t\in[0,T]} \overset{d}{=} (a^{\frac{1}{\alpha}} L_t + ct)_{t\in[0,T]},$$

where we write $X \overset{d}{=} Y$ to mean that X and Y have the same distribution; see, for example, [10, Proposition 3.15]. We now distinguish two cases:

In the case when $\alpha = 2$, L is α-stable if and only if it is Gaussian, that is, its characteristics are given by $(\lambda, \Sigma, 0)$; see, for example, [10, Proposition 3.15]. It can thus be decomposed into the sum of a Brownian motion W with covariance matrix Σ, and a linear drift term: $L_t = W_t + \lambda t$, for $t \in [0, T]$. In this case, the SDE (3.1) driven by L can therefore be reformulated as an SDE driven by W by simply absorbing the linear drift term λt into the drift of the SDE, and the resulting equation can then be treated as in Corollary 3.5.

In the case when $\alpha \in (0, 2)$, L is α-stable if and only if its characteristics are given by $(\lambda, 0, \nu)$ (i.e., $L = \varphi$ for some φ of the form in (3.15)), and there exists a finite measure ρ on S, a unit sphere on $\mathbb{R}^d$, such that

$$\nu(B) = \int_S \int_0^\infty \mathbf{1}_B(r\xi) \frac{\mathrm{d}r}{r^{1+\alpha}} \rho(\mathrm{d}\xi)$$

for all Borel sets B on $\mathbb{R}^d$; see, for example, [10, Proposition 3.15].

We then have that $\int_{|x|<1} |x|^q \nu(\mathrm{d}x) < \infty$ for $q > \alpha$, and in particular that almost all sample paths of L are of finite q-variation for $q \in (\alpha, 2)$ if $\alpha \in [1, 2)$, and are of finite 1-variation if $\alpha < 1$. This then fits into the setting of Proposition 3.8, and, since there is no Gaussian term, the resulting error estimate for the associated Euler scheme reduces to

$$\|Y^n - Y\|_{p'} \leqslant C\left(2^{-n\left(1-\frac{1}{q'}\right)} + 2^{-n\delta\left(1-\frac{q}{q'}\right)\left(1-\frac{p}{p'}\right)}\right), \qquad n \in \mathbb{N},$$

Of course, in this case, it is not necessary to utilize the rough path framework, since the integral $\int_0^t \sigma(s, Y_{s-})\,\mathrm{d}L_s$ in (3.1) can be defined as a pathwise Young integral, and by discretizing this integral, one could derive pathwise results using stability estimates for Young integrals.

3.4 | Càdlàg semimartingales

In this section, we consider the case when X is a general càdlàg semimartingale. As noted in Remark 3.7, to hope for pointwise convergence of the Euler scheme, we need to ensure that the sequence of partitions exhausts all the jump times of X. With this in mind, for each $n \in \mathbb{N}$, we introduce the stopping times $(\tau_k^n)_{k \in \mathbb{N} \cup \{0\}}$, such that $\tau_0^n = 0$, and

$$\tau_k^n = \inf\{t > \tau_{k-1}^n \,:\, |t - \tau_{k-1}^n| + |X_t - X_{\tau_{k-1}^n}| \geqslant 2^{-n}\} \wedge T, \qquad k \in \mathbb{N}. \tag{3.24}$$

We then define a sequence of adapted partitions $(\mathcal{P}_X^n)_{n \in \mathbb{N}}$ by

$$\mathcal{P}_X^n = \{\tau_k^n \,:\, k \in \mathbb{N} \cup \{0\}\}.$$

Note that, for almost every $\omega \in \Omega$, $(\mathcal{P}_X^n(\omega))_{n \in \mathbb{N}}$ is a sequence of (finite) partitions with vanishing mesh size. The next result verifies that X satisfies Property (RIE) relative to any $p \in (2, 3)$ and $(\mathcal{P}_X^n)_{n \in \mathbb{N}}$, and establishes the rate of convergence of the associated Euler scheme.

Proposition 3.11. *Let $p \in (2, 3)$, and let X be a d-dimensional càdlàg semimartingale. Let Y be the solution of the SDE (3.1) driven by X, and let Y^n be the corresponding Euler approximation along $\mathcal{P}_X^n$, as defined in (2.5).*

(i) *For almost every $\omega \in \Omega$, the sample path $X(\omega)$ satisfies Property (RIE) relative to p and $(\mathcal{P}_X^n(\omega))_{n\in\mathbb{N}}$.*
(ii) *For any $p' \in (p,3)$ and $q \in (1,2)$ such that $\frac{1}{p'} + \frac{1}{q} > 1$, and any $\varepsilon \in (0,1)$, there exists a random variable C, which does not depend on n, such that*

$$\|Y^n - Y\|_{p'} \leqslant C(2^{-n(1-\frac{1}{q})} + 2^{-n(1-\varepsilon)(1-\frac{p}{p'})}), \qquad n \in \mathbb{N}, \tag{3.25}$$

and

$$\|Y^n - Y\|_3 \leqslant C2^{-n(\frac{1}{3}-\varepsilon)}, \qquad n \in \mathbb{N}. \tag{3.26}$$

Proof.

(i) The proof is just a slight modification of the proof of [2, Proposition 4.1], and is therefore omitted here for brevity. It is actually slightly easier, as here we do not require the sequence of partitions to be nested, and the sequence of stopping times in (3.24) is constructed to ensure that the mesh size vanishes, even if X exhibits intervals of constancy.
(ii) By the definition of the partition $\mathcal{P}_X^n$, it is clear that

$$\|X^n - X\|_\infty \leqslant 2^{-n}.$$

By an application of the Burkholder–Davis–Gundy inequality and the Borel–Cantelli lemma, as in the proof of [31, Proposition 3.4], one can show that

$$\left\| \int_0^\cdot X^n_{u-} \otimes \mathrm{d}X_u - \int_0^\cdot X_{u-} \otimes \mathrm{d}X_u \right\|_\infty \lesssim 2^{-n(1-\varepsilon)}, \qquad n \in \mathbb{N},$$

where the implicit multiplicative constant is a random variable that does not depend on n.
It thus follows from Theorem 2.3 that

$$\|Y^n - Y\|_3 \leqslant \|Y^n - Y\|_{p'} \lesssim 2^{-n(1-\frac{1}{q})} + 2^{-n(1-\frac{p}{p'})} + 2^{-n(1-\varepsilon)(1-\frac{p}{p'})},$$

which leads to (3.25). Choosing p sufficiently close to 2, p' to 3, and q to $\frac{3}{2}$, and replacing ε by 3ε, then reveals (3.26). □

4 | APPLICATIONS TO DIFFERENTIAL EQUATIONS DRIVEN BY NON-SEMIMARTINGALES

While in the previous section, we considered SDEs driven by various classes of semimartingales, like the general theory of rough paths, the deterministic theory developed in Section 2 is not limited to the semimartingale framework. In this section, we investigate Property (RIE) in the context of "mixed" and "rough" SDEs. The main insight is again that the random driving signals of these equations do, indeed, satisfy Property (RIE), and thus, the pathwise convergence results regarding the Euler scheme, as presented in Theorem 2.3 and Proposition 2.13, are applicable.

Further examples of stochastic processes that fulfill Property (RIE) almost surely include *p-semimartingales* (also known as *Young semimartingales*) in the sense of Norvaiša [34], as well as *typical price paths* in the sense of Vovk, relative to suitable sequences of adapted partitions. The pathwise convergence of the Euler scheme is thus immediately applicable to differential equations driven by such p-semimartingales [30] and typical price paths [5].

4.1 | Mixed stochastic differential equations

Differential equations driven by both a Brownian motion as well as a fractional Brownian motion with Hurst parameter $H > \frac{1}{2}$ are classical objects in stochastic analysis; see, for example, [33, 41]. More precisely, a "mixed" stochastic differential equation (mixed SDE) is given by

$$Y_t = y_0 + \int_0^t b(s, Y_s)\,\mathrm{d}s + \int_0^t \sigma_1(s, Y_s)\,\mathrm{d}W_s + \int_0^t \sigma_2(s, Y_s)\,\mathrm{d}W_s^H, \qquad t \in [0, T], \tag{4.1}$$

where $b \in C_b^2(\mathbb{R}^{k+1}; \mathbb{R}^k)$, $\sigma_1 \in C_b^3(\mathbb{R}^{k+1}; \mathcal{L}(\mathbb{R}^{d_1}; \mathbb{R}^k))$, $\sigma_2 \in C_b^3(\mathbb{R}^{k+1}; \mathcal{L}(\mathbb{R}^{d_2}; \mathbb{R}^k))$ and $y_0 \in \mathbb{R}^k$. Here, W is a d_1-dimensional standard Brownian motion, and W^H is a d_2-dimensional fractional Brownian motion with Hurst parameter $H > \frac{1}{2}$, which are independent and both defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P})$ satisfying the usual conditions.

The mixed SDE (4.1) lies outside the semimartingale framework, but there are various ways to provide a rigorous meaning to its solution. Here, we consider the mixed SDE (4.1) as a random RDE, driven by the Itô rough path lift of (W, W^H), the existence of which follows from Lemma 4.1 below. In particular, it then follows from Theorem 2.2 that there exists a unique solution Y to (4.1).

Lemma 4.1. *Let W be a standard Brownian motion, and let W^H be a fractional Brownian motion with Hurst parameter $H \in (\frac{1}{2}, 1)$. Let $p \in (2, 3)$ such that $\frac{1}{p} + H > 1$, and let $\mathcal{P}^n = \{0 = t_0^n < t_1^n < \cdots < t_{N_n}^n = T\}$, $n \in \mathbb{N}$, be a sequence of equidistant partitions of the interval $[0, T]$, so that, for each $n \in \mathbb{N}$, there exists some $\pi_n > 0$ such that $t_{i+1}^n - t_i^n = \pi_n$ for each $0 \leqslant i < N_n$. If $\pi_n^{2-\frac{4}{p}} \log(n) \to 0$ as $n \to \infty$, then, for almost every $\omega \in \Omega$, the sample path $(W(\omega), W^H(\omega))$ satisfies Property (RIE) relative to p and $(\mathcal{P}^n)_{n\in\mathbb{N}}$.*

Proof. We first note that the process $(W, 0)$ satisfies the hypotheses of Theorem 3.2, and thus that almost all of its sample paths satisfy Property (RIE) relative to p and $(\mathcal{P}^n)_{n\in\mathbb{N}}$. Let $\frac{1}{H} < q < q' < 2$ such that $\frac{1}{p} + \frac{1}{q'} > 1$. Since $\frac{1}{q} < H$, it is well known that the sample paths of $(0, W^H)$ are almost surely $\frac{1}{q}$-Hölder continuous, and hence that $\|W^H\|_q < \infty$. Writing $W^{H,n}$ for the usual piecewise constant approximation of W^H along $\mathcal{P}^n$, we have by interpolation that

$$\|W^{H,n} - W^H\|_{q'} \leqslant \|W^{H,n} - W^H\|_\infty^{1-\frac{q}{q'}} \|W^{H,n} - W^H\|_q^{\frac{q}{q'}} \lesssim \|W^{H,n} - W^H\|_\infty^{1-\frac{q}{q'}} \longrightarrow 0$$

as $n \to \infty$. The result then follows by applying Proposition 2.12 to $(W, 0) + (0, W^H)$. □

Of course, since here we consider Hurst parameters $H > \frac{1}{2}$, the trajectories of W^H have in particular finite q-variation for any $q \in (\frac{1}{H}, 2)$, so we could alternatively define the integral $\int_0^t \sigma_2(s, Y_s)\,\mathrm{d}W_s^H$ in (4.1) as a pathwise Young integral, and by discretizing this integral one could in principle derive analogous pathwise convergence results; cf. Remark 2.15.

4.2 Rough stochastic differential equations

Rough stochastic differential equations (rough SDEs) are differential equations driven by both a rough path and a semimartingale. These equations first appeared in the context of robust stochastic filtering — see [11, 15] — and were recently studied in a general form in [18]. In this section, we will adapt the setting of [15], which allows to treat Hölder continuous rough paths and Brownian motion as driving signals.

We let $\eta\colon [0,T] \to \mathbb{R}^d$ be a deterministic path that is $\frac{1}{p}$-Hölder continuous for some $p \in (2,3)$, and which satisfies Property (RIE) relative to p and the dyadic partitions $(\mathcal{P}^n_D)_{n\in\mathbb{N}}$, as defined in (3.9). We write $\boldsymbol{\eta} = (\boldsymbol{\eta}^1, \boldsymbol{\eta}^2)$ for the canonical rough path lift of η, with $\boldsymbol{\eta}^2$ defined as in (2.7), so that $\boldsymbol{\eta}^2_{s,t} = \int_s^t \eta_{s,u} \otimes \mathrm{d}\eta_u$ for each $(s,t) \in \Delta_T$. We also let W be an $\mathbb{R}^e$-valued Brownian motion. For vector fields $a \in C^2_b(\mathbb{R}^k;\mathbb{R}^k)$, $b \in C^3_b(\mathbb{R}^k;\mathcal{L}(\mathbb{R}^d;\mathbb{R}^k))$ and $c \in C^3_b(\mathbb{R}^k;\mathcal{L}(\mathbb{R}^e;\mathbb{R}^k))$, and an initial value $y_0 \in \mathbb{R}^k$, we then consider the rough SDE

$$Y_t = y_0 + \int_0^t a(Y_s)\,\mathrm{d}s + \int_0^t b(Y_s)\,\mathrm{d}\boldsymbol{\eta}_s + \int_0^t c(Y_s)\,\mathrm{d}W_s, \qquad t \in [0,T]. \tag{4.2}$$

To give a rigorous meaning to the rough SDE (4.2), following the method introduced in [15], we need to construct a suitable joint rough path lift $\boldsymbol{\Lambda}(\omega)$ above the $\mathbb{R}^{d+e}$-valued path $(\eta, W(\omega))$ for almost every $\omega \in \Omega$. Indeed, the (pathwise) unique solution to the random RDE

$$Y_t = y_0 + \int_0^t a(Y_s)\,\mathrm{d}s + \int_0^t (b,c)(Y_s)\,\mathrm{d}\boldsymbol{\Lambda}_s, \qquad t \in [0,T],$$

is then defined to be the solution to the rough SDE (4.2).

To construct the Itô rough path lift of (η, W), we need the existence of the quadratic covariation of η and W along the dyadic partitions. More precisely, writing $\mathcal{P}^n_D = \{0 = t^n_0 < t^n_1 < \cdots < t^n_{2^n} = T\}$ with $t^n_k = k2^{-n}T$, we need to establish that, for almost every $\omega \in \Omega$, the limit

$$\langle \eta, W(\omega)\rangle_t := \lim_{n\to\infty} \sum_{k=0}^{2^n-1} \eta_{t^n_k\wedge t, t^n_{k+1}\wedge t} \otimes W_{t^n_k\wedge t, t^n_{k+1}\wedge t}(\omega) \tag{4.3}$$

exists and holds uniformly for $t \in [0,T]$.

Lemma 4.2. *Let $\alpha \in (0,1]$, let $\eta\colon [0,T] \to \mathbb{R}$ be an α-Hölder continuous deterministic path, and let W be a one-dimensional Brownian motion. Then, for almost every $\omega \in \Omega$, the quadratic covariation of η and $W(\omega)$ along the dyadic partitions, in the sense of (4.3), exists, and satisfies $\langle \eta, W(\omega)\rangle_t = 0$ for all $t \in [0,T]$.*

Proof. We consider the discrete-time martingale given by $t \mapsto \sum_{k\,:\,t^n_{k+1}\leqslant t} \eta_{t^n_k,t^n_{k+1}} W_{t^n_k,t^n_{k+1}}$ for $t \in \mathcal{P}^n_D$, for some fixed $n \in \mathbb{N}$. By the Burkholder–Davis–Gundy inequality, we have that

$$\mathbb{E}\left[\left\|\sum_{k\,:\,t^n_{k+1}\leqslant \cdot} \eta_{t^n_k,t^n_{k+1}} W_{t^n_k,t^n_{k+1}}\right\|_\infty^2\right] \lesssim \mathbb{E}\left[\sum_{k=0}^{2^n-1} (\eta_{t^n_k,t^n_{k+1}} W_{t^n_k,t^n_{k+1}})^2\right] = \sum_{k=0}^{2^n-1} (\eta_{t^n_k,t^n_{k+1}})^2 (t^n_{k+1} - t^n_k)$$

$$\lesssim \sum_{k=0}^{2^n-1} (t^n_{k+1} - t^n_k)^{1+2\alpha} \lesssim (2^{-n}T)^{2\alpha} \sum_{k=0}^{2^n-1} (t^n_{k+1} - t^n_k) \lesssim 2^{-2n\alpha}.$$

For any $\varepsilon \in (0,1)$, we then have, by Markov's inequality, that

$$\mathbb{P}\left(\left\|\sum_{k\,:\,t_{k+1}^n \leqslant \cdot} \eta_{t_k^n,t_{k+1}^n} W_{t_k^n,t_{k+1}^n}\right\|_\infty \geqslant 2^{-n\alpha(1-\varepsilon)}\right) \lesssim 2^{-2n\alpha\varepsilon},$$

and the Borel–Cantelli lemma then implies that

$$\left\|\sum_{k\,:\,t_{k+1}^n \leqslant \cdot} \eta_{t_k^n,t_{k+1}^n} W_{t_k^n,t_{k+1}^n}\right\|_\infty \lesssim 2^{-n\alpha(1-\varepsilon)},$$

where the implicit multiplicative constant is a random variable that does not depend on n.

For a given $t \in [0,T]$ and $n \in \mathbb{N}$, let k_0 be such that $t \in [t_{k_0}^n, t_{k_0+1}^n]$. Since η is α-Hölder continuous, and the sample paths of W are almost surely β-Hölder continuous for any $\beta \in (0,\frac{1}{2})$, we have that

$$|\eta_{t_{k_0}^n,t} W_{t_{k_0}^n,t}| \lesssim (t - t_{k_0}^n)^{\alpha+\beta} \lesssim 2^{-n(\alpha+\beta)}.$$

We thus have the bound

$$\begin{aligned}\left|\sum_{k=0}^{2^n-1} \eta_{t_k^n\wedge t,t_{k+1}^n\wedge t} W_{t_k^n\wedge t,t_{k+1}^n\wedge t}\right| &\leqslant \left|\sum_{k\,:\,t_{k+1}^n \leqslant t} \eta_{t_k^n,t_{k+1}^n} W_{t_k^n,t_{k+1}^n}\right| + |\eta_{t_{k_0}^n,t} W_{t_{k_0}^n,t}| \\ &\lesssim 2^{-n\alpha(1-\varepsilon)} + 2^{-n(\alpha+\beta)},\end{aligned}$$

where the implicit multiplicative constant is a random variable that does not depend on t or n. It follows that, almost surely,

$$\sum_{k=0}^{2^n-1} \eta_{t_k^n\wedge t,t_{k+1}^n\wedge t} W_{t_k^n\wedge t,t_{k+1}^n\wedge t} \longrightarrow 0 \qquad \text{as} \quad n \longrightarrow \infty,$$

uniformly for $t \in [0,T]$. □

It is shown in [15, Theorem 1], with integrals defined in the Stratonovich sense, that an analogous object to the process $\boldsymbol{\Lambda}$ described in (4.4) below provides a geometric rough path lift of (η, W). In the next theorem, we establish that $\boldsymbol{\Lambda}$ is the Itô rough path lift of (η, W), and, moreover, that it may be obtained as the canonical lift via Property (RIE), thus making our convergence analysis of the Euler scheme applicable to the rough SDE (4.2).

Theorem 4.3. *Let $p \in (2,3)$. Let η be a $\frac{1}{p}$-Hölder continuous $\mathbb{R}^d$-valued path that satisfies Property (RIE) relative to p and the sequence of dyadic partitions $(\mathcal{P}_D^n)_{n\in\mathbb{N}}$, and write $\boldsymbol{\eta} = (\boldsymbol{\eta}^1, \boldsymbol{\eta}^2)$ for the canonical rough path lift of η, so that $\boldsymbol{\eta}^1 = \eta$, and $\boldsymbol{\eta}^2_{s,t} = \int_s^t \eta_{s,u} \otimes \mathrm{d}\eta_u$, defined as in (2.7), for every $(s,t) \in \Delta_T$. Let W be an $\mathbb{R}^e$-valued Brownian motion, and write $\mathbf{W} = (W, \mathbb{W})$ for the Itô rough path lift of W, so that $\mathbb{W}_{s,t} = \int_s^t W_{s,u} \otimes \mathrm{d}W_u$, defined as an Itô integral, for every $(s,t) \in \Delta_T$.*

For any $p' \in (p, 3)$ and almost every $\omega \in \Omega$, the $\mathbb{R}^{d+e}$-valued path $(\eta, W(\omega))$ satisfies Property (RIE) relative to p' and $(\mathcal{P}^n_D)_{n \in \mathbb{N}}$.

Moreover, for almost every $\omega \in \Omega$, the canonical rough path lift $\mathbf{\Lambda}(\omega) = (\mathbf{\Lambda}^1(\omega), \mathbf{\Lambda}^2(\omega)) \in \mathbb{R}^{d+e} \oplus \mathbb{R}^{(d+e)\times(d+e)}$ of $(\eta, W(\omega))$ (constructed via Property (RIE) as in (2.7)) is given by $\mathbf{\Lambda}^1(\omega) = (\eta, W(\omega))$, and

$$\mathbf{\Lambda}^2_{s,t} = \begin{pmatrix} \boldsymbol{\eta}^2_{s,t} & \int_s^t \eta_{s,u} \otimes \mathrm{d}W_u \\ W_{s,t} \otimes \eta_{s,t} - (\int_s^t \eta_{s,u} \otimes \mathrm{d}W_u)^\top & \mathbb{W}_{s,t} \end{pmatrix} \tag{4.4}$$

for every $(s,t) \in \Delta_T$, where $\int_s^t \eta_{s,u} \otimes \mathrm{d}W_u$ is defined as an Itô integral, and $(\cdot)^\top$ denotes matrix transposition.

Proof. Let $p' \in (p, 3)$. It follows from the Kolmogorov criterion for rough paths (see [17, Theorem 3.1]) that, for almost every $\omega \in \Omega$,

$$\left| \left(\int_s^t \eta_{s,u} \otimes \mathrm{d}W_u \right)(\omega) \right| \lesssim |t-s|^{\frac{2}{p'}} \qquad \text{for all} \quad (s,t) \in \Delta_T, \tag{4.5}$$

and moreover, that $\mathbf{\Lambda}(\omega) = (\mathbf{\Lambda}^1(\omega), \mathbf{\Lambda}^2(\omega))$ is a $\frac{1}{p'}$-Hölder continuous rough path. We will show that $(\eta, W(\omega))$ satisfies Property (RIE), and that the associated canonical rough path is indeed given by $\mathbf{\Lambda}(\omega)$.

Step 1. As usual, we let η^n and W^n denote the piecewise constant approximations of η and W, respectively, along $\mathcal{P}^n_D$. By assumption, η satisfies Property (RIE) relative to p and $(\mathcal{P}^n_D)_{n \in \mathbb{N}}$. By Proposition 3.2 (or Proposition 3.6), for almost every $\omega \in \Omega$, the sample path $W(\omega)$ also satisfies Property (RIE) relative to p and $(\mathcal{P}^n_D)_{n \in \mathbb{N}}$.

It follows from the first condition in Property (RIE) for η and $W(\omega)$ that, for almost every $\omega \in \Omega$,

$$(\eta^n, W^n(\omega)) \longrightarrow (\eta, W(\omega)) \qquad \text{uniformly as} \quad n \longrightarrow \infty,$$

so that this condition also holds for the pair $(\eta, W(\omega))$. Moreover, it follows from the second condition in Property (RIE) that $\int_0^\cdot \eta^n_u \otimes \mathrm{d}\eta_u$ converges uniformly to $\int_0^\cdot \eta_u \otimes \mathrm{d}\eta_u$, and, for almost every $\omega \in \Omega$, that $(\int_0^\cdot W^n_u \otimes \mathrm{d}W_u)(\omega)$ converges uniformly to $(\int_0^\cdot W_u \otimes \mathrm{d}W_u)(\omega)$.

By the Burkholder–Davis–Gundy inequality, and the observation that $\|\eta^n - \eta\|_\infty \lesssim 2^{-\frac{n}{p}}$, we have that

$$\mathbb{E}\left[\left\| \int_0^\cdot \eta^n_u \otimes \mathrm{d}W_u - \int_0^\cdot \eta_u \otimes \mathrm{d}W_u \right\|_\infty^2 \right] \lesssim \mathbb{E}\left[\int_0^T |\eta^n_u - \eta_u|^2 \,\mathrm{d}u \right] \lesssim 2^{-\frac{2n}{p}}.$$

For any $\varepsilon \in (1 - \frac{2}{p}, 1)$, it then follows from Markov's inequality that

$$\mathbb{P}\left(\left\| \int_0^\cdot \eta^n_u \otimes \mathrm{d}W_u - \int_0^\cdot \eta_u \otimes \mathrm{d}W_u \right\|_\infty \geqslant 2^{-\frac{n}{2}(1-\varepsilon)} \right) \lesssim 2^{n(1-\frac{2}{p}-\varepsilon)}.$$

The Borel–Cantelli lemma then implies that, for almost every $\omega \in \Omega$,

$$\left\| \left(\int_0^\cdot \eta^n_u \otimes \mathrm{d}W_u - \int_0^\cdot \eta_u \otimes \mathrm{d}W_u \right)(\omega) \right\|_\infty \lesssim 2^{-\frac{n}{2}(1-\varepsilon)} \tag{4.6}$$

for all $n \in \mathbb{N}$, and, in particular, that $(\int_0^\cdot \eta^n_u \otimes \mathrm{d}W_u)(\omega)$ converges uniformly to $(\int_0^\cdot \eta_u \otimes \mathrm{d}W_u)(\omega)$ as $n \to \infty$.

Let us write $\mathcal{P}^n_D = \{0 = t^n_0 < t^n_1 < \cdots < t^n_{2^n} = T\}$ for $n \in \mathbb{N}$, where $t^n_k = k2^{-n}T$. It is straightforward to verify that, for any $t \in [0,T]$,

$$W_t \otimes \eta_t = \int_0^t W^n_u \otimes \mathrm{d}\eta_u + \left(\int_0^t \eta^n_u \otimes \mathrm{d}W_u \right)^\top + \langle W, \eta \rangle^n_t,$$

where, by Lemma 4.2, the discrete quadratic variation $\langle W, \eta \rangle^n_t := \sum_{k=0}^{2^n-1} W_{t^n_k \wedge t, t^n_{k+1} \wedge t} \otimes \eta_{t^n_k \wedge t, t^n_{k+1} \wedge t}$ almost surely converges uniformly to $\langle W, \eta \rangle_t = 0$ as $n \to \infty$. We then see that, for almost every $\omega \in \Omega$,

$$\int_0^t W^n_u(\omega) \otimes \mathrm{d}\eta_u \longrightarrow W_t(\omega) \otimes \eta_t - \left(\int_0^t \eta_u \otimes \mathrm{d}W_u \right)^\top (\omega)$$

as $n \to \infty$, uniformly in $t \in [0,T]$. We have thus established that, for almost every $\omega \in \Omega$, the path $(\eta, W(\omega))$ also satisfies the second condition of Property (RIE), and moreover, that the resulting canonical rough path is indeed given by (4.4).

Step 2. It remains to show that $(\eta, W(\omega))$ satisfies the third condition of Property (RIE) relative to p' and $(\mathcal{P}^n_D)_{n \in \mathbb{N}}$.

Since η satisfies Property (RIE) relative to p and $(\mathcal{P}^n_D)_{n \in \mathbb{N}}$, there exists a control function w_η such that

$$\sup_{(s,t) \in \Delta_T} \frac{|\eta_{s,t}|^p}{w_\eta(s,t)} + \sup_{n \in \mathbb{N}} \sup_{0 \leqslant k < \ell \leqslant 2^n} \frac{| \int_{t^n_k}^{t^n_\ell} \eta^n_u \otimes \mathrm{d}\eta_u - \eta_{t^n_k} \otimes \eta_{t^n_k, t^n_\ell} |^{\frac{p}{2}}}{w_\eta(t^n_k, t^n_\ell)} \leqslant 1, \tag{4.7}$$

which implies that the same inequality also holds with p replaced by p' (possibly with a different control function, but without loss of generality, we may assume that w_η remains valid for p'). Similarly, since for almost every $\omega \in \Omega$, the sample path $W(\omega)$ satisfies Property (RIE) relative to p (and therefore also to p') and $(\mathcal{P}^n_D)_{n \in \mathbb{N}}$, there exists a control function c such that

$$\sup_{(s,t) \in \Delta_T} \frac{|W_{s,t}(\omega)|^{p'}}{c(s,t)} + \sup_{n \in \mathbb{N}} \sup_{0 \leqslant k < \ell \leqslant 2^n} \frac{|(\int_{t^n_k}^{t^n_\ell} W^n_u \otimes \mathrm{d}W_u - W_{t^n_k} \otimes W_{t^n_k, t^n_\ell})(\omega)|^{\frac{p'}{2}}}{c(t^n_k, t^n_\ell)} \leqslant 1. \tag{4.8}$$

Step 3. Let $\beta \in (0, \frac{1}{2})$. Since η is $\frac{1}{p}$-Hölder continuous, and the sample paths of W are almost surely β-Hölder continuous, we have that

$$|\eta_{t^n_{i-1}} \otimes W_{t^n_{i-1}, t^n_i} + \eta_{t^n_i} \otimes W_{t^n_i, t^n_{i+1}} - \eta_{t^n_{i-1}} \otimes W_{t^n_{i-1}, t^n_{i+1}}| = |\eta_{t^n_{i-1}, t^n_i} \otimes W_{t^n_i, t^n_{i+1}}| \lesssim |t^n_{i+1} - t^n_{i-1}|^{\frac{1}{p}+\beta}$$

for any $i = 1, \dots, N_n - 1$, where the implicit multiplicative constant is a random variable, and we can follow the proof of [31, Lemma 3.2] to deduce that, for almost any fixed $\omega \in \Omega$, for any $k < \ell$, and writing $N = \ell - k = 2^n |t^n_\ell - t^n_k| T^{-1}$,

$$\left| \left(\int_{t^n_k}^{t^n_\ell} \eta^n_u \otimes \mathrm{d}W_u \right)(\omega) - \eta_{t^n_k} \otimes W_{t^n_k, t^n_\ell}(\omega) \right| \lesssim N^{1-\frac{2}{\rho}} |t^n_\ell - t^n_k|^{\frac{2}{\rho}} \lesssim 2^{n(1-\frac{2}{\rho})} |t^n_\ell - t^n_k|,$$

where $\frac{2}{\rho} = \frac{1}{p} + \beta$.

Let $\varepsilon \in (1 - \frac{2}{p}, 1)$. If $2^{-n} \geqslant |t^n_\ell - t^n_k|^{\frac{4}{p(1-\varepsilon)}}$, then

$$\left| \left(\int_{t^n_k}^{t^n_\ell} \eta^n_u \otimes \mathrm{d}W_u \right)(\omega) - \eta_{t^n_k} \otimes W_{t^n_k, t^n_\ell}(\omega) \right| \lesssim |t^n_\ell - t^n_k|^{1 - \frac{4}{p(1-\varepsilon)}(1-\frac{2}{\rho})}.$$

By choosing ε close to $1 - \frac{2}{p}$, we can make the above exponent $1 - \frac{4}{p(1-\varepsilon)}(1 - \frac{2}{\rho})$ arbitrarily close to $\frac{4}{\rho} - 1 = \frac{2}{p} + 2\beta - 1$. By then choosing β close to $\frac{1}{2}$, we can make this value arbitrarily close to $\frac{2}{p}$ from below. In particular, by making suitable choices of ε and β, we can ensure that $1 - \frac{4}{p(1-\varepsilon)}(1 - \frac{2}{\rho}) = \frac{2}{p'}$, and we obtain

$$\left| \left(\int_{t^n_k}^{t^n_\ell} \eta^n_u \otimes \mathrm{d}W_u \right)(\omega) - \eta_{t^n_k} \otimes W_{t^n_k, t^n_\ell}(\omega) \right| \lesssim |t^n_\ell - t^n_k|^{\frac{2}{p'}}. \tag{4.9}$$

We will now aim to obtain the same estimate in the case that $2^{-n} < |t^n_\ell - t^n_k|^{\frac{4}{p(1-\varepsilon)}}$, with ε chosen as above. Recalling (4.5) and (4.6), we have that

$$\begin{aligned}
&\left| \left(\int_{t^n_k}^{t^n_\ell} \eta^n_u \otimes \mathrm{d}W_u \right)(\omega) - \eta_{t^n_k} \otimes W_{t^n_k, t^n_\ell}(\omega) \right| \\
&= \left| \left(\int_{t^n_k}^{t^n_\ell} \eta^n_u \otimes \mathrm{d}W_u \right)(\omega) - \left(\int_{t^n_k}^{t^n_\ell} \eta_u \otimes \mathrm{d}W_u \right)(\omega) + \left(\int_{t^n_k}^{t^n_\ell} \eta_u \otimes \mathrm{d}W_u \right)(\omega) - \eta_{t^n_k} \otimes W_{t^n_k, t^n_\ell}(\omega) \right| \\
&\leqslant 2 \left\| \left(\int_0^\cdot \eta^n_u \otimes \mathrm{d}W - \int_0^\cdot \eta_u \otimes \mathrm{d}W \right)(\omega) \right\|_\infty + \left| \left(\int_{t^n_k}^{t^n_\ell} \eta_{t^n_k, u} \otimes \mathrm{d}W_u \right)(\omega) \right| \\
&\lesssim 2^{-\frac{n}{2}(1-\varepsilon)} + |t^n_\ell - t^n_k|^{\frac{2}{p'}} \\
&\lesssim |t^n_\ell - t^n_k|^{\frac{2}{p'}}.
\end{aligned}$$

Combining this with (4.9), we conclude that

$$\sup_{n \in \mathbb{N}} \sup_{0 \leqslant k < \ell \leqslant 2^n} \frac{|(\int_{t^n_k}^{t^n_\ell} \eta^n_u \otimes \mathrm{d}W_u)(\omega) - \eta_{t^n_k} \otimes W_{t^n_k, t^n_\ell}(\omega)|^{\frac{p'}{2}}}{C(\omega) |t^n_\ell - t^n_k|} \leqslant 1, \tag{4.10}$$

for a suitable random variable C.

Step 4. For any $n \in \mathbb{N}$ and $0 \leqslant k < \ell \leqslant 2^n$, it is straightforward to verify that

$$|\eta_{t_k^n,t_\ell^n}|^2 = 2\int_{t_k^n}^{t_\ell^n} \eta^n_{t_k^n,u} \cdot \mathrm{d}\eta_u + \sum_{i=k}^{\ell-1} |\eta_{t_i^n,t_{i+1}^n}|^2,$$

where $\cdot$ denotes the Euclidean inner product. It follows from (4.7) that $|\eta_{t_k^n,t_\ell^n}|^2 \lesssim w_\eta(t_k^n,t_\ell^n)^{\frac{2}{p'}}$, and that

$$\sup_{n\in\mathbb{N}} \sup_{0\leqslant k<\ell\leqslant 2^n} \frac{|\int_{t_k^n}^{t_\ell^n} \eta^n_{t_k^n,u} \cdot \mathrm{d}\eta_u|^{\frac{p'}{2}}}{w_\eta(t_k^n,t_\ell^n)} \lesssim 1,$$

from which we then have that

$$\sup_{n\in\mathbb{N}} \sup_{0\leqslant k<\ell\leqslant 2^n} \frac{|\sum_{i=k}^{\ell-1} |\eta_{t_i^n,t_{i+1}^n}|^2|^{\frac{p'}{2}}}{w_\eta(t_k^n,t_\ell^n)} \lesssim 1.$$

The same argument holds for the sample paths of W, and since

$$\left|\sum_{i=k}^{\ell-1} W_{t_i^n,t_{i+1}^n} \otimes \eta_{t_i^n,t_{i+1}^n}\right| \lesssim \sum_{i=k}^{\ell-1} |W_{t_i^n,t_{i+1}^n}|^2 + \sum_{i=k}^{\ell-1} |\eta_{t_i^n,t_{i+1}^n}|^2,$$

we deduce that

$$\sup_{n\in\mathbb{N}} \sup_{0\leqslant k<\ell\leqslant 2^n} \frac{|\sum_{i=k}^{\ell-1} W_{t_i^n,t_{i+1}^n} \otimes \eta_{t_i^n,t_{i+1}^n}|^{\frac{p'}{2}}}{w_\eta(t_k^n,t_\ell^n) + c(t_k^n,t_\ell^n)} \lesssim 1. \tag{4.11}$$

By the Hölder continuity of η and W, it is clear that $|W_{t_k^n,t_\ell^n} \otimes \eta_{t_k^n,t_\ell^n}| \lesssim |t_\ell^n - t_k^n|^{\frac{2}{p'}}$, so that

$$\sup_{n\in\mathbb{N}} \sup_{0\leqslant k<\ell\leqslant 2^n} \frac{|W_{t_k^n,t_\ell^n} \otimes \eta_{t_k^n,t_\ell^n}|^{\frac{p'}{2}}}{|t_\ell^n - t_k^n|} \lesssim 1. \tag{4.12}$$

For any $n \in \mathbb{N}$ and $0 \leqslant k < \ell \leqslant 2^n$, it is straightforward to verify that

$$W_{t_k^n,t_\ell^n} \otimes \eta_{t_k^n,t_\ell^n} = \int_{t_k^n}^{t_\ell^n} W^n_{t_k^n,u} \otimes \mathrm{d}\eta_u + \left(\int_{t_k^n}^{t_\ell^n} \eta^n_{t_k^n,u} \otimes \mathrm{d}W_u\right)^\top + \sum_{i=k}^{\ell-1} W_{t_i^n,t_{i+1}^n} \otimes \eta_{t_i^n,t_{i+1}^n}.$$

Recalling (4.10), (4.11), and (4.12), we thus have that

$$\sup_{n\in\mathbb{N}} \sup_{0\leqslant k<\ell\leqslant 2^n} \frac{|\int_{t_k^n}^{t_\ell^n} W^n_{t_k^n,u} \otimes \mathrm{d}\eta_u|^{\frac{p'}{2}}}{\hat{w}(t_k^n,t_\ell^n)} \leqslant 1$$

for a suitable random control function $\hat{w}$. Combining this with (4.7), (4.8), and (4.10), we conclude that, for almost every $\omega \in \Omega$, the path $(\eta, W(\omega))$ indeed satisfies the third condition of Property (RIE). □

Remark 4.4. A joint rough path lift of (η, W) is constructed in [15, Section 2] that allows (4.2) to be treated as a rough Stratonovich SDE. Since the construction of the joint lift $\mathbf{\Lambda}$ above is based on a piecewise constant approximation, as in Property (RIE), rather than on linear interpolations as considered in [15], Theorem 4.3 provides a joint Itô-type rough path lift of (η, W), and thus, an Itô interpretation of the rough SDE (4.2), consistent with that in [18].

APPENDIX A: PROOF OF THEOREM 2.2

Proof of Theorem 2.2. *Step 1.* Let $L > 0$ such that $\|A\|_r, \|H\|_r, \|\mathbf{X}\|_p \leqslant L$, and let $w : \Delta_T \to [0, \infty)$ be the right-continuous control function given by

$$w(s,t) = \|A\|^r_{r,[s,t]} + \|H\|^r_{r,[s,t]} + \|X\|^p_{p,[s,t]} + \|\mathbb{X}\|^{\frac{p}{2}}_{\frac{p}{2},[s,t]}, \qquad \text{for} \quad (s,t) \in \Delta_T.$$

For $t \in (0, T]$, we define the map $\mathcal{M}_t : \mathcal{V}^{q,r}_X([0,t];\mathbb{R}^k) \to \mathcal{V}^{q,r}_X([0,t];\mathbb{R}^k)$ by

$$\mathcal{M}_t(Y, Y') = \left(y_0 + \int_0^\cdot b(H_s, Y_s)\,\mathrm{d}A_s + \int_0^\cdot \sigma(H_s, Y_s)\,\mathrm{d}\mathbf{X}_s, \sigma(H, Y)\right),$$

and, for $\delta \geqslant 1$, introduce the subset of controlled paths

$$\mathcal{B}^{(\delta)}_t = \left\{(Y, Y') \in \mathcal{V}^{q,r}_X([0,t];\mathbb{R}^k) : (Y_0, Y'_0) = (y_0, \sigma(H_0, y_0)),\ \|Y, Y'\|^{(\delta)}_{X,q,r} \leqslant 1\right\},$$

where

$$\|Y, Y'\|^{(\delta)}_{X,q,r} := \|Y'\|_{q,[0,t]} + \delta\|R^Y\|_{r,[0,t]}.$$

Applying standard estimates for Young and rough integrals (e.g., [22, Proposition 2.4 and Lemma 3.6]), for any $(Y, Y') \in \mathcal{B}^{(\delta)}_t$, we deduce that

$$\|\mathcal{M}_t(Y, Y')\|^{(\delta)}_{X,q,r} \leqslant C_1\left(\frac{1}{\delta} + \delta(\|A\|_{r,[0,t]} + \|H\|_{r,[0,t]} + \|\mathbf{X}\|_{p,[0,t]})\right),$$

for a constant $C_1 \geqslant \frac{1}{2}$ which depends only on $p, q, r, \|b\|_{C^2_b}, \|\sigma\|_{C^3_b}$, and L. Let $\delta = \delta_1 := 2C_1$, so that

$$\|\mathcal{M}_t(Y, Y')\|^{(\delta_1)}_{X,q,r} \leqslant \frac{1}{2} + 2C_1^2(2w(0,t)^{\frac{1}{r}} + w(0,t)^{\frac{1}{p}} + w(0,t)^{\frac{2}{p}}).$$

By the right continuity of w, we can then take $t = t_1$ sufficiently small such that

$$\|\mathcal{M}_{t_1}(Y, Y')\|^{(\delta_1)}_{X,q,r} \leqslant 1,$$

and we have that $\mathcal{B}^{(\delta_1)}_{t_1}$ is invariant under $\mathcal{M}_{t_1}$.

Step 2. Let $(Y,Y'),(\widetilde{Y},\widetilde{Y}')\in\mathcal{B}_t^{(\delta)}$, for some (new) $\delta\geqslant 1$ and $t\in(0,t_1]$. Applying standard estimates for Young and rough integrals (e.g., [22, Proposition 2.4, Lemma 3.1 and Lemma 3.7]), we deduce that

$$\begin{aligned}&\|\mathcal{M}_t(Y,Y')-\mathcal{M}_t(\widetilde{Y},\widetilde{Y}')\|_{X,q,r}^{(\delta)}\\&\quad\leqslant C_2\Big(\|R^Y-R^{\widetilde{Y}}\|_{r,[0,t]}+\delta(\|Y'-\widetilde{Y}'\|_{q,[0,t]}+\|R^Y-R^{\widetilde{Y}}\|_{r,[0,t]})(\|A\|_{r,[0,t]}+\|\mathbf{X}\|_{p,[0,t]})\Big),\end{aligned}$$

where $C_2>\frac{1}{2}$ depends only on $p,q,r,\|b\|_{C_b^2},\|\sigma\|_{C_b^3}$ and L. Let $\delta=\delta_2:=2C_2>1$, so that

$$\begin{aligned}&\|\mathcal{M}_t(Y,Y')-\mathcal{M}_t(\widetilde{Y},\widetilde{Y}')\|_{X,q,r}^{(\delta_2)}\\&\quad\leqslant\frac{\delta_2}{2}\|R^Y-R^{\widetilde{Y}}\|_{r,[0,t]}\\&\qquad+2C_2^2(\|Y'-\widetilde{Y}'\|_{q,[0,t]}+\|R^Y-R^{\widetilde{Y}}\|_{r,[0,t]})(w(0,t)^{\frac{1}{r}}+w(0,t)^{\frac{1}{p}}+w(0,t)^{\frac{2}{p}}).\end{aligned}$$

Again by the right continuity of w, we then take $t=t_2\leqslant t_1$ sufficiently small such that

$$\begin{aligned}\|\mathcal{M}_{t_2}(Y,Y')-\mathcal{M}_{t_2}(\widetilde{Y},\widetilde{Y}')\|_{X,q,r}^{(\delta_2)}&\leqslant\frac{1}{2}\|Y'-\widetilde{Y}'\|_{q,[0,t_2]}+\frac{\delta_2+1}{2}\|R^Y-R^{\widetilde{Y}}\|_{r,[0,t_2]}\\&\leqslant\frac{\delta_2+1}{2\delta_2}\|(Y,Y')-(\widetilde{Y},\widetilde{Y}')\|_{X,q,r}^{(\delta_2)},\end{aligned}$$

from which it follows that $\mathcal{M}_{t_2}$ is a contraction on the Banach space $(\mathcal{B}_{t_2}^{(\delta_1)},\|\cdot\|_{X,q,r}^{(\delta_2)})$. The fixed point of this map is the unique solution of the RDE (2.2) over the time interval $[0,t_2]$.

Step 3. Now let $\widetilde{A}\in D^{q_1},\widetilde{H}\in D^{q_2},\widetilde{\mathbf{X}}=(\widetilde{X},\widetilde{\mathbb{X}})\in\mathcal{D}^p$ and $\widetilde{y}_0\in\mathbb{R}^n$, such that $\|\widetilde{A}\|_r,\|\widetilde{H}\|_r,\|\widetilde{\mathbf{X}}\|_p\leqslant L$. By considering instead the control function w given by

$$\begin{aligned}w(s,t)&=\|A\|_{r,[s,t]}^r+\|H\|_{r,[s,t]}^r+\|X\|_{p,[s,t]}^p+\|\mathbb{X}\|_{\frac{p}{2},[s,t]}^{\frac{p}{2}}\\&\quad+\|\widetilde{A}\|_{r,[s,t]}^r+\|\widetilde{H}\|_{r,[s,t]}^r+\|\widetilde{X}\|_{p,[s,t]}^p+\|\widetilde{\mathbb{X}}\|_{\frac{p}{2},[s,t]}^{\frac{p}{2}},\qquad\text{for}\quad(s,t)\in\Delta_T,\end{aligned}$$

it follows from the above that there exist unique solutions $(Y,Y')\in\mathcal{V}_X^{q,r}([0,t_2];\mathbb{R}^k)$ and $(\widetilde{Y},\widetilde{Y}')\in\mathcal{V}_{\widetilde{X}}^{q,r}([0,t_2];\mathbb{R}^k)$ of the RDE (2.2), with data $(A,H,\mathbf{X},y_0)$ and $(\widetilde{A},\widetilde{H},\widetilde{\mathbf{X}},\widetilde{y}_0)$, respectively, over a sufficiently small time interval $[0,t_2]$. Standard estimates for Young and rough integrals (e.g., [22, Proposition 2.4, Lemma 3.1 and Lemma 3.7]) imply, after some calculation, that for any $\delta\geqslant 1$ and $t\in(0,t_2]$,

$$\begin{aligned}&\|Y'-\widetilde{Y}'\|_{q,[0,t]}+\delta\|R^Y-R^{\widetilde{Y}}\|_{r,[0,t]}\\&\quad\leqslant C_3\Big(|y_0-\widetilde{y}_0|+|H_0-\widetilde{H}_0|+\|H-\widetilde{H}\|_{r,[0,t]}+\|R^Y-R^{\widetilde{Y}}\|_{r,[0,t]}\\&\qquad+\delta(\|A-\widetilde{A}\|_{r,[0,t]}+\|\mathbf{X};\widetilde{\mathbf{X}}\|_{p,[0,t]})\end{aligned}$$

$$+\delta(|y_0-\widetilde{y}_0|+|H_0-\widetilde{H}_0|+\|H-\widetilde{H}\|_{r,[0,t]}+\|Y'-\widetilde{Y}'\|_{q,[0,t]}+\|R^Y-R^{\widetilde{Y}}\|_{r,[0,t]})$$
$$\times(\|A\|_{r,[0,t]}+\|\mathbf{X}\|_{p,[0,t]})),$$

where $C_3>0$ depends only on $p,q,r,\|b\|_{C_b^2},\|\sigma\|_{C_b^3}$ and L. Let $\delta=\delta_3:=C_3+1$, so that

$$\begin{aligned}&\|Y'-\widetilde{Y}'\|_{q,[0,t]}+\|R^Y-R^{\widetilde{Y}}\|_{r,[0,t]}\\&\leqslant C_3\big(|y_0-\widetilde{y}_0|+|H_0-\widetilde{H}_0|+\|H-\widetilde{H}\|_{r,[0,t]}+\delta_3(\|A-\widetilde{A}\|_{r,[0,t]}+\|\mathbf{X};\widetilde{\mathbf{X}}\|_{p,[0,t]})\\&\quad+\delta_3(|y_0-\widetilde{y}_0|+|H_0-\widetilde{H}_0|+\|H-\widetilde{H}\|_{r,[0,t]}+\|Y'-\widetilde{Y}'\|_{q,[0,t]}+\|R^Y-R^{\widetilde{Y}}\|_{r,[0,t]})\\&\quad\times(w(0,t)^{\frac{1}{r}}+w(0,t)^{\frac{1}{p}}+w(0,t)^{\frac{2}{p}})\Big).\end{aligned}$$

By taking $t=t_3\leqslant t_2$ sufficiently small, we deduce that

$$\begin{aligned}&\|Y-\widetilde{Y}\|_{p,[0,t_3]}+\|Y'-\widetilde{Y}'\|_{q,[0,t_3]}+\|R^Y-R^{\widetilde{Y}}\|_{r,[0,t_3]}\\&\leqslant C_4\Big(|y_0-\widetilde{y}_0|+|H_0-\widetilde{H}_0|+\|H-\widetilde{H}\|_{r,[0,t_3]}+\|A-\widetilde{A}\|_{r,[0,t_3]}+\|\mathbf{X};\widetilde{\mathbf{X}}\|_{p,[0,t_3]}\Big),\end{aligned}\tag{A.1}$$

for a new constant C_4, still depending only on $p,q,r,\|b\|_{C_b^2},\|\sigma\|_{C_b^3}$, and L.

Step 4. We infer from the above that there exists a constant $\varepsilon>0$, which depends only on $p,q,r,\|b\|_{C_b^2},\|\sigma\|_{C_b^3}$, and L, such that, given initial values $Y_s,\widetilde{Y}_s\in\mathbb{R}^k$, the local solutions (Y,Y') and $(\widetilde{Y},\widetilde{Y}')$ established above exist on any interval $[s,t]$ such that $w(s,t)\leqslant\varepsilon$. Moreover, these local solutions satisfy an estimate on this interval of the form in (A1).

By [22, Lemma 1.5], there exists a partition $\mathcal{P}=\{0=t_0<t_1<\cdots<t_N=T\}$, such that $w(t_i,t_{i+1}-)<\varepsilon$ for every $i=0,1,\ldots,N-1$. We can then define the solutions (Y,Y') and $(\widetilde{Y},\widetilde{Y}')$ on each of the half-open intervals $[t_i,t_{i+1})$. Given the solutions on $[t_i,t_{i+1})$, the values $Y_{t_{i+1}}$ and $\widetilde{Y}_{t_{i+1}}$ at the right end point of the interval are uniquely determined by the jumps of $A,\widetilde{A},\mathbf{X}$ and $\widetilde{\mathbf{X}}$ at time t_{i+1}. We thus deduce the existence of unique solutions (Y,Y') and $(\widetilde{Y},\widetilde{Y}')$ of the RDE on the entire interval $[0,T]$.

Since w is superadditive, we have that

$$w(t_0,t_1-)+w(t_1-,t_1)+w(t_1,t_2-)+\cdots+w(t_{N-1},t_N-)+w(t_N-,t_N)\leqslant w(0,T).$$

It is then straightforward to see that the partition $\mathcal{P}$ may be chosen such that the number of partition points in $\mathcal{P}$ may be bounded by a constant depending only on ε and $w(0,T)$. Thus, we may combine the local estimates in (A1) on each of the subintervals, together with simple estimates on the jumps at the end points of these subintervals, to obtain the global estimate in (2.3). □

APPENDIX B: THE CONVERGENCE OF PIECEWISE CONSTANT APPROXIMATIONS

In the following, we adopt the notation

$$\liminf_{n\to\infty}\mathcal{P}^n:=\bigcup_{m\in\mathbb{N}}\bigcap_{n\geq m}\mathcal{P}^n$$

for the times $t \in [0,T]$ which, as $n \to \infty$, eventually belong to all subsequent partitions in the sequence $(\mathcal{P}^n)_{n\in\mathbb{N}}$. The following proposition generalizes the result of [2, Proposition 2.14] so that the sequence of partitions is no longer assumed to be nested.

Proposition B.1. *Let $\mathcal{P}^n = \{0 = t_0^n < t_1^n < \cdots < t_{N_n}^n = T\}$, $n \in \mathbb{N}$, be a sequence of partitions with vanishing mesh size, so that $|\mathcal{P}^n| \to 0$ as $n \to \infty$. Let $F : [0,T] \to \mathbb{R}^d$ be a càdlàg path, and let*

$$F_t^n = F_T \mathbf{1}_{\{T\}}(t) + \sum_{k=0}^{N_n-1} F_{t_k^n} \mathbf{1}_{[t_k^n, t_{k+1}^n)}(t), \qquad t \in [0,T],$$

be the piecewise constant approximation of F along $\mathcal{P}^n$. Let

$$J_F := \{t \in (0,T] : F_{t-} \neq F_t\}$$

be the set of jump times of F. The following are equivalent:

(i) $J_F \subseteq \liminf_{n\to\infty} \mathcal{P}^n$,
(ii) *the sequence $(F^n)_{n\in\mathbb{N}}$ converges pointwise to F,*
(iii) *the sequence $(F^n)_{n\in\mathbb{N}}$ converges uniformly to F.*

Proof. We first show that conditions (i) and (ii) are equivalent. To this end, suppose that $J_F \subseteq \liminf_{n\to\infty} \mathcal{P}^n$ and let $t \in (0,T]$. If $t \in J_F$, then there exists $m \geqslant 1$ such that $t \in \mathcal{P}^n$ for all $n \geqslant m$. In this case, we then have that $F_t^n = F_t$ for all $n \geqslant m$. If $t \notin J_F$, then F is continuous at time t, and, since the mesh size $|\mathcal{P}^n| \to 0$, it follows that $F_t^n \to F_t$ as $n \to \infty$.

Now suppose instead that there exists a $t \in J_F$ such that $t \notin \liminf_{n\to\infty} \mathcal{P}^n$. Then there exists a subsequence $(n_j)_{j\in\mathbb{N}}$ such that $F_t^{n_j} \to F_{t-}$ as $j \to \infty$. Since $F_{t-} \neq F_t$, it follows that $F_t^n \nrightarrow F_t$. This establishes the equivalence of (i) and (ii).

Since (iii) clearly implies (ii), it only remains to show that (ii) implies (iii). By [16, Theorem 3.3], it is enough to show that the family of paths $\{F^n : n \in \mathbb{N}\}$ is equiregulated in the sense of [16, Definition 3.1]. That is, we need to show that, for every $t \in (0,T]$ and $\varepsilon > 0$, there exists a $u \in [0,t)$ such that $|F_s^n - F_{t-}^n| < \varepsilon$ for every $s \in (u,t)$ and every $n \in \mathbb{N}$, and moreover, that for every $t \in [0,T)$ and $\varepsilon > 0$, there exists a $u \in (t,T]$ such that $|F_s^n - F_t^n| < \varepsilon$ for every $s \in (t,u)$ and every $n \in \mathbb{N}$.

Step 1. Let $t \in (0,T]$ and $\varepsilon > 0$. Since the left limit F_{t-} exists, there exists $\delta > 0$ with $t - \delta > 0$, such that

$$|F_s - F_{t-}| < \frac{\varepsilon}{2} \qquad \text{for all} \quad s \in (t-\delta, t).$$

Since $|\mathcal{P}^n| \to 0$ as $n \to \infty$, there exists an $m \in \mathbb{N}$ such that, for every $n \geqslant m$, there exists a partition point $t_k^n \in \mathcal{P}^n$ such that $t - \delta < t_k^n < t - \frac{\delta}{2}$.

Let

$$u := \max\left(\left(t - \frac{\delta}{2}, t\right) \cap \bigcup_{n<m} \mathcal{P}^n\right),$$

where here we define $\max(\emptyset) := t - \frac{\delta}{2}$.

Take any $s \in (u,t)$ and any $n \in \mathbb{N}$. Let $i = \max\{k : t^n_k \leqslant s\}$ and $j = \max\{k : t^n_k < t\}$, so that $F^n_s = F_{t^n_i}$ and $F^n_{t-} = F_{t^n_j}$.

If $n \geqslant m$, then there exists a point $t^n_k \in \mathcal{P}^n$ such that $t - \delta < t^n_k < t - \frac{\delta}{2} \leqslant u < s$, and it follows that $t^n_i, t^n_j \in (t-\delta, t)$. If instead $n < m$, and if there exists a partition point $t^n_k \in (t - \frac{\delta}{2}, t)$, then $t - \frac{\delta}{2} < t^n_k \leqslant u < s$, and it again follows that $t^n_i, t^n_j \in (t-\delta, t)$. In either case, we then have that

$$|F^n_s - F^n_{t-}| = |F_{t^n_i} - F_{t^n_j}| \leqslant |F_{t^n_i} - F_{t-}| + |F_{t^n_j} - F_{t-}| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

The remaining case is when $n < m$ but $(t - \frac{\delta}{2}, t) \cap \mathcal{P}^n = \emptyset$. In this case, the path F^n is constant on the interval $[t - \frac{\delta}{2}, t)$ and, since $s \in (t - \frac{\delta}{2}, t)$, we have that $F^n_s = F^n_{t-}$.

In each case, we have that $|F^n_s - F^n_{t-}| < \varepsilon$ for all $s \in (u,t)$ and all $n \in \mathbb{N}$.

Step 2. Let $t \in (J_F \cup \{0\}) \setminus \{T\}$ and $\varepsilon > 0$. Since F is right-continuous, there exists a $\delta > 0$ with $t + \delta < T$, such that

$$|F_s - F_t| < \varepsilon \qquad \text{for all} \quad s \in [t, t+\delta).$$

Since condition (ii) implies condition (i), we know that $t \in \liminf_{n\to\infty} \mathcal{P}^n$, so that there exists an $m \in \mathbb{N}$ such that $t \in \cap_{n \geqslant m} \mathcal{P}^n$. Let

$$u := \min\left((t, t+\delta) \cap \bigcup_{n<m} \mathcal{P}^n \right),$$

where here we define $\min(\emptyset) := t + \delta$.

Take any $s \in (t,u)$, and any $n \in \mathbb{N}$. Let $i = \max\{k : t^n_k \leqslant s\}$, so that $F^n_s = F_{t^n_i}$.

If $n \geqslant m$, then $t \in \mathcal{P}^n$, so $F^n_t = F_t$ and, moreover, $t \leqslant t^n_i \leqslant s < u \leqslant t + \delta$, so that in particular $t^n_i \in [t, t+\delta)$, and hence

$$|F^n_s - F^n_t| = |F_{t^n_i} - F_t| < \varepsilon.$$

If $n < m$, then there does not exist any partition point $t^n_k \in (t,u) \cap \mathcal{P}^n$. It follows that the path F^n is constant on the interval $[t,u)$, so that, in particular, $F^n_s = F^n_t$.

In each case, we have that $|F^n_s - F^n_t| < \varepsilon$ for all $s \in (t,u)$ and all $n \in \mathbb{N}$.

Step 3. Let $t \in (0,T) \setminus J_F$ and $\varepsilon > 0$. Since F is continuous at time t, there exists a $\delta > 0$ with $0 < t - \delta$ and $t + \delta < T$, such that

$$|F_s - F_t| < \frac{\varepsilon}{2} \qquad \text{for all} \quad s \in (t-\delta, t+\delta).$$

Since $|\mathcal{P}^n| \to 0$ as $n \to \infty$, there exists an $m \in \mathbb{N}$ such that, for every $n \geqslant m$, there exists a partition point $t^n_k \in \mathcal{P}^n$ such that $t - \delta < t^n_k < t$. Let

$$u := \min\left((t, t+\delta) \cap \bigcup_{n<m} \mathcal{P}^n \right),$$

where here we define $\min(\emptyset) := t + \delta$.

Take any $s \in (t,u)$ and any $n \in \mathbb{N}$. Let $i = \max\{k : t_k^n \leqslant s\}$ and $j = \max\{k : t_k^n \leqslant t\}$, so that $F_s^n = F_{t_i^n}$ and $F_t^n = F_{t_j^n}$.

If $n \geqslant m$, then there exists a point $t_k^n \in \mathcal{P}^n$ such that $t_k^n \in (t-\delta, t)$, and it follows that $t_i^n, t_j^n \in (t-\delta, t+\delta)$, so that

$$|F_s^n - F_t^n| = |F_{t_i^n} - F_{t_j^n}| \leqslant |F_{t_i^n} - F_t| + |F_{t_j^n} - F_t| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

If $n < m$, then there does not exist any partition point $t_k^n \in (t,u) \cap \mathcal{P}^n$. It follows that the path F^n is constant on the interval $[t,u)$, so that, in particular, $F_s^n = F_t^n$.

In each case, we have that $|F_s^n - F_t^n| < \varepsilon$ for all $s \in (t,u)$ and all $n \in \mathbb{N}$. It follows that the family of paths $\{F^n : n \in \mathbb{N}\}$ is indeed equiregulated. □

Theorem B.2. *Let $p \in (2,3)$, $q \in [p,\infty)$ and $r \in [\frac{p}{2}, 2)$ such that $\frac{1}{p} + \frac{1}{r} > 1$ and $\frac{1}{p} + \frac{1}{q} = \frac{1}{r}$, and let $\mathcal{P}^n = \{0 = t_0^n < t_1^n < \cdots < t_{N_n}^n = T\}$, $n \in \mathbb{N}$, be a sequence of partitions with vanishing mesh size. Suppose that X satisfies Property (RIE) relative to p and $(\mathcal{P}^n)_{n\in\mathbb{N}}$, and let $\mathbf{X}$ be the canonical rough path lift of X, as constructed in (2.7). Let $(F,F') \in \mathcal{V}_X^{q,r}$ be a controlled path with respect to X, and suppose that $J_F \subseteq \liminf_{n\to\infty} \mathcal{P}^n$, where J_F is the set of jump times of F. Then the rough integral of (F,F') against $\mathbf{X}$ is given by*

$$\int_0^t F_u \,\mathrm{d}\mathbf{X}_u = \lim_{n\to\infty} \sum_{k=0}^{N_n-1} F_{t_k^n} X_{t_k^n \wedge t, t_{k+1}^n \wedge t},$$

where the convergence is uniform in $t \in [0,T]$.

The previous theorem generalizes the result of [2, Theorem 2.15] so that the sequence of partitions is no longer assumed to be nested. The proof of Theorem B.2 follows the proof of [2, Theorem 2.15] almost verbatim. The only difference is that, rather than using [2, Proposition 2.14] to establish the uniform convergence of F^n to F, we can instead use Proposition B.1 (which does not require the sequence of partitions to be nested).

ACKNOWLEDGMENTS

A. P. Kwossek and D. J. Prömel gratefully acknowledge financial support by the Baden-Württemberg Stiftung, and would like to thank A. Neuenkirch for fruitful discussions that helped to improve the present work. A. P. Kwossek was affiliated with the University of Mannheim for the majority of this project's duration.

JOURNAL INFORMATION

The *Journal of the London Mathematical Society* is wholly owned and managed by the London Mathematical Society, a not-for-profit Charity registered with the UK Charity Commission. All surplus income from its publishing programme is used to support mathematicians and mathematics research in the form of research grants, conference grants, prizes, initiatives for early career researchers and the promotion of mathematics.

ORCID

Andrew L. Allan https://orcid.org/0000-0003-3324-6237

REFERENCES

1. A. L. Allan, *Robust filtering and propagation of uncertainty in hidden Markov models*, Electron. J. Probab. **26** (2021), no. 73, 1–37.
2. A. L. Allan, C. Liu, and D. J. Prömel, *A càdlàg rough path foundation for robust finance*, Finance Stoch. **28** (2024), 215–257.
3. D. Applebaum, *Lévy processes and stochastic calculus*, 2nd ed., Cambridge Studies in Advanced Mathematics, vol. 116, Cambridge University Press, Cambridge, 2009.
4. I. Bailleul, *Flows driven by rough paths*, Rev. Mat. Iberoam. **31** (2015), no. 3, 901–934.
5. D. Bartl, M. Kupper, and A. Neufeld, *Stochastic integration and differential equations for typical paths*, Electron. J. Probab. **24** (2019), no. 97, 1–21.
6. D. Bartl, M. Kupper, D. J. Prömel, and L. Tangpi, *Duality for pathwise superhedging in continuous time*, Finance Stoch. **23** (2019), no. 3, 697–728.
7. K. Bichteler, *Stochastic integration and L^p-theory of semimartingales*, Ann. Probab. **9** (1981), no. 1, 49–89.
8. J. Bretagnolle, *p-variation de fonctions aléatoires. 2ième partie: processus à accroissements indépendants*, Séminaire de probabilités de Strasbourg **6** (1972), 64–71.
9. R. Cont and P. Das, *Quadratic variation and quadratic roughness*, Bernoulli **29** (2023), no. 1, 496–522.
10. R. Cont and P. Tankov, *Financial modelling with jump processes*, Chapman & Hall/CRC Financial Mathematics Series, Chapman & Hall/CRC, Boca Raton, FL, 2004.
11. D. Crisan, J. Diehl, P. K. Friz, and H. Oberhauser, *Robust filtering: correlated noise and multidimensional observation*, Ann. Appl. Probab. **23** (2013), no. 5, 2139–2160.
12. A. M. Davie, *Differential equations driven by rough paths: an approach via discrete approximation*, Appl. Math. Res. Express. (2008), no. 2, Art. ID abm009, 1–40.
13. S. Dereich and F. Heidenreich, *A multilevel Monte Carlo algorithm for Lévy-driven stochastic differential equations*, Stochastic Process. Appl. **121** (2011), no. 7, 1565–1587.
14. A. Deya, A. Neuenkirch, and S. Tindel, *A Milstein-type scheme without Lévy area terms for SDEs driven by fractional Brownian motion*, Ann. Inst. Henri Poincaré Probab. Stat. **48** (2012), no. 2, 518–550.
15. J. Diehl, H. Oberhauser, and S. Riedel, *A Lévy area between Brownian motion and rough paths with applications to robust nonlinear filtering and rough partial differential equations*, Stochastic Process. Appl. **125** (2015), 161–181.
16. D. Fraňková, *Regulated functions with values in Banach space*, Math. Bohem. **144** (2019), no. 4, 437–456.
17. P. K. Friz and M. Hairer, *A course on rough paths with an introduction to regularity structures*, 2nd ed., Springer, Cham, 2020.
18. P. K. Friz, A. Hocquet, and K. Lê, *Rough stochastic differential equations*, Preprint, 2021 arXiv:2106.10340.
19. P. K. Friz and A. Shekhar, *General rough integration, Lévy rough paths and a Lévy-Kintchine-type formula*, Ann. Probab. **45** (2017), no. 4, 2707–2765.
20. P. Friz and N. Victoir, *Euler estimates for rough differential equations*, J. Differential Equations **244** (2008), no. 2, 388–412.
21. P. K. Friz and N. B. Victoir, *Multidimensional stochastic processes as rough paths*, Cambridge Studies in Advanced Mathematics, vol. 120, Cambridge University Press, Cambridge, 2010, Theory and applications.
22. P. K. Friz and H. Zhang, *Differential equations driven by rough paths with jumps*, J. Differential Equations **264** (2018), no. 10, 6226–6301.
23. I. Gyöngy, *A note on Euler's approximations*, Potential Anal. **8** (1998), no. 3, 205–216.
24. J. Hong, C. Huang, and X. Wang, *Symplectic Runge-Kutta methods for Hamiltonian systems driven by Gaussian rough paths*, Appl. Numer. Math. **129** (2018), 120–136.
25. J. Jacod, T. G. Kurtz, S. Méléard, and P. Protter, *The approximate Euler method for Lévy driven stochastic differential equations*, Ann. Inst. H. Poincaré Probab. Statist. **41** (2005), no. 3, 523–558.
26. J. Jacod and A. N. Shiryaev, *Limit theorems for stochastic processes*, 2nd ed., Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 288, Springer, Berlin, 2003.
27. R. L. Karandikar, *On almost sure convergence of modified Euler-Peano approximation of solution to an S.D.E. driven by a semimartingale*, Séminaire de Probabilités, XXV, Lecture Notes in Math., vol. 1485, Springer, Berlin, 1991, pp. 113–120.

28. P. E. Kloeden and A. Neuenkirch, *The pathwise convergence of approximation schemes for stochastic differential equations*, LMS J. Comput. Math. **10** (2007), 235–253.
29. P. E. Kloeden and E. Platen, *Numerical solution of stochastic differential equations*, Applications of Mathematics (New York), vol. 23, Springer, Berlin, 1992.
30. K. Kubilius, *The existence and uniqueness of the solution of an integral equation driven by a p-semimartingale of special type*, Stochastic Process. Appl. **98** (2002), no. 2, 289–315.
31. C. Liu and D. J. Prömel, *Examples of Itô càdlàg rough paths*, Proc. Amer. Math. Soc. **146** (2018), no. 11, 4937–4950.
32. Y. Liu and S. Tindel, *First-order Euler scheme for SDEs driven by fractional Brownian motions: the rough case*, Ann. of Appl. Probab. **29** (2019), no. 2, 758–826.
33. Y. S. Mishura and G. M. Shevchenko, *Existence and uniqueness of the solution of stochastic differential equation involving Wiener process and fractional Brownian motion with Hurst index* $H > 1/2$, Comm. Statist. Theory Methods **40** (2011), no. 19–20, 3492–3508.
34. R. Norvaiša, *The p-variation and an extension of the class of semimartingales*, Acta Appl. Math. **79** (2003), no. 1–2, 47–53.
35. N. Perkowski and D. J. Prömel, *Pathwise stochastic integrals for model free finance*, Bernoulli **22** (2016), no. 4, 2486–2520.
36. P. E. Protter, *Stochastic integration and differential equations*, Stochastic modelling and applied probability, vol. 21, Springer, Berlin, 2005, 2nd ed., Version 2.1, Corrected third printing.
37. M. Redmann and S. Riedel, *Runge-Kutta methods for rough differential equations*, J. Stoch. Anal. **3** (2022), no. 4, Art. 6, 24.
38. S. Rubenthaler, *Numerical simulation of the solution of a stochastic differential equation driven by a Lévy process*, Stochastic Process. Appl. **103** (2003), no. 2, 311–349.
39. T. Shardlow and P. Taylor, *On the pathwise approximation of stochastic differential equations*, BIT **56** (2016), no. 3, 1101–1129.
40. R. Vershynin, *High-dimensional probability*, Cambridge Series in Statistical and Probabilistic Mathematics, vol. 47, Cambridge University Press, Cambridge, 2018, An introduction with applications in data science, With a foreword by Sara van de Geer.
41. M. Zähle, *Integration with respect to fractal functions and stochastic calculus. II*, Math. Nachr. **225** (2001), 145–183.