



On the support of solutions to stochastic differential equations with path-dependent coefficients

Rama Cont^{a,*}, Alexander Kalinin^{b,1}

^a *Mathematical Institute, University of Oxford, England*

^b *Department of Mathematics, Imperial College London, England*

Received 8 October 2018; received in revised form 16 July 2019; accepted 30 July 2019

Available online 6 August 2019

Abstract

Given a stochastic differential equation with path-dependent coefficients driven by a multidimensional Wiener process, we show that the support of the law of the solution is given by the image of the Cameron–Martin space under the flow of mild solutions to a system of path-dependent ordinary differential equations. Our result extends the Stroock–Varadhan support theorem for diffusion processes to the case of SDEs with path-dependent coefficients. The proof is based on functional Itô calculus.
© 2019 Elsevier B.V. All rights reserved.

MSC: 60H10; 28C20; 34K50

Keywords: Support theorem; Stochastic differential equation; Functional equation; Semimartingale; Wiener space; Functional Itô calculus

1. Overview

1.1. Support theorems for stochastic differential equations

A stochastic process may be viewed as a random variable taking values in a space of paths. The (topological) support of this random variable then describes the closure of the set of all attainable paths and provides insight into the structure of sample paths of the process. The nature of the support has been investigated for various classes of stochastic processes under different function space topologies, with a focus on stochastic differential equations.

* Corresponding author.

E-mail addresses: Rama.Cont@maths.ox.ac.uk (R. Cont), alex.kalinin@mail.de (A. Kalinin).

¹ Alexander Kalinin's research was supported by a Chapman Fellowship from the Department of Mathematics, Imperial College London.

For diffusion processes the support with respect to the supremum norm was first described by Stroock and Varadhan [19,20], a result known as the ‘Stroock–Varadhan support theorem’. An extension to unbounded coefficients was given by Gyöngy [13]. The support of more general Wiener functionals and extensions to SDEs in Hilbert spaces are discussed in Aida et al. [1,2]. These results were extended to the Hölder topology by Ben Arous et al. [4] and, using different techniques, by Millet and Sanz-Solé [16]. Bally et al. [3] use similar methods to derive a support theorem in Hölder norm for parabolic SPDEs. Support theorems in p-variation topology are discussed by Ledoux et al. [14], using rough path techniques. Support theorems in Hölder and p-variation topologies are discussed in [12] and Pakkanen [17] gives conditions for a stochastic integral to have full support. In this work we extend some of these results to stochastic differential equations with path-dependent coefficients.

Let $r, T \geq 0$ with $r < T$ and $d, m \in \mathbb{N}$ and suppose that $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in [0, T]}, P)$ is a filtered probability space satisfying the usual conditions and on which there is a standard d -dimensional $(\mathcal{F}_t)_{t \in [0, T]}$ -Brownian motion W . We consider the following path-dependent stochastic differential equation:

$$dX_t = b(t, X) dt + \sigma(t, X) dW_t \quad \text{for } t \in [r, T] \tag{1.1}$$

with initial condition $X_s = \hat{x}(s)$ for $s \in [0, r]$, where $\hat{x} \in C([0, T], \mathbb{R}^m)$ and the coefficients $b : [r, T] \times C([0, T], \mathbb{R}^m) \rightarrow \mathbb{R}^m$ and $\sigma : [r, T] \times C([0, T], \mathbb{R}^m) \rightarrow \mathbb{R}^{m \times d}$ are product measurable and non-anticipative in the sense that $b(t, x)$ and $\sigma(t, x)$ depend on the path $x \in C([0, T], \mathbb{R}^m)$ up to time $t \in [r, T]$ only.

Under Lipschitz continuity and affine growth conditions on b and σ , this SDE admits a unique strong solution for which a.e. sample path lies in the delayed Hölder space $C_r^\alpha([0, T], \mathbb{R}^m)$ for every $\alpha \in [0, 1/2)$. Our main result is a description of the support of the solution in the Hölder topology: we show that the support of its law is given by the image of the Cameron–Martin space under the flow associated with a system of functional differential equations.

1.2. Statement of the main result

Denote by $|\cdot|$ the Euclidean norm in \mathbb{R}^d and the Hilbert–Schmidt norms in $\mathbb{R}^{d \times d}$, $\mathbb{R}^{m \times d}$ and $\mathbb{R}^{m \times m}$. We let \mathbb{I}_d be the $d \times d$ identity matrix and for a matrix $A \in \mathbb{R}^{m \times d}$ we denote by A' its transpose. For a path $x : [0, T] \rightarrow \mathbb{R}^m$ we let x^t be the path stopped at $t \in [0, T]$: $x^t(s) = x(s \wedge t)$ for $s \in [0, T]$. Throughout, $C([0, T], \mathbb{R}^m)$ is the separable Banach space of all \mathbb{R}^m -valued continuous maps on $[0, T]$, endowed with the supremum norm given by $\|x\|_\infty = \sup_{t \in [0, T]} |x(t)|$.

For $\alpha \in (0, 1]$ we introduce the non-separable Banach space $C_r^\alpha([0, T], \mathbb{R}^m)$ of all $x \in C([0, T], \mathbb{R}^m)$ that are α -Hölder continuous on $[r, T]$, equipped with the ‘delayed α -Hölder norm’ defined via

$$\|x\|_{\alpha,r} := \|x^r\|_\infty + \sup_{s,t \in [r,T]: s \neq t} \frac{|x(s) - x(t)|}{|s - t|^\alpha}. \tag{1.2}$$

We set $C_r^0([0, T], \mathbb{R}^m) := C([0, T], \mathbb{R}^m)$ and $\|\cdot\|_{0,r} := \|\cdot\|_\infty$, by convention, and let $H_r^1([0, T], \mathbb{R}^m)$ be the separable Banach space of all $x \in C([0, T], \mathbb{R}^m)$ that are absolutely continuous on $[r, T]$ and whose weak derivative \dot{x} is square-integrable with respect to the Lebesgue measure, endowed with the ‘delayed Cameron–Martin norm’ given by

$$\|x\|_{H,r} := \|x^r\|_\infty + \left(\int_r^T |\dot{x}(s)|^2 ds \right)^{1/2}. \tag{1.3}$$

Then it holds that $H_r^1([0, T], \mathbb{R}^m) \subsetneq C_r^{1/2}([0, T], \mathbb{R}^m)$ and $\|x\|_{1/2,r} \leq \|x\|_{H,r}$ for all $x \in H_r^1([0, T], \mathbb{R}^m)$. In the case of no delay the Cameron–Martin space is a Hilbert space. Further, we allow infinite values and extend the definitions of $\|\cdot\|_\infty$ and $\|\cdot\|_{\alpha,r}$ at (1.2) to any map $x : [0, T] \rightarrow \mathbb{R}^m$ and the definition of $\|\cdot\|_{H,r}$ at (1.3) to any map $x : [0, T] \rightarrow \mathbb{R}^m$ that is absolutely continuous on $[r, T]$.

We assume the diffusion coefficient σ to be of class $\mathcal{C}^{1,2}$ on $[r, T] \times C([0, T], \mathbb{R}^m)$ in the sense of horizontal and vertical differentiability [7,11], as discussed in Section 2.1, and consider the map $\rho : [r, T] \times C([0, T], \mathbb{R}^m) \rightarrow \mathbb{R}^m$ with components

$$\rho_k(t, x) := \sum_{l=1}^d \partial_x \sigma_{k,l}(t, x) \sigma(t, x) e_l, \tag{1.4}$$

if $t < T$, and $\rho_k(t, x) := 0$, otherwise. Here, $\{e_1, \dots, e_d\}$ denotes the canonical basis of \mathbb{R}^d and $\partial_x \sigma_{k,l} : [r, T] \times C([0, T], \mathbb{R}^m) \rightarrow \mathbb{R}^{1 \times m}$ is the vertical derivative of the (k, l) -coordinate of σ for every $k \in \{1, \dots, m\}$ and $l \in \{1, \dots, d\}$. Note that $\rho = \partial_x \sigma \sigma$ for $m = d = 1$. Moreover, the horizontal and the second-order vertical derivative of σ are denoted by $\partial_t \sigma$ and $\partial_{xx} \sigma$, respectively.

We characterize the support of the unique strong solution to (1.1) in terms of the following path-dependent ordinary differential equation driven by an element $h \in H_r^1([0, T], \mathbb{R}^d)$:

$$\dot{x}_h(t) = (b - (1/2)\rho)(t, x_h) + \sigma(t, x_h) \dot{h}(t) \quad \text{for } t \in [r, T]. \tag{1.5}$$

Based on these preliminaries, our main result may be stated as follows:

Theorem 1 (Support Theorem for Path-dependent SDEs). *Let σ be of class $\mathcal{C}^{1,2}$ on $[r, T] \times C([0, T], \mathbb{R}^m)$. Assume that σ and $\partial_x \sigma$ are bounded and there are $c, \eta, \lambda \geq 0$ and $\kappa \in [0, 1)$ such that*

$$\begin{aligned} |b(t, x)| &\leq c(1 + \|x\|_\infty^\kappa), & |b(t, x) - b(t, y)| &\leq \lambda \|x - y\|_\infty, \\ |\partial_t \sigma_{k,l}(t, x)| + |\partial_{xx} \sigma_{k,l}(t, x)| &\leq c(1 + \|x\|_\infty^\eta) \quad \text{and} \\ |\sigma(t, x) - \sigma(s, y)| + |\partial_x \sigma_{k,l}(t, x) - \partial_x \sigma_{k,l}(s, y)| &\leq \lambda(|s - t|^{1/2} + \|x^t - y^s\|_\infty) \end{aligned}$$

for all $s, t \in [r, T]$, $x, y \in C([0, T], \mathbb{R}^m)$, $k \in \{1, \dots, m\}$ and $l \in \{1, \dots, d\}$. Then the following three assertions are valid:

- (i) Pathwise uniqueness holds for (1.1) and there is a unique strong solution X to (1.1) satisfying $X_s = \hat{x}(s)$ for all $s \in [0, r]$ a.s. Further, $E[\|X\|_{\alpha,r}^p] < \infty$ for each $\alpha \in [0, 1/2)$ and $p \geq 1$.
- (ii) For $h \in H_r^1([0, T], \mathbb{R}^d)$ there is a unique mild solution x_h to (1.5) such that $x_h(s) = \hat{x}(s)$ for all $s \in [0, r]$ and we have $x_h \in H_r^1([0, T], \mathbb{R}^m)$. In addition, the map $H_r^1([0, T], \mathbb{R}^d) \rightarrow H_r^1([0, T], \mathbb{R}^m)$, $h \mapsto x_h$ is Lipschitz continuous on bounded sets.
- (iii) For $\alpha \in [0, 1/2)$ the support of the image measure $P \circ X^{-1}$ in the delayed Hölder space $C_r^\alpha([0, T], \mathbb{R}^m)$ is the closure of the set of all mild solutions x_h to (1.5), where $h \in H_r^1([0, T], \mathbb{R}^d)$. That is,

$$\text{supp}(P \circ X^{-1}) = \overline{\{x_h \mid h \in H_r^1([0, T], \mathbb{R}^d)\}} \quad \text{in } C_r^\alpha([0, T], \mathbb{R}^m). \tag{1.6}$$

This theorem extends previous results [2,4,16,19] on the support of diffusion processes to the case of path-dependent coefficients. Moreover, in the diffusion case we retrieve the results of [4,16] under weaker regularity conditions.

Our proof uses the functional Itô calculus [5,7,8,11] to generalize the approach used by Millet and Sanz-Solé [16] to the path-dependent case, providing the correct Girsanov changes of measures to deduce (2.7). Based on adapted linear interpolations of Brownian motion, we construct Hölder continuous approximations of solutions to (1.1) and (1.5) and show that these approximations converge in Hölder norm in probability to the respective solutions. A key ingredient is the functional Itô formula in [7,9,11], combined with interpolation error estimates in supremum norm for stochastic processes.

Outline. The remainder of this paper is devoted to the proof of Theorem 1. Section 2 discusses the various building blocks of the proof. Section 2.1 recalls several functional calculus concepts from [5,7,8,11] that are useful in our setting. Section 2.2 gives conditions for the existence and uniqueness of a mild solution to the path-dependent ODE (1.5); Section 2.3 gives conditions for the existence of a unique strong solution to the path-dependent SDE (1.1). Section 2.4 discusses the interpolation method used to characterize the support in the Hölder topology.

Section 3.1 deals with Hölder spaces for stochastic processes and the notion of convergence in Hölder norm in probability in more depth. Section 3.2 derives a quantitative version of the Kolmogorov–Chentsov Theorem with an explicit estimate of the Hölder norm (Proposition 12). Section 3.3 provides a sufficient criterion, based on a sequence of partitions, for convergence in Hölder norm in probability (Lemma 13). While Section 3.4 discusses adapted linear interpolations of Brownian motion, Section 3.5 deduces interpolation error estimates for stochastic processes and several required moment estimates, improving in particular a convergence result from [16][Lemma 3.2].

Section 4 proves the existence and uniqueness of mild solutions to path-dependent ODEs (Section 4.1) and strong solutions to path-dependent SDEs (Section 4.2). Finally, Section 5 combines these ingredients to give a proof of the main result.

2. Preliminaries

2.1. Non-anticipative functional calculus

Let $D([0, T], \mathbb{R}^m)$ denote the Banach space of all \mathbb{R}^m -valued càdlàg maps on $[0, T]$, equipped with the supremum norm $\|\cdot\|_\infty$, and recall the following notions from [5,8]. A functional $G : [r, T] \times D([0, T], \mathbb{R}^m) \rightarrow \mathbb{R}$ is *non-anticipative* if

$$\forall t \in [r, T], \forall x \in D([0, T], \mathbb{R}^m), \quad G(t, x) = G(t, x^t),$$

where $x^t = x(t \wedge \cdot)$ is the path x stopped at t . G is called *boundedness-preserving* if it is bounded on bounded sets: for each $n \in \mathbb{N}$ there is $c_n \geq 0$ such that

$$\forall t \in [r, T], \forall x \in D([0, T], \mathbb{R}^m), \quad \|x\|_\infty \leq n \quad \Rightarrow \quad |G(t, x)| \leq c_n.$$

We notice that the following pseudometric on $[r, T] \times D([0, T], \mathbb{R}^m)$ given by

$$d_\infty((t, x), (s, y)) := |t - s|^{1/2} + \|x^t - y^s\|_\infty$$

is complete and if G is d_∞ -continuous, then it is non-anticipative. As observed in [10], Lipschitz continuity with respect to d_∞ allows for a Hölder smoothness of degree 1/2 in the time variable.

Let us recall the definitions of the horizontal and vertical derivative from [7,11]. A non-anticipative functional G on $[r, T] \times D([0, T], \mathbb{R}^m)$ is *horizontally differentiable* if for each $t \in [r, T]$ and $x \in D([0, T], \mathbb{R}^m)$, the function

$$[0, T - t] \rightarrow \mathbb{R}, \quad h \mapsto G(t + h, x^t)$$

is differentiable at 0. In this case, its derivative there is denoted by $\partial_t G(t, x)$. We say that G is *vertically differentiable* if for all $t \in [r, T)$ and $x \in D([0, T], \mathbb{R}^m)$, the function

$$\mathbb{R}^m \rightarrow \mathbb{R}, \quad h \mapsto G(t, x + h \mathbb{1}_{[t, T)})$$

is differentiable at 0. If this is the case, then we denote its derivative there by $\partial_x G(t, x)$. G is *partially vertically differentiable* if for every $k \in \{1, \dots, m\}$, $t \in [r, T)$ and $x \in D([0, T], \mathbb{R}^m)$, the function

$$\mathbb{R} \rightarrow \mathbb{R}, \quad h \mapsto G(t, x + h \bar{e}_k \mathbb{1}_{[t, T)})$$

is differentiable at 0, where $\{\bar{e}_1, \dots, \bar{e}_m\}$ is the canonical basis of \mathbb{R}^m . In this case, its derivative there is denoted by $\partial_{x_k} G(t, x)$. If G is vertically differentiable, then it is partially vertically differentiable and $\partial_x G = (\partial_{x_1} G, \dots, \partial_{x_m} G)$.

We call G *twice vertically differentiable* if it is vertically differentiable and the same holds for $\partial_x G$. In this case, we set $\partial_{xx} G := \partial_x(\partial_x G)$ and

$$\partial_{x_k x_l} G := \partial_{x_k}(\partial_{x_l} G) \quad \text{for all } k, l \in \{1, \dots, m\}.$$

It follows from Schwarz’s Lemma that if G is twice vertically differentiable and $\partial_{xx} G$ is d_∞ -continuous, then $\partial_{xx} G$ is symmetric: $\partial_{x_k x_l} G = \partial_{x_l x_k} G$ for all $k, l \in \{1, \dots, m\}$. Finally, G is of class $\mathbb{C}^{1,2}$ if it is once horizontally and twice vertically differentiable such that G and its derivatives $\partial_t G$, $\partial_x G$ and $\partial_{xx} G$ are boundedness-preserving and d_∞ -continuous.

While horizontal differentiability can be readily extended to functionals that are merely defined on $[r, T) \times C([0, T], \mathbb{R}^m)$, vertical differentiability is based on càdlàg perturbations, as it involves indicator functions. A functional F on $[r, T) \times C([0, T], \mathbb{R}^m)$ is said to be of class $\mathbb{C}^{1,2}$ if it admits an extension $G : [r, T) \times D([0, T], \mathbb{R}^m) \rightarrow \mathbb{R}$ that satisfies this property. Then it follows from Theorems 5.4.1 and 5.4.2 in [5] that the definitions

$$\partial_x F := \partial_x G \quad \text{and} \quad \partial_{xx} F := \partial_{xx} G \quad \text{on } [r, T) \times C([0, T], \mathbb{R}^m)$$

are independent of the choice of the extension G . Based on this uniqueness result, we can apply the functional Itô formula [9] to characterize the support in the proof of Proposition 31.

Examples 2. (i) Let $\alpha \in (0, 1]$ and $\varphi : \mathbb{R}^m \rightarrow \mathbb{R}_+$ be α -Hölder continuous. Then the *running supremum functional* $G : [r, T) \times D([0, T], \mathbb{R}^m) \rightarrow \mathbb{R}$ given by

$$G(t, x) := \sup_{s \in [0, t]} \varphi(x(s))$$

is boundedness-preserving and α -Hölder continuous with respect to d_∞ . To be more precise, there exists $\lambda \geq 0$ such that

$$|G(t, x) - G(s, y)| \leq \lambda \|x^t - y^s\|_\infty^\alpha$$

for all $s, t \in [r, T)$ and $x, y \in D([0, T], \mathbb{R}^m)$. Further, we have $\partial_t G = 0$. However, if φ fails to be differentiable at some $\bar{x} \in \mathbb{R}^m$ such that $\varphi(\bar{x} + h) > \varphi(\bar{x})$ for all $h \in \mathbb{R}^m \setminus \{0\}$, then G is not vertically differentiable. For instance, take $\varphi = |\cdot|^\alpha$.

(ii) Let $\alpha \in (0, 1]$, $\varphi : \mathbb{R}^m \rightarrow \mathbb{R}^d$ be α -Hölder continuous and $\beta : [r, T) \rightarrow [0, T]$ be right-continuous and satisfy $\beta(t) < t$ for all $t \in (r, T)$. Then the *non-anticipative delayed map* $G : [r, T) \times D([0, T], \mathbb{R}^m) \rightarrow \mathbb{R}^d$ given by

$$G(t, x) := \varphi((x \circ \beta)(t))$$

is boundedness-preserving and α -Hölder continuous in $x \in D([0, T], \mathbb{R}^m)$, uniformly in $t \in [r, T]$. Unless φ is constant, G fails to be horizontally differentiable, yet it is vertically differentiable on $(r, T) \times D([0, T], \mathbb{R}^m)$ and $\partial_x G = 0$ there.

(iii) Assume that $\varphi : [0, T] \times \mathbb{R}^m \rightarrow \mathbb{R}^{m \times d}$ is a Borel measurable bounded map that is Lipschitz continuous in $\bar{x} \in \mathbb{R}^m$, uniformly in $t \in [0, T]$. Then the *non-anticipative integral map* $G : [r, T] \times D([0, T], \mathbb{R}^m) \rightarrow \mathbb{R}^{m \times d}$ given by

$$G(t, x) := \int_0^t \varphi(s, x(s)) ds$$

is bounded and d_∞ -Lipschitz continuous. Moreover, if φ is continuous, then G is of class $\mathbb{C}^{1,2}$. In fact, it holds that $\partial_t G(t, x) = \varphi(t, x(t))$ for any $t \in [r, T)$ and $x \in D([0, T], \mathbb{R}^m)$ and $\partial_x G = 0$.

2.2. Mild solutions to path-dependent ODEs

In this section we show that the ODE (1.5) admits a unique mild solution which belongs to the delayed Cameron–Martin space $H_r^1([0, T], \mathbb{R}^m)$. To this end, let us consider the path-dependent ordinary differential equation:

$$\dot{x}(t) = F(t, x) \quad \text{for } t \in [r, T], \tag{2.1}$$

where $F : [r, T] \times C([0, T], \mathbb{R}^m) \rightarrow \mathbb{R}^m$ denotes a non-anticipative product measurable map. Then for each $h \in H_r^1([0, T], \mathbb{R}^d)$ the choice $F = b - (1/2)\rho + \sigma \dot{h}$ yields (1.5) with the correction term ρ given by (1.4).

In general the map $[r, T] \rightarrow \mathbb{R}^m, t \mapsto F(t, x)$ may fail to be continuous, so one may not expect solutions in the classical sense. A *mild solution* to (2.1) is a path $x \in C([0, T], \mathbb{R}^m)$ satisfying

$$\int_r^T |F(s, x)| ds < \infty \quad \text{and} \quad x(t) = x(r) + \int_r^t F(s, x) ds$$

for all $t \in [r, T]$. By definition, a mild solution x is absolutely continuous on $[r, T]$ and it becomes a classical solution if and only if the measurable map $[r, T] \rightarrow \mathbb{R}^m, t \mapsto F(t, x)$ is continuous.

Let us introduce the following regularity conditions, which are satisfied under the assumptions of Theorem 1 for the choice of F mentioned before.

(C.1) There is a measurable function $c_0 : [r, T] \rightarrow \mathbb{R}_+$ satisfying $\int_r^T c_0(s)^2 ds < \infty$ and

$$|F(t, x)| \leq c_0(t) \left(1 + \|x^r\|_\infty + \int_r^T |\dot{x}(s)| ds \right)$$

for all $t \in [r, T]$ and $x \in C([0, T], \mathbb{R}^m)$ that is absolutely continuous on $[r, T]$.

(C.2) For each $n \in \mathbb{N}$ there is a measurable function $\lambda_n : [r, T] \rightarrow \mathbb{R}_+$ such that $\int_r^T \lambda_n(s)^2 ds < \infty$ and

$$|F(t, x) - F(t, y)| \leq \lambda_n(t) \|x - y\|_{H,r}$$

for every $t \in [r, T]$ and $x, y \in H_r^1([0, T], \mathbb{R}^m)$ with $\|x\|_{H,r} \vee \|y\|_{H,r} \leq n$.

Under the above affine growth condition and Lipschitz smoothness on bounded sets, we obtain a unique mild solution that can be approximated by a Picard iteration in the complete norm $\|\cdot\|_{H,r}$ defined in (1.3).

Proposition 3. Let (C.1) and (C.2) hold. Then (2.1) admits a unique mild solution x satisfying

$$x(s) = \hat{x}(s) \text{ for all } s \in [0, r].$$

Moreover $x \in H_r^1([0, T], \mathbb{R}^m)$ and the sequence $(x_n)_{n \in \mathbb{N}_0}$ in $H_r^1([0, T], \mathbb{R}^m)$, recursively defined by

$$x_0(t) := \hat{x}(r \wedge t), \quad x_{n+1}(t) := x_0(t) + \int_r^{r \vee t} F(s, x_n) ds$$

converges to x in the delayed Cameron–Martin norm $\|\cdot\|_{H,r}$.

2.3. Strong solutions to path-dependent SDEs

We turn to the derivation of a unique strong solution to the SDE (1.1) whose paths lie in $C_r^\alpha([0, T], \mathbb{R}^m)$ for every $\alpha \in [0, 1/2)$. By a solution to (1.1) we mean an $(\mathcal{F}_t)_{t \in [0, T]}$ -adapted continuous process $X : [0, T] \times \Omega \rightarrow \mathbb{R}^m$ satisfying

$$\int_r^T |b(s, X)| ds + \int_r^T |\sigma(s, X)|^2 ds < \infty \text{ a.s.}$$

$$\text{and } X_t = X_r + \int_r^t b(s, X) ds + \int_r^t \sigma(s, X) dW_s$$

for every $t \in [r, T]$ a.s. A solution is said to be strong if it is adapted to the augmented natural filtration of the underlying Brownian motion W .

We now state the required conditions on the coefficients b and σ , which are valid in the setting of Theorem 1.

(C.3) There are a measurable function $c_0 : [r, T] \rightarrow \mathbb{R}_+$ and $\tilde{c}_0 \geq 0$ such that $\int_r^T c_0(s)^2 ds < \infty$,

$$|b(t, x)| \leq c_0(t)(1 + \|x\|_\infty) \text{ and } |\sigma(t, x)| \leq \tilde{c}_0(1 + \|x\|_\infty)$$

for all $t \in [r, T]$ and $x \in C([0, T], \mathbb{R}^m)$.

(C.4) There are $\alpha_0 \in [0, 1/2)$, a measurable function $\lambda_0 : [r, T] \rightarrow \mathbb{R}_+$ and $\tilde{\lambda}_0 \geq 0$ such that $\int_r^T \lambda_0(s)^2 ds < \infty$,

$$|b(t, x) - b(t, y)| \leq \lambda_0(t)\|x - y\|_{\alpha_0, r} \text{ and } |\sigma(t, x) - \sigma(t, y)| \leq \tilde{\lambda}_0\|x - y\|_{\alpha_0, r}$$

for each $t \in [r, T]$ and $x, y \in C_r^{\alpha_0}([0, T], \mathbb{R}^m)$.

Remark 4. If (C.4) holds, then it is also true if α_0 is replaced by any $\alpha \in (\alpha_0, 1]$. Thus, it is strongest in the case that $\alpha_0 = 0$, since $\|\cdot\|_0 = \|\cdot\|_\infty$, by convention.

Let $\alpha \in [0, 1]$ and $p \geq 1$. By Proposition 11, the space $\mathcal{C}_{r,p}^\alpha([0, T], \mathbb{R}^m)$ of all $(\mathcal{F}_t)_{t \in [0, T]}$ -adapted continuous processes $X : [0, T] \times \Omega \rightarrow \mathbb{R}^m$ for which $E[\|X\|_{\alpha,r}^p]$ is finite, equipped with the seminorm

$$\mathcal{C}_{r,p}^\alpha([0, T], \mathbb{R}^m) \rightarrow \mathbb{R}_+, \quad X \mapsto (E[\|X\|_{\alpha,r}^p])^{1/p} \tag{2.2}$$

is complete. Note that if a sequence $({}_n X)_{n \in \mathbb{N}}$ in this seminormed space converges, then it also converges in the norm $\|\cdot\|_{\alpha,r}$ in probability. We define

$$\mathcal{C}_\infty([0, T], \mathbb{R}^m) := \bigcap_{p \geq 1} \mathcal{C}_{r,p}^0([0, T], \mathbb{R}^m) \text{ and}$$

$$\mathcal{C}_{r,\infty}^{1/2-}([0, T], \mathbb{R}^m) := \bigcap_{\alpha \in [0, 1/2), p \geq 1} \mathcal{C}_{r,p}^\alpha([0, T], \mathbb{R}^m),$$

which are completely pseudometrizable topological spaces.

Proposition 5. *Let (C.3) and (C.4) hold. Then, up to indistinguishability, there is a unique strong solution $X \in \mathcal{C}_{r,\infty}^{1/2-}([0, T], \mathbb{R}^m)$ to (1.1) such that*

$$X_s = \hat{x}(s) \text{ for all } s \in [0, r] \text{ a.s.}$$

and the sequence $({}_nX)_{n \in \mathbb{N}_0}$ in $\mathcal{C}_{r,\infty}^{1/2-}([0, T], \mathbb{R}^m)$, recursively given by

$${}_0X_t := \hat{x}(r \wedge t), \quad {}_{n+1}X_t = {}_nX_t + \int_r^{r \vee t} b(s, {}_nX) ds + \int_r^{r \vee t} \sigma(s, {}_nX) dW_s$$

converges to X in the norm $\|\cdot\|_{\alpha,r}$ in p -th moment for $p \geq 1$:

$$\forall \alpha \in [0, 1/2), \quad \forall p \geq 1, \quad \lim_{n \uparrow \infty} E[\|{}_nX - X\|_{\alpha,r}^p] = 0.$$

Remark 6. Pathwise uniqueness for (1.1) is shown in Lemma 27, requiring only the following Lipschitz condition on bounded sets, which follows from (C.4) in the strongest case $\alpha_0 = 0$:

(C.5) For each $n \in \mathbb{N}$ there is a measurable function $\lambda_n : [r, T] \rightarrow \mathbb{R}_+$ satisfying $\int_r^T \lambda_n(s)^2 ds < \infty$ and

$$|b(t, x) - b(t, y)| + |\sigma(t, x) - \sigma(t, y)| \leq \lambda_n(t) \|x - y\|_\infty$$

for every $t \in [r, T]$ and $x, y \in C([0, T], \mathbb{R}^m)$ with $\|x\|_\infty \vee \|y\|_\infty \leq n$.

2.4. Characterization of the support in Hölder topology

Sections 2.2 and 2.3 provide the main arguments to prove the first two assertions of Theorem 1. Let us now describe how we will prove the characterization (1.6) of the support. For $n \in \mathbb{N}$ let \mathbb{T}_n be a partition of $[r, T]$ that we write in the form

$$\mathbb{T}_n = \{t_{0,n}, \dots, t_{k_n,n}\}$$

for some $k_n \in \mathbb{N}$ and $t_{0,n}, \dots, t_{k_n,n} \in [r, T]$ such that $r = t_{0,n} < \dots < t_{k_n,n} = T$ and we denote its mesh by $|\mathbb{T}_n| = \max_{i \in \{0, \dots, k_n-1\}} (t_{i+1,n} - t_{i,n})$. We assume that $\lim_{n \uparrow \infty} |\mathbb{T}_n| = 0$ and that the sequence $(\mathbb{T}_n)_{n \in \mathbb{N}}$ of partitions is *balanced* in the sense of [6]. That is, there is $c_{\mathbb{T}} \geq 1$ such that

$$|\mathbb{T}_n| \leq c_{\mathbb{T}} \min_{i \in \{0, \dots, k_n-1\}} (t_{i+1,n} - t_{i,n}) \text{ for all } n \in \mathbb{N}. \tag{2.3}$$

For $k, n \in \mathbb{N}$ and every path $x : [0, T] \rightarrow \mathbb{R}^k$ we let $L_n(x) : [0, T] \rightarrow \mathbb{R}^k$ denote the unique map satisfying $L_n(x)(t) = x(r \wedge t)$ for $t \in [0, t_{1,n}]$ and

$$L_n(x)(t) = x(t_{i-1,n}) + \frac{t - t_{i,n}}{t_{i+1,n} - t_{i,n}} (x(t_{i,n}) - x(t_{i-1,n})) \tag{2.4}$$

for $i \in \{1, \dots, k_n - 1\}$ and $t \in [t_{i,n}, t_{i+1,n}]$. Then $L_n(x)$ is piecewise continuously differentiable on $[r, T]$ and can be regarded as *delayed linear interpolation* of x along \mathbb{T}_n on this interval.

Hence, we may define an adapted process ${}_nW : [0, T] \times \Omega \rightarrow \mathbb{R}^d$, whose paths belong to $H_r^1([0, T], \mathbb{R}^d)$, by setting

$${}_nW_t := L_n(W)(t). \tag{2.5}$$

Let us for the moment suppose that the assumptions and the first two claims of **Theorem 1** hold. To establish that $\text{supp}(P \circ X^{-1})$ is included in the closure of $\{x_h \mid h \in H_r^1([0, T], \mathbb{R}^d)\}$ in $C_r^\alpha([0, T], \mathbb{R}^m)$ for $\alpha \in [0, 1/2)$, we will justify in Section 5.4 that it suffices to check that

$$\lim_{n \uparrow \infty} P(\|x_{{}_nW} - X\|_{\alpha,r} \geq \varepsilon) = 0 \quad \text{for all } \varepsilon > 0. \tag{2.6}$$

We remark that, by the definition of mild solution to (2.1), for each $n \in \mathbb{N}$ the adapted process $x_{{}_nW}$, whose paths lie in $H_r^1([0, T], \mathbb{R}^m)$, is a strong solution to the degenerate SDE

$$d_n Y_t = ((b - (1/2)\rho)(t, {}_nY) + \sigma(t, {}_nY)_n \dot{W}_t) dt \quad \text{for } t \in [r, T]$$

with initial condition ${}_n Y^r = \hat{x}^r$. Next, we will show that the converse inclusion in (1.6) follows if for every $h \in H_r^1([0, T], \mathbb{R}^d)$ we can find a sequence $(P_{h,n})_{n \in \mathbb{N}}$ of probability measures on (Ω, \mathcal{F}) that are absolutely continuous to P such that

$$\lim_{n \uparrow \infty} P_{h,n}(\|X - x_h\|_{\alpha,r} \geq \varepsilon) = 0 \quad \text{for any } \varepsilon > 0. \tag{2.7}$$

We emphasize that for each $n \in \mathbb{N}$ the measure $P_{h,n}$ will be constructed by means of Girsanov’s theorem such that X is a strong solution to the SDE

$$d_n Y_t = (b(t, {}_nY) + \sigma(t, {}_nY)(\dot{h} - \dot{L}_n({}_nW)))(t) dt + \sigma(t, {}_nY) d_{h,n} W_t \tag{2.8}$$

for $t \in [r, T]$ under $P_{h,n}$ with initial condition ${}_n Y^r = \hat{x}^r$, where $h, {}_nW$ is some standard d -dimensional $(\mathcal{F}_t)_{t \in [0, T]}$ -Brownian motion under $P_{h,n}$. Hence, to prove (2.6) and (2.7) at the same time, we consider the following general framework.

Assume \underline{B} is an \mathbb{R}^m -valued and B_H, \bar{B} and Σ are $\mathbb{R}^{m \times d}$ -valued non-anticipative product measurable maps on $[r, T] \times C([0, T], \mathbb{R}^m)$. Then for each $n \in \mathbb{N}$ we introduce the SDE

$$d_n Y_t = (\underline{B}(t, {}_nY) + B_H(t, {}_nY)\dot{h}(t) + \bar{B}(t, {}_nY)_n \dot{W}_t) dt + \Sigma(t, {}_nY) dW_t \tag{2.9}$$

for $t \in [r, T]$. Under the hypothesis that \bar{B} is of class $\mathbb{C}^{1,2}$ on $[r, T] \times C([0, T], \mathbb{R}^m)$, we also consider the SDE

$$dY_t = ((\underline{B} + R)(t, Y) + B_H(t, Y)\dot{h}(t) + (\bar{B} + \Sigma)(t, Y) dW_t) \tag{2.10}$$

for $t \in [r, T]$, where the \mathbb{R}^m -valued non-anticipative product measurable map R on $[r, T] \times C([0, T], \mathbb{R}^m)$ is given coordinatewise by

$$R_k(t, x) := \sum_{l=1}^d \partial_x \bar{B}_{k,l}(t, x)((1/2)\bar{B} + \Sigma)(t, x)e_l, \tag{2.11}$$

if $t < T$, and $R_k(t, x) := 0$, otherwise. In **Theorem 7** we show that whenever ${}_nY$ and Y are respectively solutions to (2.9) and (2.10) such that ${}_n Y^r = Y^r = \hat{x}^r$ a.s. for all $n \in \mathbb{N}$, then

$$\lim_{n \uparrow \infty} P(\|{}_nY - Y\|_{\alpha,r} \geq \varepsilon) = 0 \quad \text{for each } \varepsilon > 0. \tag{2.12}$$

Then the choice $\underline{B} = b - (1/2)\rho$, $B_H = 0$, $\bar{B} = \sigma$ and $\Sigma = 0$ yields (2.6), since $R = (1/2)\rho$ in this case. Moreover, by choosing $\underline{B} = b$, $B_H = \sigma$, $\bar{B} = -\sigma$ and $\Sigma = \sigma$ instead, (2.7) follows. Since these are the two desired results, let us consider the following required conditions:

(C.6) \bar{B} is of class $\mathbb{C}^{1,2}$ on $[r, T] \times C([0, T], \mathbb{R}^m)$ and there are $c, \eta \geq 0$ and $\kappa \in [0, 1)$ such that $|\underline{B}(t, x)| + |B_H(t, x)| \leq c(1 + \|x\|_\infty^\kappa)$,

$$|\bar{B}(t, x)| + \left(\sum_{k=1}^m \sum_{l=1}^d |\partial_x \bar{B}_{k,l}(t, x)|^2 \right)^{1/2} + |\Sigma(t, x)| \leq c \quad \text{and}$$

$$|\partial_t \bar{B}(t, x)| + \left(\sum_{k=1}^m \sum_{l=1}^d |\partial_{xx} \bar{B}_{k,l}(t, x)|^2 \right)^{1/2} \leq c(1 + \|x\|_\infty^\eta)$$

for all $t \in [r, T]$ and $x \in C([0, T], \mathbb{R}^m)$.

(C.7) \underline{B} is Lipschitz continuous in $x \in C([0, T], \mathbb{R}^m)$, uniformly in $t \in [r, T]$, and $B_H, \bar{B}, \partial_x \bar{B}$ and Σ are d_∞ -Lipschitz continuous.

(C.8) There are $\bar{b}_0 \in \mathbb{R}$ and a measurable function $\bar{b} : [r, T] \rightarrow \mathbb{R}$ such that $\int_r^T |\bar{b}(s)|^2 ds < \infty$ and $\bar{b}_0 \bar{B}(t, x) = \bar{b}(t) \Sigma(t, x)$ for all $t \in [r, T]$ and $x \in C([0, T], \mathbb{R}^m)$.

Theorem 7. Let (C.6)–(C.8) be satisfied and $h \in H_r^1([0, T], \mathbb{R}^d)$ and $(x_n)_{n \in \mathbb{N}}$ be a bounded sequence in $C([0, T], \mathbb{R}^m)$. Then the following assertions hold:

- (i) For any $n \in \mathbb{N}$ there is a unique strong solution ${}_n Y$ to (2.9) such that ${}_n Y^r = x_n^r$ a.s. Moreover, $\sup_{n \in \mathbb{N}} E[\|{}_n Y\|_{\alpha,r}^p] < \infty$ for all $\alpha \in [0, 1/2)$ and $p \geq 1$.
- (ii) There is a unique strong solution Y to (2.10) such that $Y^r = \hat{x}^r$ a.s. and we have $E[\|Y\|_{\alpha,r}^p] < \infty$ for every $\alpha \in [0, 1/2)$ and $p \geq 1$.
- (iii) Let $\alpha \in [0, 1/2)$ satisfy $\lim_{n \rightarrow \infty} |\mathbb{T}_n|^{-2\alpha} \|x_n^r - \hat{x}^r\|_\infty^2 = 0$. Then

$$\lim_{n \rightarrow \infty} |\mathbb{T}_n|^{-2\alpha} E \left[\max_{j \in \{0, \dots, k_n\}} |{}_n Y_{t_{j,n}} - Y_{t_{j,n}}|^2 \right] = 0. \tag{2.13}$$

In particular, (2.12) holds: $({}_n Y)_{n \in \mathbb{N}}$ converges in the norm $\|\cdot\|_{\alpha,r}$ in probability to Y .

Remark 8. Condition (C.8) allows us to perform a change of measure and obtain a unique strong solution to (2.9). However, when deriving (2.13) in Sections 5.1–5.3 we merely assume (C.6) and (C.7).

3. Convergence in Hölder norm in probability

3.1. Hölder spaces for stochastic processes

For $\alpha \in [0, 1]$ we let $\mathcal{C}_r^\alpha([0, T], \mathbb{R}^m)$ denote the linear space of all adapted continuous processes $X : [0, T] \times \Omega \rightarrow \mathbb{R}^m$ satisfying $X \in C_r^\alpha([0, T], \mathbb{R}^m)$ a.s., endowed with the pseudometric

$$\mathcal{C}_r^\alpha([0, T], \mathbb{R}^m) \times \mathcal{C}_r^\alpha([0, T], \mathbb{R}^m) \rightarrow \mathbb{R}_+, \tag{3.1}$$

$$(X, Y) \mapsto E[\|X - Y\|_{\alpha,r} \wedge 1].$$

Then a sequence $({}_n X)_{n \in \mathbb{N}}$ in this space converges to some $X \in \mathcal{C}_r^\alpha([0, T], \mathbb{R}^m)$ if and only if it converges in the norm $\|\cdot\|_{\alpha,r}$ in probability to this process. Put differently, $(\|{}_n X - X\|_{\alpha,r})_{n \in \mathbb{N}}$

converges in probability to zero. Further, $({}_nX)_{n \in \mathbb{N}}$ is Cauchy if and only if

$$\lim_{k \uparrow \infty} \sup_{n \in \mathbb{N}: n \geq k} P(\|{}_nX - kX\|_{\alpha,r} \geq \varepsilon) = 0 \quad \text{for all } \varepsilon > 0.$$

To abbreviate notation, we set $\mathcal{C}([0, T], \mathbb{R}^m) := \mathcal{C}_r^0([0, T], \mathbb{R}^m)$, which represents the linear space of all \mathbb{R}^m -valued adapted continuous processes. The fact that $\mathcal{C}([0, T], \mathbb{R}^m)$ is complete can be extended, as the following result shows.

Lemma 9. *The linear space $\mathcal{C}_r^\alpha([0, T], \mathbb{R}^m)$ endowed with the pseudometric (3.1) is complete.*

Proof. Let $({}_nX)_{n \in \mathbb{N}}$ be a Cauchy sequence in $\mathcal{C}_r^\alpha([0, T], \mathbb{R}^m)$. By Lemma 4.3.3 in [20] for instance, there is $X \in \mathcal{C}([0, T], \mathbb{R}^m)$ to which $({}_nX)_{n \in \mathbb{N}}$ converges uniformly in probability. For given $\varepsilon, \eta > 0$ there is $n_0 \in \mathbb{N}$ such that

$$P\left(\sup_{s,t \in [r,T]: s \neq t} \frac{|({}_nX_s - kX_s) - ({}_nX_t - kX_t)|}{|s - t|^\alpha} \geq \frac{\varepsilon}{2}\right) < \frac{\eta}{2}$$

for all $k, n \in \mathbb{N}$ with $k \wedge n \geq n_0$. We fix $l \in \mathbb{N}$ and set $\delta_l := (T - r)/l$, then there exists $k_l \in \mathbb{N}$ such that $k_l \geq n_0$ and $P(\|k_l X - X\|_\infty \geq (\varepsilon/4)\delta_l^\alpha) < \eta/2$. Hence,

$$P\left(\sup_{s,t \in [r,T]: |s-t| \geq \delta_l} \frac{|({}_nX_s - X_s) - ({}_nX_t - X_t)|}{|s - t|^\alpha} > \varepsilon\right) < \eta$$

for each $n \in \mathbb{N}$ with $n \geq n_0$. By the continuity of measures, $(\|{}_nX - X\|_{\alpha,r})_{n \in \mathbb{N}}$ converges in probability to zero. In particular, $\|X\|_{\alpha,r} < \infty$ a.s. \square

For $p \geq 1$ we recall the linear space $\mathcal{C}_{r,p}^\alpha([0, T], \mathbb{R}^m)$ of all $X \in \mathcal{C}_r^\alpha([0, T], \mathbb{R}^m)$ for which $E[\|X\|_{\alpha,r}^p]$ is finite, endowed with the seminorm (2.2). We say that a sequence $({}_nX)_{n \in \mathbb{N}}$ in this space is *p-fold uniformly integrable* if $(\|{}_nX\|_{\alpha,r})_{n \in \mathbb{N}}$ satisfies this property in the usual sense.

Lemma 10. *Any Cauchy sequence $({}_nX)_{n \in \mathbb{N}}$ in $\mathcal{C}_{r,p}^\alpha([0, T], \mathbb{R}^m)$ with respect to the seminorm (2.2) is p-fold uniformly integrable.*

Proof. Let $\varepsilon > 0$, then there is $n_0 \in \mathbb{N}$ such that $E[\|{}_nX - kX\|_{\alpha,r}^p] < \varepsilon/2^p$ for all $k, n \in \mathbb{N}$ with $k \wedge n \geq n_0$. As the random variable $Y := \max_{n \in \{1, \dots, n_0\}} \|{}_nX\|_{\alpha,r}$ is p-fold integrable, we obtain that

$$\sup_{n \in \mathbb{N}} (E[\mathbb{1}_A \|{}_nX\|_{\alpha,r}^p])^{1/p} \leq (E[\mathbb{1}_A Y^p])^{1/p} + \varepsilon^{1/p}/2$$

for each $A \in \mathcal{F}$. First, by choosing $A = \Omega$, this gives $\sup_{n \in \mathbb{N}} E[\|{}_nX\|_{\alpha,r}^p] < \infty$. Secondly, by setting $\delta := \varepsilon/2^p$, it follows that $\sup_{n \in \mathbb{N}} E[\mathbb{1}_A \|{}_nX\|_{\alpha,r}^p] < \varepsilon$ for every $A \in \mathcal{F}$ with $E[\mathbb{1}_A Y^p] < \delta$. \square

We conclude with the following convergence characterization.

Proposition 11. *A sequence $({}_nX)_{n \in \mathbb{N}}$ in $\mathcal{C}_{r,p}^\alpha([0, T], \mathbb{R}^m)$ converges with respect to the seminorm (2.2) if and only if it is p-fold uniformly integrable and there is $X \in \mathcal{C}_r^\alpha([0, T], \mathbb{R}^m)$ such that*

$$\lim_{n \uparrow \infty} P(\|{}_nX - X\|_{\alpha,r} \geq \varepsilon) = 0 \quad \text{for all } \varepsilon > 0.$$

In the latter case, $E[\|X\|_{\alpha,r}^p] < \infty$ and $\lim_{n \uparrow \infty} E[\|{}_nX - X\|_{\alpha,r}^p] = 0$. Moreover, $\mathcal{C}_{r,p}^\alpha([0, T], \mathbb{R}^m)$ equipped with (2.2) is complete.

Proof. By Lemmas 9 and 10, it suffices to show the if-direction of the first claim. We let $(\nu_n)_{n \in \mathbb{N}}$ be a strictly increasing sequence in \mathbb{N} such that $(\|\nu_n X - X\|_{\alpha,r})_{n \in \mathbb{N}}$ converges almost-surely to zero, then

$$E[\|X\|_{\alpha,r}^p] \leq \liminf_{n \uparrow \infty} E[\|\nu_n X\|_{\alpha,r}^p] \leq \sup_{n \in \mathbb{N}} E[\|nX\|_{\alpha,r}^p] < \infty,$$

by Fatou’s Lemma. For $\varepsilon > 0$ there are $\delta > 0$ and $n_0 \in \mathbb{N}$ such that $E[\mathbb{1}_A \|nX\|_{\alpha,r}^p] < (\varepsilon/3)^p$ and $P(\|nX - X\|_{\alpha,r} \geq \varepsilon/3) < \delta$ for every $A \in \mathcal{F}$ and $n \in \mathbb{N}$ with $P(A) < \delta$ and $n \geq n_0$. Thus,

$$(E[\|nX - X\|_{\alpha,r}^p])^{1/p} \leq (E[\mathbb{1}_{\{\|nX - X\|_{\alpha,r} \geq \varepsilon/3\}} \|nX - X\|_{\alpha,r}^p])^{1/p} + \varepsilon/3 < \varepsilon$$

for any such $n \in \mathbb{N}$, because similar reasoning as before gives us that $E[\mathbb{1}_A \|X\|_{\alpha,r}^p] \leq \sup_{n \in \mathbb{N}} E[\mathbb{1}_A \|nX\|_{\alpha,r}^p] < \infty$ for all $A \in \mathcal{F}$. This completes the proof. \square

3.2. A general kolmogorov–chentsov estimate

We revisit the proof of the Kolmogorov–Chentsov Theorem (see e.g. Theorem 2.1 in [18]) to obtain a quantitative estimate of the Hölder norm. Let

$$k_{\alpha,p,q} := 2^{p+q} (2^{q/p-\alpha} - 1)^{-p} \tag{3.2}$$

for $p \geq 1, q > 0$ and $\alpha \in [0, q/p)$ and note that the function $[0, q/p) \rightarrow (2^p, \infty), \alpha \mapsto k_{\alpha,p,q}$ is strictly increasing with $\lim_{\alpha \uparrow q/p} k_{\alpha,p,q} = \infty$. Then the following result holds, in which the process in question and not necessarily a modification appears.

Proposition 12. Assume that X is an \mathbb{R}^m -valued right-continuous process for which there are $c_0 \geq 0, p \geq 1$ and $q > 0$ such that

$$E[|X_s - X_t|^p] \leq c_0 |s - t|^{1+q} \tag{3.3}$$

for all $s, t \in [r, T]$. Then for each $\alpha \in [0, q/p)$ it holds that

$$E \left[\sup_{s,t \in [r,T]: s \neq t} \frac{|X_s - X_t|^p}{|s - t|^{\alpha p}} \right] \leq k_{\alpha,p,q} c_0 (T - r)^{1+q-\alpha p}.$$

In particular if $q \leq p$, then the sample paths of X are a.s. α -Hölder continuous on $[r, T]$ for all $\alpha \in [0, q/p)$.

Proof. For given $n \in \mathbb{N}_0$ let \mathbb{D}_n be the n th dyadic partition of $[r, T]$ whose points are $d_{i,n} := r + i2^{-n}(T - r)$, where $i \in \{0, \dots, 2^n\}$. We define

$$\Delta_n := \{(s, t) \in \mathbb{D}_n \times \mathbb{D}_n \mid |s - t| \leq 2^{-n}(T - r)\},$$

then it is readily seen that there are 2^n tuples $(s, t) \in \Delta_n$ satisfying $s < t$. For $Y_n := \sup_{(s,t) \in \Delta_n} |X_s - X_t|$ condition (3.3) gives

$$E[Y_n^p] \leq \sum_{(s,t) \in \Delta_n: s < t} E[|X_s - X_t|^p] \leq 2^{-nq} c_0 (T - r)^{1+q}. \tag{3.4}$$

We set $\mathbb{D} := \bigcup_{n \in \mathbb{N}_0} \mathbb{D}_n$ and let $s, t \in \mathbb{D}$ satisfy $0 < t - s < 2^{-n}(T - r)$ for some $n \in \mathbb{N}_0$. Then for each $k \in \mathbb{N}_0$ there are unique $i_k, j_k \in \{1, \dots, 2^k\}$ such that $d_{i_k-1,k} \leq s < d_{i_k,k}$, and

$$d_{j_k-1,k} \leq t < d_{j_k,k}, \quad \text{if } t < T, \quad \text{and } d_{j_k,k} = T, \quad \text{otherwise.}$$

As $(d_{i_k,k})_{k \in \mathbb{N}_0}$ and $(d_{j_k,k})_{k \in \mathbb{N}_0}$ are two decreasing sequences converging to s and t , respectively, two telescoping sums yield that

$$X_s - X_t = X_{d_{i_n,n}} - X_{d_{j_n,n}} + \sum_{k=n}^{\infty} (X_{d_{i_{k+1},k+1}} - X_{d_{i_k,k}}) + \sum_{k=n}^{\infty} (X_{d_{j_{k+1},k+1}} - X_{d_{j_k,k}}).$$

We notice that either $i_n = j_n$ or instead $n \geq 1$, $j_n \geq 2$ and $i_n = j_n - 1$, since $0 < t - s < 2^{-n}(T - r)$. In both cases, we have $(d_{i_n,n}, d_{j_n,n}) \in \Delta_n$. Moreover, $(d_{i_k,k}, d_{i_{k+1},k+1}), (d_{j_k,k}, d_{j_{k+1},k+1}) \in \Delta_{k+1}$ for all $k \in \mathbb{N}_0$, by construction. So,

$$|X_s - X_t| \leq 2 \sum_{k=n}^{\infty} Y_k.$$

Clearly, for each $s, t \in \mathbb{D}$ with $0 < t - s < T - r$ there is a unique $n \in \mathbb{N}_0$ satisfying $2^{-n-1}(T - r) \leq t - s < 2^{-n}(T - r)$. This entails that

$$\sup_{s,t \in [r,T]: s \neq t} \frac{|X_s - X_t|}{|s - t|^\alpha} \leq 2^{1+\alpha} (T - r)^{-\alpha} \sum_{k=0}^{\infty} 2^{\alpha k} Y_k, \tag{3.5}$$

as \mathbb{D} is a countable dense set in $[r, T]$ containing T and X is right-continuous. Hence, (3.5), the triangle inequality, monotone convergence and (3.4) yield that

$$\left(E \left[\sup_{s,t \in [r,T]: s \neq t} \frac{|X_s - X_t|^p}{|s - t|^{\alpha p}} \right] \right)^{1/p} \leq 2^{1+\alpha} c_0^{1/p} (T - r)^{(1+q)/p - \alpha} \sum_{k=0}^{\infty} 2^{(\alpha - q/p)k}.$$

Since the power series on the right-hand side converges absolutely to the inverse of $1 - 2^{\alpha - q/p}$, the proposition follows. \square

3.3. Convergence along a sequence of partitions

We state a sufficient criterion for a sequence of processes to converge in the norm $\|\cdot\|_{\alpha,r}$ in probability, where $\alpha \in [0, 1]$. For this purpose, condition (2.3) on the sequence of partitions is crucial.

Lemma 13. *Let $({}_n X)_{n \in \mathbb{N}}$ be a sequence of \mathbb{R}^m -valued right-continuous processes for which there are $p, q > 0$ with $q \leq p$ such that for each $\beta \in (0, q/p)$ there is $c_\beta \geq 0$ satisfying*

$$P \left(\max_{j \in \{0, \dots, k_n - 1\}} \sup_{s,t \in [t_{j,n}, t_{j+1,n}]: s \neq t} \frac{|{}_n X_s - {}_n X_t|}{|s - t|^\beta} \geq \lambda \right) \leq c_\beta \lambda^{-p} \tag{3.6}$$

for every $n \in \mathbb{N}$ and $\lambda > 0$. If $(\|{}_n X^r\|_\infty)_{n \in \mathbb{N}}$ and $(\max_{j \in \{1, \dots, k_n\}} |{}_n X_{t_{j,n}}| / |\mathbb{T}_n|^\alpha)_{n \in \mathbb{N}}$ converge in probability to zero, then so does $(\|{}_n X\|_{\alpha,r})_{n \in \mathbb{N}}$ for any $\alpha \in [0, q/p)$.

Proof. Let $\beta \in (\alpha, q/p)$ and $n \in \mathbb{N}$. First, a case distinction shows that

$$\begin{aligned} \sup_{s,t \in [r,T]: s \neq t} \frac{|{}_n X_s - {}_n X_t|}{|s - t|^\alpha} &\leq 2 \max_{j \in \{0, \dots, k_n - 1\}} \sup_{s,t \in [t_{j,n}, t_{j+1,n}]: s \neq t} \frac{|{}_n X_s - {}_n X_t|}{|s - t|^\alpha} \\ &\quad + \max_{i,j \in \{1, \dots, k_n\}: i \neq j} \frac{|{}_n X_{t_{i,n}} - {}_n X_{t_{j,n}}|}{|t_{i,n} - t_{j,n}|^\alpha}. \end{aligned}$$

By using the facts that $|s - t|^{\beta-\alpha} \leq |\mathbb{T}_n|^{\beta-\alpha}$ and $|t_{i,n} - t_{j,n}| \geq |\mathbb{T}_n|/c_{\mathbb{T}}$ for all $i, j \in \{0, \dots, k_n\}$ with $i \neq j$ and $s, t \in [t_{j,n}, t_{j+1,n}]$, we see that

$$P\left(\sup_{s,t \in [r,T]: s \neq t} \frac{|{}_n X_s - {}_n X_t|}{|s - t|^\alpha} \geq \varepsilon\right) \leq c_\beta (4/\varepsilon)^p |\mathbb{T}_n|^{(\beta-\alpha)p} + P\left(\max_{j \in \{1, \dots, k_n\}} |{}_n X_{t_{j,n}}| / |\mathbb{T}_n|^\alpha > (\varepsilon/4) c_{\mathbb{T}}^{-\alpha}\right)$$

for any $\varepsilon > 0$. As the terms on the right-hand side converge to zero as $n \uparrow \infty$, the assertion is shown. \square

Remark 14. Let $p \geq 1$ and $c_0 \geq 0$ be such that $E[|{}_n X_s - {}_n X_t|^p] \leq c_0 |s - t|^{1+q}$ for all $n \in \mathbb{N}$, $j \in \{0, \dots, k_n - 1\}$ and $s, t \in [t_{j,n}, t_{j+1,n}]$. Then Chebyshev’s inequality in combination with Proposition 12 ensures that condition (3.6) is satisfied.

3.4. Adapted linear interpolations of brownian motion

We study the sequence $({}_n W)_{n \in \mathbb{N}}$ of adapted linear interpolations of W that are given by (2.5) and whose paths lie in $H_r^1([0, T], \mathbb{R}^d)$. To this end, we introduce the following notation. For given $n \in \mathbb{N}$ and $t \in [r, T)$, let $i \in \{0, \dots, k_n - 1\}$ be such that $t \in [t_{i,n}, t_{i+1,n})$, then we set

$$\underline{t}_n := t_{(i-1) \vee 0, n}, \quad t_n := t_{i,n} \quad \text{and} \quad \bar{t}_n := t_{i+1, n}.$$

That is, \underline{t}_n is the predecessor of t_n with respect to \mathbb{T}_n , unless $i = 0$, and \bar{t}_n is the successor of t_n . We also set $\underline{T}_n := t_{k_n-1, n}$, $T_n := T$ and $\bar{T}_n := T$. In addition, we use the following abbreviations:

$$\Delta t_{i,n} := t_{i,n} - t_{(i-1) \vee 0, n} \quad \text{and} \quad \Delta W_{t_{i,n}} := W_{t_{i,n}} - W_{t_{(i-1) \vee 0, n}}$$

for each $i \in \{0, \dots, k_n\}$. After these preparations, let us begin with a general integral representation.

Lemma 15. Let $n \in \mathbb{N}$ and $s, t \in [r, T]$ be such that $s < t$. Then each $\mathbb{R}^{m \times d}$ -valued progressively measurable process X satisfies

$$\int_s^t X_{\underline{u}_n} d_n W_u = \frac{t - s}{\Delta t_{i+1, n}} \int_{t_{i-1, n}}^{t_{i, n}} X_{u_n} dW_u \quad \text{a.s.,}$$

whenever $i \in \{1, \dots, k_n - 1\}$ is such that $s, t \in [t_{i,n}, t_{i+1,n})$, and

$$\int_s^t X_{\underline{u}_n} d_n W_u = \frac{t_{i+1, n} - s}{\Delta t_{i+1, n}} \int_{t_{i-1, n}}^{t_{i, n}} X_{u_n} dW_u + \int_{t_{i, n}}^{t_{j-1, n}} X_{u_n} dW_u + \frac{t - t_{j, n}}{\Delta t_{j+1, n}} \int_{t_{j-1, n}}^{t_{j, n}} X_{u_n} dW_u \quad \text{a.s.,}$$

if $i, j \in \{1, \dots, k_n - 1\}$ are such that $i < j$, $s \in [t_{i,n}, t_{i+1,n})$ and $t \in [t_{j,n}, t_{j+1,n})$.

Proof. The first identity follows immediately from the definition of ${}_n W$. To obtain the second, we simply use the decomposition

$$\int_s^t X_{\underline{u}_n} d_n W_u = \int_s^{t_{i+1, n}} X_{\underline{u}_n} d_n W_u + \sum_{k=i+1}^{j-1} \int_{t_{k, n}}^{t_{k+1, n}} X_{\underline{u}_n} d_n W_u + \int_{t_{j, n}}^t X_{\underline{u}_n} d_n W_u$$

together with the first identity. \square

Let us recall an explicit moment estimate for stochastic integrals driven by W from [15] [Theorem 7.2]. For $p \geq 2$ we set $w_p := ((p^3/2)/(p - 1))^{p/2}$, then for any $\mathbb{R}^{m \times d}$ -valued progressively measurable process X with $\int_r^T E[|X_u|^p] du < \infty$,

$$E \left[\sup_{v \in [s, t]} \left| \int_s^v X_u dW_u \right|^p \right] \leq w_p (t - s)^{p/2-1} \int_s^t E[|X_u|^p] du \tag{3.7}$$

for all $s, t \in [r, T]$ with $s \leq t$. We derive a corresponding result for the sequence $({}_nW)_{n \in \mathbb{N}}$ of adapted linear interpolations of W .

Proposition 16. *For each $p \geq 2$ there is $\hat{w}_p > 0$ such that each $\mathbb{R}^{m \times d}$ -valued progressively measurable process X satisfies*

$$E \left[\max_{v \in [s, t]} \left| \int_s^v X_{\underline{u}_n} d_n W_u \right|^p \right] \leq \hat{w}_p (t - s)^{p/2} \max_{j \in \{0, \dots, k_n\}; t_{j,n} \in [s_n, t_n]} E[|X_{t_{j,n}}|^p]$$

for each $n \in \mathbb{N}$ and $s, t \in [r, T]$ with $s \leq t$.

Proof. We assume that $E[|X_{t_{j,n}}|^p] < \infty$ for all $j \in \{0, \dots, k_n\}$ with $t_{j,n} \in [s_n, t_n]$, as otherwise there is nothing to show. First, if $t \leq t_{1,n}$, then $\int_s^v X_{\underline{u}_n} d_n W_u = 0$ for each $v \in [s, t]$. For $s < t_{1,n}$ and $t \geq t_{1,n}$ we have

$$\int_s^v X_{\underline{u}_n} d_n W_u = \int_{t_{1,n}}^v X_{\underline{u}_n} d_n W_u \quad \text{for all } v \in [t_{1,n}, t].$$

Thus, let us use Lemma 15 and assume at first that $s, t \in [t_{i,n}, t_{i+1,n}]$ for some $i \in \{1, \dots, k_n - 1\}$. Then (3.7) yields that

$$E \left[\max_{v \in [s, t]} \left| \int_s^v X_{\underline{u}_n} d_n W_u \right|^p \right] \leq w_p c_{\mathbb{T}}^{p/2} (t - s)^{p/2} E[|X_{t_{i-1,n}}|^p],$$

where $c_{\mathbb{T}}$ is the constant appearing in (2.3). Let now $i, j \in \{1, \dots, k_n - 1\}$ be such that $i < j$, $s \in [t_{i,n}, t_{i+1,n}]$ and $t \in [t_{j,n}, t_{j+1,n}]$, then

$$\begin{aligned} \max_{v \in [s, t]} \left| \int_s^v X_{\underline{u}_n} d_n W_u \right| &\leq \frac{t_{i+1,n} - s}{\Delta t_{i+1,n}} \left| \int_{t_{i-1,n}}^{t_{i,n}} X_{u_n} dW_u \right| \\ &+ \max_{k \in \{i, \dots, j-1\}} \left| \int_{t_{i,n}}^{t_{k,n}} X_{u_n} dW_u \right| + \frac{t - t_{j,n}}{\Delta t_{j+1,n}} \left| \int_{t_{j-1,n}}^{t_{j,n}} X_{u_n} dW_u \right| \quad \text{a.s.} \end{aligned}$$

This is due to Lemma 15, which asserts that the adapted process $[s, t] \times \Omega \rightarrow \mathbb{R}^m, (v, \omega) \mapsto \int_s^v X_{\underline{u}_n}(\omega) d_n W_u(\omega)$ is piecewise linear. Hence, from (3.7) we obtain that

$$E \left[\max_{v \in [s, t]} \left| \int_s^v X_{\underline{u}_n} d_n W_u \right|^p \right] \leq \hat{w}_p (t - s)^{p/2} \max_{k \in \{i-1, \dots, j-1\}} E[|X_{t_{k,n}}|^p]$$

for $\hat{w}_p := 3^p w_p c_{\mathbb{T}}^{p/2}$, which yields the claim. \square

Finally, we derive an explicit integral moment estimate for $({}_nW)_{n \in \mathbb{N}}$.

Lemma 17. *For each $p, q \geq 1$ there exists $\hat{w}_{p,q} > 0$ satisfying*

$$E \left[\left(\int_s^t |{}_n \dot{W}_u|^q du \right)^p \right] \leq \hat{w}_{p,q} |\mathbb{T}_n|^{-pq/2} (t - s)^p \tag{3.8}$$

for all $n \in \mathbb{N}$ and $s, t \in [r, T]$ with $s \leq t$.

Proof. Since ${}_nW$ is constant on $[r, t_{1,n}]$, we let at first $s, t \in [t_{i,n}, t_{i+1,n}]$ for some $i \in \{1, \dots, k_n - 1\}$ and Z be an \mathbb{R}^d -valued random vector such that $Z \sim \mathcal{N}(0, \mathbb{I}_d)$. Then

$$E \left[\left(\int_s^t |{}_n\dot{W}_u|^q du \right)^p \right] = E[|Z|^{pq}] \frac{(\Delta t_{i,n})^{pq/2}}{(\Delta t_{i+1,n})^{pq}} (t-s)^p \leq \hat{w}_{p,q} |\mathbb{T}_n|^{-pq/2} (t-s)^p$$

for $\hat{w}_{p,q} := E[|Z|^{pq}] c_{\mathbb{T}}^{pq}$, where $c_{\mathbb{T}}$ is the constant in (2.3). Next, assume instead $i, j \in \{1, \dots, k_n - 1\}$ are such that $i < j, s \in [t_{i,n}, t_{i+1,n}]$ and $t \in [t_{j,n}, t_{j+1,n}]$. In this case,

$$\left(E \left[\left(\int_s^t |{}_n\dot{W}_u|^q du \right)^p \right] \right)^{1/p} \leq \hat{w}_{p,q}^{1/p} |\mathbb{T}_n|^{-q/2} (t-s),$$

by the triangle inequality. Therefore, the assertion holds. \square

3.5. Auxiliary convergence results

In this section we provide interpolation error estimates in the supremum norm for stochastic processes and several moment estimates, required to prove (2.13).

Lemma 18. *Let $n \in \mathbb{N}$ and $x : [0, T] \rightarrow \mathbb{R}^m$. Then the map $L_n(x) : [0, T] \rightarrow \mathbb{R}^m$ given at (2.4) satisfies $\|L_n(x)^t\|_\infty \leq \|x^r\|_\infty \vee \max_{j \in \{1, \dots, k_n - 1\}: t_{j,n} < t} |x(t_{j,n})|$ and*

$$\|L_n(x)^t - x^t\|_\infty \leq \max_{j \in \{0, \dots, k_n - 1\}: t_{j,n} \leq t} \sup_{s \in [t_{j,n}, t_{j+1,n}]} |x(t_{(j-1) \vee 0, n}) - x^t(s)| \vee |x(t_{j,n}) - x^t(s)|$$

for each $t \in [t_{1,n}, T]$.

Proof. Fix $s \in [t_{1,n}, t]$ and let $i \in \{1, \dots, k_n - 1\}$ be such that $s \in [t_{i,n}, t_{i+1,n}]$, then $|L_n(x)(s)| \leq |x(t_{i-1,n})| \vee |x(t_{i,n})|$, since $L_n(x)$ is linear on $[t_{i,n}, t_{i+1,n}]$. In addition,

$$\begin{aligned} |L_n(x)(s) - x(s)| &\leq \frac{t_{i+1,n} - s}{\Delta t_{i+1,n}} |x(t_{i-1,n}) - x(s)| + \frac{s - t_{i,n}}{\Delta t_{i+1,n}} |x(t_{i,n}) - x(s)| \\ &\leq \sup_{u \in [t_{i,n}, t_{i+1,n}]} |x(t_{i-1,n}) - x^t(u)| \vee |x(t_{i,n}) - x^t(u)|, \end{aligned}$$

which is readily seen, and the assertions follow. \square

In combination with Proposition 12, this directly gives the following result.

Lemma 19. *Let $({}_nX)_{n \in \mathbb{N}}$ be a sequence of \mathbb{R}^m -valued right-continuous processes for which there are $c_0 \geq 0, p \geq 1$ and $q > 0$ such that*

$$E[|{}_nX_s - {}_nX_t|^p] \leq c_0 |s - t|^{1+q}$$

for all $n \in \mathbb{N}, j \in \{0, \dots, k_n - 1\}$ and $s, t \in [t_{j,n}, t_{j+1,n}]$. Then there is $c_{p,q} > 0$ such that

$$E[\|L_n({}_nX) - {}_nX\|_\infty^p] \leq c_{p,q} c_0 |\mathbb{T}_n|^q \quad \text{for any } n \in \mathbb{N}.$$

Proof. For $\alpha \in [0, 1]$ Lemma 18 entails that

$$\begin{aligned} \|L_n({}_nX) - {}_nX\|_\infty &\leq |\mathbb{T}_n|^\alpha \max_{j \in \{0, \dots, k_n - 1\}} \sup_{s, t \in [t_{j,n}, t_{j+1,n}]: s \neq t} \frac{|{}_nX_s - {}_nX_t|}{|s - t|^\alpha} \\ &\quad + \max_{j \in \{1, \dots, k_n - 1\}} |{}_nX_{t_{j,n}} - {}_nX_{t_{j-1,n}}|. \end{aligned}$$

Let additionally $\alpha < q/p$. Since we have $\sum_{j=0}^{k_n-1} (t_{j+1,n} - t_{j,n}) = T - r$, it follows by virtue of Proposition 12 that

$$E \left[\max_{j \in \{0, \dots, k_n-1\}} \sup_{s, t \in [t_{j,n}, t_{j+1,n}]: s \neq t} \frac{|{}_n X_s - {}_n X_t|^p}{|s - t|^{\alpha p}} \right] \leq k_{\alpha, p, q} c_0 (T - r) |\mathbb{T}_n|^{q - \alpha p},$$

where the constant $k_{\alpha, p, q}$ is given by (3.2). Moreover,

$$E \left[\max_{j \in \{1, \dots, k_n-1\}} |{}_n X_{t_{j,n}} - {}_n X_{t_{j-1,n}}|^p \right] \leq c_0 (T - r) |\mathbb{T}_n|^q.$$

As the function $[0, q/p] \rightarrow (2^p, \infty)$, $\beta \mapsto k_{\beta, p, q}$ is strictly increasing, we choose $\alpha = 0$ and set $c_{p, q} := 2^{p-1} (1 + k_{0, p, q}) (T - r)$, which completes the proof. \square

A consequence of Lemma 17 is the following moment bound.

Lemma 20. *Let $({}_n X)_{n \in \mathbb{N}}$ be a sequence of \mathbb{R}_+ -valued measurable processes for which there are $p > 2$ and $c_p > 0$ such that $E[{}_n X_s^p] \leq c_p |\mathbb{T}_n|^p$ for each $s \in [r, T)$ and $n \in \mathbb{N}$. Then there is $c_2 > 0$ satisfying*

$$E \left[\left(\int_r^T {}_n X_s |{}_n \dot{W}_s| ds \right)^2 \right] \leq c_2 |\mathbb{T}_n| \quad \text{for all } n \in \mathbb{N}.$$

Proof. Let $q > 1$ satisfy $2/p + 1/q = 1$, then it follows from the inequalities of Cauchy–Schwarz and Hölder that

$$E \left[\left(\int_r^T {}_n X_s |{}_n \dot{W}_s| ds \right)^2 \right] \leq \left(E \left[\left(\int_r^T {}_n X_s^2 ds \right)^{p/2} \right] \right)^{2/p} c_{2,1} |\mathbb{T}_n|^{-1} \leq c_2 |\mathbb{T}_n|,$$

where we have set $c_{2,1} := \hat{w}_{q,2}^{1/q} (T - r)$ and $c_2 := c_p^{2/p} (T - r) c_{2,1}$, by using the constant $\hat{w}_{q,2}$ constructed in Lemma 17. \square

To shorten the notation for the next and several other estimates in Section 5, we introduce for each $n \in \mathbb{N}$ the function $\gamma_n : [r, T] \rightarrow [0, c_{\mathbb{T}}]$ defined via

$$\gamma_n(s) := \frac{\Delta s_n}{\Delta \bar{s}_n}. \tag{3.9}$$

So, γ_n vanishes on $[r, t_{1,n})$ and agrees with the constant $\Delta t_{i,n} / \Delta t_{i+1,n}$ on $[t_{i,n}, t_{i+1,n})$ for each $i \in \{1, \dots, k_n - 1\}$ and we have $\gamma_n(T) = 1$.

Lemma 21. *Let $F : [r, T] \times C([0, T], \mathbb{R}^m) \rightarrow \mathbb{R}^m$ be d_∞ -Lipschitz continuous and $({}_n Y)_{n \in \mathbb{N}}$ be a sequence in $\mathcal{C}([0, T], \mathbb{R}^m)$ for which there are $c_0, c_{2,0} \geq 0$ such that $|F(t, x)| \leq c_0 (1 + \|x\|_\infty)$ and*

$$E[\|{}_n Y\|_\infty^2] + E[\|{}_n Y^s - {}_n Y^t\|_\infty^2] / |s - t| \leq c_{2,0} (1 + E[\|{}_n Y^r\|_\infty^2])$$

for all $n \in \mathbb{N}$, $s, t \in [r, T]$ with $s \neq t$ and $x \in C([0, T], \mathbb{R}^m)$. Then there is $c_2 > 0$ satisfying

$$E \left[\max_{j \in \{0, \dots, k_n\}} \left| \int_r^{t_{j,n}} F(s_n, {}_n Y)(\gamma_n(s) - 1) ds \right|^2 \right] \leq c_2 |\mathbb{T}_n| (1 + E[\|{}_n Y^r\|_\infty^2])$$

for each $n \in \mathbb{N}$.

Proof. We may assume that $E[\|_n Y^r\|_\infty^2] < \infty$ and, by decomposing the integral, we can rewrite that

$$\int_r^{t_j,n} F(\underline{s}_n, {}_n Y) \gamma_n(s) ds = \int_r^{t_{j-1},n} F(s_n, {}_n Y) ds$$

for each $j \in \{1, \dots, k_n\}$. Thus, let $\lambda_0 \geq 0$ be a Lipschitz constant for F , then

$$E \left[\max_{j \in \{1, \dots, k_n\}} \left| \int_r^{t_{j-1},n} F(s_n, {}_n Y) - F(\underline{s}_n, {}_n Y) ds \right|^2 \right] \leq c_{2,1} |\mathbb{T}_n| (1 + E[\|_n Y^r\|_\infty^2])$$

with $c_{2,1} := 2(T - r)^2 \lambda_0^2 (1 + c_{2,0})$. In addition, we estimate that

$$E \left[\max_{j \in \{1, \dots, k_n\}} \left| \int_{t_{j-1},n}^{t_j,n} F(\underline{s}_n, {}_n Y) ds \right|^2 \right] \leq c_{2,2} |\mathbb{T}_n|^2 (1 + E[\|_n Y^r\|_\infty^2]),$$

where $c_{2,2} := 2c_0^2 (1 + c_{2,0})$. So, the constant $c_2 := 2(c_{2,1} + (T - r)c_{2,2})$ yields the claim. \square

The last moment estimate involves integrals with respect to ${}_n W$ and W , where $n \in \mathbb{N}$, and it extends Lemma 3.2 in [16]:

Proposition 22. *Let $F : [r, T] \times C([0, T], \mathbb{R}^m) \rightarrow \mathbb{R}^{m \times d}$ be d_∞ -Lipschitz continuous and $({}_n Y)_{n \in \mathbb{N}}$ be a sequence in $\mathcal{C}([0, T], \mathbb{R}^m)$. Suppose there are $c_0 \geq 0$, $p \geq 2$ and $c_{p,0} \geq 0$ such that $|F(t, x)| \leq c_0(1 + \|x\|_\infty)$ and*

$$E[\|_n Y\|_\infty^p] + E[\|_n Y^s - {}_n Y^t\|_\infty^p] / |s - t|^{p/2} \leq c_{p,0} (1 + E[\|_n Y^r\|_\infty^p])$$

for each $n \in \mathbb{N}$, $s, t \in [r, T]$ with $s \neq t$ and $x \in C([0, T], \mathbb{R}^m)$. Then there is $c_p > 0$ such that

$$E \left[\max_{j \in \{0, \dots, k_n\}} \left| \int_r^{t_j,n} F(\underline{s}_n, {}_n Y) d({}_n W_s - W_s) \right|^p \right] \leq c_p |\mathbb{T}_n|^{p/2-1} (1 + E[\|_n Y^r\|_\infty^p])$$

for any $n \in \mathbb{N}$.

Proof. We let $E[\|_n Y^r\|_\infty^p] < \infty$ and apply Lemma 15 to get that

$$\int_r^{t_j,n} F(\underline{s}_n, {}_n Y) d{}_n W_s = \int_r^{t_{j-1},n} F(s_n, {}_n Y) dW_s \quad \text{a.s.}$$

for all $j \in \{1, \dots, k_n\}$. Let $\lambda_0 \geq 0$ denote a Lipschitz constant for F , then

$$E \left[\max_{j \in \{1, \dots, k_n\}} \left| \int_r^{t_{j-1},n} F(s_n, {}_n Y) - F(\underline{s}_n, {}_n Y) dW_s \right|^p \right] \leq c_{p,1} |\mathbb{T}_n|^{p/2} (1 + E[\|_n Y^r\|_\infty^p])$$

with $c_{p,1} := 2^{p-1} w_p (T - r)^{p/2} \lambda_0^p (1 + c_{p,0})$, where w_p satisfies (3.7). Moreover,

$$E \left[\max_{j \in \{1, \dots, k_n\}} \left| \int_{t_{j-1},n}^{t_j,n} F(\underline{s}_n, {}_n Y) dW_s \right|^p \right] \leq \sum_{j=1}^{k_n} E \left[\left| \int_{t_{j-1},n}^{t_j,n} F(\underline{s}_n, {}_n Y) dW_s \right|^p \right] \leq c_{p,2} |\mathbb{T}_n|^{p/2-1} (1 + E[\|_n Y^r\|_\infty^p])$$

for $c_{p,2} := 2^{p-1} w_p (T - r) c_0^p (1 + c_{p,0})$. So, we set $c_p := 2^{p-1} ((T - r)c_{p,1} + c_{p,2})$ and obtain the asserted estimate. \square

4. Path-dependent ODEs and SDEs: proofs

We give existence and uniqueness proofs for mild solutions to path-dependent ODEs in Section 2.2 and strong solutions to path-dependent SDEs in Section 2.3.

4.1. Proof of Proposition 3

Let us first derive a global estimate for mild solutions to the ODE (2.1).

Lemma 23. Under (C.1), any mild solution x to (2.1) satisfies

$$\|x^t\|_{H,r}^2 \leq c_H e^{c_H \int_r^t c_0(s)^2 ds} \left(\|x^r\|_\infty^2 + \int_r^t c_0(s)^2 ds \right) \tag{4.1}$$

for each $t \in [r, T]$ with $c_H := 2^2 \max\{1, T - r\}$.

Proof. We readily estimate that

$$\|x^t\|_{H,r}^2 \leq 2\|x^r\|_\infty^2 + 2 \int_r^t c_0(s)^2 \left(1 + \|x^r\|_\infty + \int_r^s |\dot{x}(u)| du \right)^2 ds$$

for all $t \in [r, T]$, which shows that $\|x\|_{H,r}$ is finite. Hence, the claim follows from Gronwall’s inequality. \square

Now we check the uniqueness of mild solutions, which implies uniqueness for classical solutions.

Lemma 24. Assume that (C.1) and (C.2) hold. Then any two mild solutions x and \tilde{x} to (2.1) that satisfy $x^r = \tilde{x}^r$ must coincide.

Proof. By Lemma 23, there is $n \in \mathbb{N}$ such that $\|x\|_{H,r} \vee \|\tilde{x}\|_{H,r} \leq n$. Thus,

$$\|x^t - \tilde{x}^t\|_{H,r}^2 \leq \int_r^t \lambda_n^2(s) \|x^s - \tilde{x}^s\|_{H,r}^2 ds$$

for every $t \in [r, T]$ and Gronwall’s inequality implies that $x = \tilde{x}$. \square

Proof of Proposition 3. Since uniqueness follows from Lemma 24, we directly turn to the existence assertion and define \mathcal{H} to be the (bounded) set of all $x \in H_r^1([0, T], \mathbb{R}^m)$ satisfying $x^r = \hat{x}^r$ and the estimate (4.1).

According to Lemma 23, a path $x \in C([0, T], \mathbb{R}^m)$ is a mild solution to (2.1) such that $x^r = \hat{x}^r$ if and only if $x \in \mathcal{H}$ and it is a fixed-point of the operator $\psi : \mathcal{H} \rightarrow H_r^1([0, T], \mathbb{R}^m)$ given by

$$\psi(y)(t) := x_0(t) + \int_r^{r \vee t} F(s, y) ds.$$

We remark that ψ maps \mathcal{H} into itself. Indeed, this follows by inserting (4.1) into the inequality $\|\psi(x)\|_{H,r}^2 \leq c_H \|x_0\|_\infty^2 + c_H \int_r^t c_0(s)^2 (1 + \|x^s\|_{H,r}^2) ds$, valid for every $x \in \mathcal{H}$ and $t \in [r, T]$. Because $x_n = \psi(x_{n-1})$ for each $n \in \mathbb{N}$, we now know that $(x_n)_{n \in \mathbb{N}_0}$ is a sequence in \mathcal{H} .

Next, let us choose $l \in \mathbb{N}$ such that $\|x\|_{H,r} \leq l$ for all $x \in \mathcal{H}$. Then we get that $\|\psi(x)^t - \psi(\tilde{x})^t\|_{H,r}^2 \leq \int_r^t \lambda_l(s)^2 \|x^s - \tilde{x}^s\|_{H,r}^2 ds$ for any $x, \tilde{x} \in \mathcal{H}$ and $t \in [r, T]$, which

shows that ψ is $\|\cdot\|_{H,r}$ -Lipschitz continuous. It follows inductively that

$$\|x_{n+1}^t - x_n^t\|_{H,r}^2 \leq \frac{\delta^2}{n!} \left(\int_r^t \lambda_l(s)^2 ds \right)^n$$

for every $n \in \mathbb{N}_0$ and $t \in [r, T]$, where we have set $\delta := \|\psi(x_0) - x_0\|_{H,r}$. Hence, the triangle inequality gives us that

$$\|x_n - x_k\|_{H,r} \leq \delta \sum_{i=k}^{n-1} \left(\frac{1}{i!} \right)^{1/2} \left(\int_r^T \lambda_l(s)^2 ds \right)^{i/2}$$

for all $k, n \in \mathbb{N}_0$ with $k < n$. The ratio test yields that the series $\sum_{i=0}^\infty (1/i!)^{1/2} u^{i/2}$ converges absolutely for all $u \geq 0$. So, $\lim_{k \uparrow \infty} \sup_{n \in \mathbb{N}: n \geq k} \|x_n - x_k\|_{H,r} = 0$.

Since \mathcal{H} is closed with respect to the complete norm $\|\cdot\|_{H,r}$, there is a unique $x \in \mathcal{H}$ such that $\lim_{n \uparrow \infty} \|x_n - x\|_{H,r} = 0$. Lipschitz continuity of ψ implies that $\lim_{n \uparrow \infty} \|x_{n+1} - \psi(x)\|_{H,r} = 0$. For this reason, $x = \psi(x)$ and the proposition is established. \square

4.2. Proof of Proposition 5

At first, let us deduce a global estimate for any solution to the SDE (1.1). For this purpose, whenever $p > 4$ and (C.3) holds, we set $c_p := (\int_r^T c_0(s)^2 ds)^{p/2} + \tilde{c}_0^p w_p$, where w_p is the constant appearing in (3.7).

Lemma 25. *Let (C.3) be valid. Then for each $p > 4$ and $\alpha \in [0, 1/2 - 2/p]$, any solution X to (1.1) satisfies*

$$E[\|X^t\|_{\alpha,r}^p] \leq c_{\alpha,p} e^{c_{\alpha,p} c_p (t-r)} (E[\|X^r\|_\infty^p] + c_p(t-r)) \tag{4.2}$$

for all $t \in [r, T]$ with $c_{\alpha,p} := 8^{p-1} \max\{1, T-r\}^{p/2-1} k_{\alpha,p,p/2-2}$, where $k_{\alpha,p,p/2-2}$ is given by (3.2) for the choice $q = p/2 - 2$.

Proof. We assume that $E[\|X^r\|_\infty^p] < \infty$ and let $n \in \mathbb{N}$. Then the stopping time $\tau_n := \inf\{t \in [0, T] \mid |X_t| \geq n\} \vee r$ satisfies $\|X^{\tau_n}\|_\infty \leq \|X^r\|_\infty \vee n$ and we get that

$$E[|X_u^{\tau_n} - X_v^{\tau_n}|^p] \leq 4^{p-1} c_p (v-u)^{p/2-1} \int_r^t 1 + E[\|X^{s \wedge \tau_n}\|_\infty^p] ds$$

for every $t \in [r, T]$ and $u, v \in [r, t]$ with $u \leq v$, by the inequalities of Jensen and Cauchy–Schwarz and (3.7). Therefore, it follows from Proposition 12 that

$$E[\|X^{t \wedge \tau_n}\|_{\alpha,r}^p] \leq 2^{p-1} E[\|X^r\|_\infty^p] + 8^{p-1} k_{\alpha,p,p/2-2} c_p (t-r)^{p(1/2-\alpha)-1} \int_r^t 1 + E[\|X^{s \wedge \tau_n}\|_\infty^p] ds,$$

showing in particular that $E[\|X^{\tau_n}\|_{\alpha,r}^p]$ is finite. Thus, Gronwall’s inequality and Fatou’s lemma lead to the claimed estimate. \square

Remark 26. If instead $\alpha \in [0, 1/2]$ and $p \geq 1$ satisfy $\alpha \geq 1/2 - 2/p$, then we still have that $E[\|X\|_{\alpha,r}^p] \leq (E[\|X\|_{\alpha,r}^q])^{p/q} < \infty$ for any $q > \max\{p, 4\}$ with $\alpha < 1/2 - 2/q$, by Hölder’s inequality.

Lemma 27. *Under (C.5), pathwise uniqueness holds for (1.1).*

Proof. Let X and \tilde{X} be two weak solutions to (1.1) defined on a common filtered probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, (\tilde{\mathcal{F}}_t)_{t \in [0, T]}, \tilde{P})$ that satisfies the usual conditions and on which there is a standard d -dimensional $(\tilde{\mathcal{F}}_t)_{t \in [0, T]}$ -Brownian motion \tilde{W} such that $X^r = \tilde{X}^r$ a.s.

For fixed $n \in \mathbb{N}$ we set $\tau_n := \inf\{t \in [0, T] \mid |X_t| \geq n \text{ or } |\tilde{X}_t| \geq n\} \vee r$, ensuring that $\|X^{\tau_n}\|_\infty \vee \|\tilde{X}^{\tau_n}\|_\infty \leq \|X^r\|_\infty \vee n$ and $\|X^{\tau_n} - \tilde{X}^{\tau_n}\|_\infty \leq 2n$ a.s. Then with $c_2 := 2(T - r + w_2)$ we obtain that

$$\tilde{E}[\|X^{t \wedge \tau_n} - \tilde{X}^{t \wedge \tau_n}\|_\infty^2] \leq c_2 \int_r^t \lambda_n(s)^2 \tilde{E}[\|X^{s \wedge \tau_n} - \tilde{X}^{s \wedge \tau_n}\|_\infty^2] ds$$

for any $t \in [r, T]$. So, $X^{\tau_n} = \tilde{X}^{\tau_n}$ a.s., by Gronwall’s inequality. As $\tau_n \leq \tau_{n+1}$ for all $n \in \mathbb{N}$ and $\sup_{n \in \mathbb{N}} \tau_n = \infty$, we get that $X_t = \lim_{n \uparrow \infty} X_t^{\tau_n} = \lim_{n \uparrow \infty} \tilde{X}_t^{\tau_n} = \tilde{X}_t$ a.s. for all $t \in [r, T]$. Right-continuity of paths implies that $X = \tilde{X}$ a.s. \square

Proof of Proposition 5. Let us define \mathcal{H} to be the closed and bounded set of all $X \in \mathcal{C}([0, T], \mathbb{R}^m)$ satisfying $X^r = \hat{x}^r$ a.s. and the estimate (4.2) for any $p > 4$ and $\alpha \in [0, 1/2 - 2/p)$. Then Remark 26 entails that $\mathcal{H} \subset \mathcal{C}_{r, \infty}^{1/2-}([0, T], \mathbb{R}^m)$.

By Lemma 25, a process $X \in \mathcal{C}([0, T], \mathbb{R}^m)$ is a solution to (1.1) satisfying $X^r = \hat{x}^r$ a.s. if and only if $X \in \mathcal{H}$ and it is an a.s. fixed-point of the operator $\Psi : \mathcal{H} \rightarrow \mathcal{C}_{r, \infty}^{1/2-}([0, T], \mathbb{R}^m)$ specified by requiring that

$$\Psi(Y)_t = {}_0X_t + \int_r^{r \vee t} b(s, Y) ds + \int_r^{r \vee t} \sigma(s, Y) dW_s$$

for all $t \in [0, T]$ a.s. We stress the fact that, due to Proposition 12, for every $X \in \mathcal{H}$, $p > 4$ and $\alpha \in [0, 1/2 - 2/p)$ it follows that

$$E[\|\Psi(X)^t\|_{\alpha, r}^p] \leq c_{\alpha, p} E[\|{}_0X\|_\infty^p] + c_{\alpha, p} c_p \int_r^t 1 + E[\|X^s\|_{\alpha, r}^p] ds$$

for each $t \in [r, T]$. Thus, $\Psi(\mathcal{H}) \subset \mathcal{H}$ follows from plugging (4.2) into the above inequality. Since ${}_nX = \Psi({}_{n-1}X)$ a.s. for all $n \in \mathbb{N}$, we have shown that $({}_nX)_{n \in \mathbb{N}_0}$ is a sequence in \mathcal{H} .

Next, we choose $\alpha \in [0, 1/2)$ and $p > 4$ such that $\alpha_0 \leq \alpha < 1/2 - 2/p$, where α_0 is the constant in (C.4). We set $\bar{c}_p := (\int_r^T \lambda_0(s)^2 ds)^{p/2} + \tilde{\lambda}_0^p w_p$, then it follows from Proposition 12 that any $X, \tilde{X} \in \mathcal{H}$ satisfy

$$E[\|\Psi(X)^t - \Psi(\tilde{X})^t\|_{\alpha, r}^p] \leq c_{\alpha, p} \bar{c}_p \int_r^t E[\|X^s - \tilde{X}^s\|_{\alpha, r}^p] ds$$

for all $t \in [r, T]$, since we can use that $\|x\|_{\alpha_0, r} \leq \max\{1, T - r\}^{\alpha - \alpha_0} \|x\|_{\alpha, r}$ for every $x \in C_r^\alpha([0, T], \mathbb{R}^m)$. Hence, Gronwall’s inequality entails that there is at most a unique solution X to (1.1) such that $X^r = \hat{x}^r$ a.s.

We infer from the above inequality that Ψ is Lipschitz continuous with respect to the seminorm (2.2). In addition,

$$E[\|{}_{n+1}X^t - {}_nX^t\|_{\alpha, r}^p] \leq \frac{\delta^p}{n!} (c_{\alpha, p} \bar{c}_p)^n (t - r)^n$$

for each $n \in \mathbb{N}_0$ and $t \in [r, T]$, by induction with $\delta := (E[\|\Psi({}_0X) - {}_0X\|_{\alpha, r}^p])^{1/p}$. Hence, the triangle inequality gives

$$(E[\|{}_nX - {}_kX\|_{\alpha, r}^p])^{1/p} \leq \delta \sum_{i=k}^{n-1} \left(\frac{1}{i!}\right)^{1/p} (c_{\alpha, p} \bar{c}_p)^{i/p} (T - r)^{i/p}$$

for all $k, n \in \mathbb{N}_0$ with $k < n$. The ratio test shows that the series $\sum_{i=0}^\infty (1/i!)^{1/p} u^{i/p}$ converges absolutely for any $u \geq 0$. So, $\lim_{k \uparrow \infty} \sup_{n \in \mathbb{N}: n \geq k} E[\|_n X - {}_k X\|_{\alpha,r}^p] = 0$.

Due to Proposition 11, there exists, up to indistinguishability, a unique process $X \in \mathcal{H}$ such that

$$\lim_{n \uparrow \infty} E[\|_n X - X\|_{\alpha,r}^p] = 0. \tag{4.3}$$

Lipschitz continuity of Ψ implies that $\lim_{n \uparrow \infty} E[\|_{n+1} X - \Psi(X)\|_{\alpha,r}^p] = 0$. For this reason, $X = \Psi(X)$ a.s. As α and p have been arbitrarily chosen, (4.3) must hold for any $\alpha \in [0, 1/2)$ and $p > 4$ such that $\alpha_0 \leq \alpha < 1/2 - 2/p$.

Finally, if $\alpha \in [0, 1/2)$ and $p \geq 1$ are such that $\alpha_0 \leq \alpha < 1/2 - 2/p$ fails, then (4.3) is still true. Indeed, when proving this fact, we may, if necessary, replace α by α_0 to ensure that $\alpha_0 \leq \alpha$ is valid, since we have

$$\|x\|_{\alpha,r} \leq \max\{1, T - r\}^{\alpha_0 - \alpha} \|x\|_{\alpha_0,r} \quad \text{for all } x \in C_r^{\alpha_0}([0, T], \mathbb{R}^m)$$

whenever $\alpha_0 > \alpha$. Next, if $\alpha_0 \leq \alpha$ but $\alpha \geq 1/2 - 2/p$, we take $q > \max\{p, 4\}$ with $\alpha < 1/2 - 2/q$ and use that $E[\|_n X - X\|_{\alpha,r}^p] \leq (E[\|_n X - X\|_{\alpha,r}^q])^{p/q}$ for all $n \in \mathbb{N}$ to infer the desired result, which completes the proof. \square

5. Proof of the main result

5.1. Decomposition into remainder terms

Let us deduce a moment estimate for any solution to (2.9) that is independent of $n \in \mathbb{N}$, which generalizes Proposition 3.1 in [16].

Proposition 28. *Let (C.6) be valid, $h \in H_r^1([0, T], \mathbb{R}^d)$ and \bar{B} be d_∞ -Lipschitz continuous. Then for each $p \geq 2$ there is $c_p > 0$ such that any $n \in \mathbb{N}$ and each solution ${}_n Y$ to (2.9) satisfy*

$$E[\|_n Y\|_\infty^p] + E[\|_n Y^s - {}_n Y^t\|_\infty^p] / |s - t|^{p/2} \leq c_p (1 + E[\|_n Y^r\|_\infty^p]) \tag{5.1}$$

for all $s, t \in [r, T]$ with $s \neq t$.

Proof. We may assume that $E[\|_n Y^r\|_\infty^p]$ is finite and the constant κ appearing in (C.6) is positive. For $l \in \mathbb{N}$ the stopping time $\tau_{l,n} := \inf\{t \in [0, T] \mid |{}_n Y_t| \geq l\} \vee r$ satisfies $\|_n Y^{\tau_{l,n}}\|_\infty \leq \|_n Y^r\|_\infty \vee l$ and for $s, t \in [r, T]$ with $s \leq t$ we have

$$\begin{aligned} & (E[\|_n Y^{s \wedge \tau_{l,n}} - {}_n Y^{t \wedge \tau_{l,n}}\|_\infty^p])^{1/p} \\ & \leq \left(\bar{c}_p (t - s)^{p/2 - 1} \int_s^t 1 + E[\|_n Y^{u \wedge \tau_{l,n}}\|_\infty^{kp}] du \right)^{1/p} \\ & \quad + \left(E \left[\sup_{v \in [s, t]} \left| \int_s^{v \wedge \tau_{l,n}} \bar{B}(u, {}_n Y) d_n W_u \right|^p \right] \right)^{1/p}, \end{aligned} \tag{5.2}$$

where $\bar{c}_p := 6^{p-1}((T - r)^{p/2} + \|h\|_{H,r}^p + w_p)c^p$ and w_p is the constant in (3.7). Lemma 17 provides the constant $\hat{w}_{p/\kappa,1}$ such that (3.8) holds when p and q are replaced by p/κ and 1,

respectively. So,

$$\left(E \left[\left(\int_{\underline{u}_n}^{u \wedge \tau_{l,n}} |\overline{B}(v, {}_n Y)_n \dot{W}_v| dv \right)^{p/\kappa} \right] \right)^\kappa \leq c_{p,1} (u - \underline{u}_n)^{p/2}$$

for any given $u \in [s, T]$ and $c_{p,1} := 2^{p/2} \hat{w}_{p/\kappa,1}^\kappa c^p$. By virtue of (5.2), we may define $\bar{c}_{p/\kappa}$ just as \bar{c}_p above when p is replaced by p/κ and obtain that

$$(E[\|{}_n Y^{u \wedge \tau_{l,n}} - {}_n Y^{\underline{u}_n \wedge \tau_{l,n}}\|_\infty^{p/\kappa}])^\kappa \leq c_{p,2} (u - \underline{u}_n)^{p/2} (1 + E[\|{}_n Y^{u \wedge \tau_{l,n}}\|_\infty^p])^\kappa$$

with $c_{p,2} := 2^{p-1} (\bar{c}_{p/\kappa}^\kappa + c_{p,1})$. Consequently, by letting $\lambda \geq 0$ denote a Lipschitz constant for \overline{B} and using Hölder’s inequality, we can estimate that

$$\begin{aligned} & E \left[\left(\int_s^{t \wedge \tau_{l,n}} |(\overline{B}(u, {}_n Y) - \overline{B}(\underline{u}_n, {}_n Y))_n \dot{W}_u| du \right)^p \right] \\ & \leq (t - s)^{p/2-1} \int_s^t E \left[|\overline{B}(u, {}_n Y^{\tau_{l,n}}) - \overline{B}(\underline{u}_n, {}_n Y^{\tau_{l,n}})|^p \left(\int_s^t |{}_n \dot{W}_v|^2 dv \right)^{p/2} \right] du \\ & \leq c_{p,3} (t - s)^{p/2-1} \int_s^t (1 + E[\|{}_n Y^{u \wedge \tau_{l,n}}\|_\infty^p])^\kappa du \end{aligned}$$

for $c_{p,3} := 2^{3p/2-1} \hat{w}_{(p/2)/(1-\kappa),2}^{1-\kappa} (T - r)^{p/2} \lambda^p (1 + c_{p,2})$. In addition, Proposition 16 directly yields that

$$E \left[\sup_{v \in [s,t]} \left| \int_s^{v \wedge \tau_{l,n}} \overline{B}(\underline{u}_n, {}_n Y) d{}_n W_u \right|^p \right] \leq \hat{w}_p c^p (t - s)^{p/2}.$$

Hence, we set $c_{p,4} := 3^{p-1} (2\bar{c}_p + c_{p,3} + \hat{w}_p c^p)$, then from (5.2) we in total obtain that

$$E[\|{}_n Y^{s \wedge \tau_{l,n}} - {}_n Y^{t \wedge \tau_{l,n}}\|_\infty^p] \leq c_{p,4} (t - s)^{p/2-1} \int_s^t (1 + E[\|{}_n Y^{u \wedge \tau_{l,n}}\|_\infty^p]) du. \tag{5.3}$$

Now Gronwall’s inequality and Fatou’s lemma entail that

$$E[\|{}_n Y^t\|_\infty^p] \leq \liminf_{t \uparrow \infty} E[\|{}_n Y^{t \wedge \tau_{l,n}}\|_\infty^p] \leq c_{p,5} (1 + E[\|{}_n Y^r\|_\infty^p]),$$

where $c_{p,5} := 2^{p-1} \max\{1, T - r\}^{p/2} \max\{1, c_{p,4}\} e^{2^{p-1}(T-r)^{p/2} c_{p,4}}$. Thus, if we set $c_p := (1 + c_{p,4})(1 + c_{p,5})$, the claim follows from (5.3) and an application of Fatou’s lemma. \square

Corollary 29. *Let (C.6) hold and $h \in H_r^1([0, T], \mathbb{R}^d)$. Then for each $p \geq 2$ there is $c_p > 0$ such that any solution Y to (2.10) satisfies*

$$E[\|Y\|_\infty^p] + E[\|Y^s - Y^t\|_\infty^p] / |s - t|^{p/2} \leq c_p (1 + E[\|Y^r\|_\infty^p]) \tag{5.4}$$

for every $s, t \in [r, T]$ with $s \neq t$.

Proof. Since the map R defined via (2.11) is bounded, we may apply Proposition 28 in the case that \underline{B} is replaced by $\underline{B} + R$, \overline{B} is replaced by 0 and Σ is replaced by $\overline{B} + \Sigma$. From this the claim follows immediately. \square

With the derived moment estimates we deduce a crucial decomposition estimate that involves the linear operator L_n and the function γ_n defined in (2.4) and (3.9), respectively.

Proposition 30. *Let (C.6) and (C.7) be satisfied and $h \in H_r^1([0, T], \mathbb{R}^d)$. Then there is $c_2 > 0$ such that for each $n \in \mathbb{N}$ and any solutions ${}_nY$ and Y to (2.9) and (2.10), respectively,*

$$\begin{aligned}
 E\left[\max_{j \in \{0, \dots, k_n\}} |{}_nY_{t_{j,n}} - Y_{t_{j,n}}|^2\right] / c_2 &\leq |\mathbb{T}_n| (1 + E[\|{}_nY^r\|_\infty^2 + \|Y^r\|_\infty^2]) \\
 &+ E[\|{}_nY^r - Y^r\|_\infty^2 + \|L_n({}_nY) - {}_nY\|_\infty^2 + \|L_n(Y) - Y\|_\infty^2] \\
 &+ E\left[\max_{j \in \{0, \dots, k_n\}} \left| \int_r^{t_{j,n}} R(\underline{s}_n, {}_nY)(\gamma_n(s) - 1) ds \right|^2\right] \\
 &+ E\left[\max_{j \in \{0, \dots, k_n\}} \left| \int_r^{t_{j,n}} \overline{B}(\underline{s}_n, {}_nY) d({}_nW_s - W_s) \right|^2\right] \\
 &+ E\left[\max_{j \in \{0, \dots, k_n\}} \left| \int_r^{t_{j,n}} (\overline{B}(s, {}_nY) - \overline{B}(\underline{s}_n, {}_nY)) {}_n\dot{W}_s - R(\underline{s}_n, {}_nY)\gamma_n(s) ds \right|^2\right].
 \end{aligned}$$

Proof. Let $E[\|{}_nY^r\|_\infty^2]$ and $E[\|Y^r\|_\infty^2]$ be finite. We define an increasing function $\varphi_n : [r, T] \rightarrow \mathbb{R}_+$ by

$$\varphi_n(t) := E\left[\max_{j \in \{0, \dots, k_n\}: t_{j,n} \leq t} |{}_nY_{t_{j,n}} - Y_{t_{j,n}}|^2\right]$$

and seek to apply Gronwall’s inequality. For this purpose, we write the difference of ${}_nY$ and Y in the form

$$\begin{aligned}
 {}_nY_t - Y_t &= {}_nY_r - Y_r + \int_r^t \underline{B}(s, {}_nY) - \underline{B}(s, Y) + (B_H(s, {}_nY) - B_H(s, Y))\dot{h}(s) ds \\
 &+ {}_n\Delta_t + \int_r^t \Sigma(s, {}_nY) - \Sigma(s, Y) dW_s
 \end{aligned}$$

for all $t \in [r, T]$ a.s., where the process ${}_n\Delta \in \mathcal{C}([0, T], \mathbb{R}^m)$ is chosen such that

$${}_n\Delta_t = \int_r^t \overline{B}(s, {}_nY) {}_n\dot{W}_s - R(s, Y) ds - \int_r^t \overline{B}(s, Y) dW_s$$

for each $t \in [r, T]$ a.s. Hence, let $\lambda \geq 0$ denote a Lipschitz constant for $\underline{B}(s, \cdot)$, B_H , \overline{B} , R and Σ for any $s \in [r, T]$, then we obtain that

$$\varphi_n(t)^{1/2} \leq \delta_{n,1}^{1/2} + \delta_n(t)^{1/2} + \left(c_{2,1} \int_r^{t_n} \delta_{n,1} + \delta_{n,2}(s) + \varepsilon_n(s) + \varphi_n(s) ds \right)^{1/2} \tag{5.5}$$

for every $t \in [r, T]$, where we have set $c_{2,1} := 15(T - r + \|h\|_{H,r}^2 + w_2)\lambda^2$ and $\delta_{n,1} := E[\|{}_nY^r - Y^r\|_\infty^2]$ and the functions $\delta_n, \delta_{n,2}, \varepsilon_n : [r, T] \rightarrow \mathbb{R}_+$, which are readily seen to be measurable, are defined via

$$\begin{aligned}
 \delta_n(t) &:= E\left[\max_{j \in \{0, \dots, k_n\}: t_{j,n} \leq t} |{}_n\Delta_{t_{j,n}}|^2\right], \\
 \delta_{n,2}(s) &:= E\left[\|L_n({}_nY)^{\xi_n} - {}_nY^{\xi_n}\|_\infty^2 + \|L_n(Y)^{\xi_n} - Y^{\xi_n}\|_\infty^2\right] \quad \text{and} \\
 \varepsilon_n(s) &:= E\left[\|{}_nY^s - {}_nY^{\xi_n}\|_\infty^2 + \|Y^s - Y^{\xi_n}\|_\infty^2\right].
 \end{aligned}$$

In deriving (5.5), we used that $E[\|L_n({}_nY)^{\xi_n} - L_n(Y)^{\xi_n}\|_\infty^2] \leq \delta_{n,1} + \varphi_n(s)$ for all $s \in [r, T]$, which follows from Lemma 18, since L_n is linear.

To estimate the function δ_n we introduce three processes ${}_{n,3}\Delta, {}_{n,4}\Delta, {}_{n,5}\Delta \in \mathcal{C}([0, T], \mathbb{R}^m)$ by setting ${}_{n,3}\Delta_t := \int_r^t R(\underline{s}_n, nY)(\gamma_n(s) - 1) ds$ and

$${}_{n,5}\Delta_t := \int_r^t (\overline{B}(s, nY) - \overline{B}(\underline{s}_n, nY))_n \dot{W}_s - R(\underline{s}_n, nY)\gamma_n(s) ds$$

and requiring that ${}_{n,4}\Delta_t = \int_r^t \overline{B}(\underline{s}_n, nY) d({}_nW_s - W_s)$ for each $t \in [r, T]$ a.s. Then ${}_n\Delta$ can be rewritten in the following way:

$$\begin{aligned} {}_n\Delta_t &= {}_{n,3}\Delta_t + {}_{n,4}\Delta_t + {}_{n,5}\Delta_t + \int_r^t R(\underline{s}_n, nY) - R(s, Y) ds \\ &\quad + \int_r^t \overline{B}(\underline{s}_n, nY) - \overline{B}(s, Y) dW_s \end{aligned}$$

for all $t \in [r, T]$ a.s. Thus, we set $c_{2,2} := 10(T - r + w_2)\lambda^2$, then it follows readily that

$$\begin{aligned} \delta_n(t)^{1/2} &\leq \delta_{n,3}(t)^{1/2} + \delta_{n,4}(t)^{1/2} + \delta_{n,5}(t)^{1/2} \\ &\quad + \left(c_{2,2} \int_r^t \delta_{n,1} + (s - \underline{s}_n) + \delta_{n,2}(s) + \varepsilon_n(s) + \varphi_n(s) ds \right)^{1/2} \end{aligned} \tag{5.6}$$

for each $t \in [r, T]$, where the increasing function $\delta_{n,i} : [r, T] \rightarrow \mathbb{R}_+$ is given by

$$\delta_{n,i}(t) := E \left[\max_{j \in \{0, \dots, k_n\}: t_{j,n} \leq t} |{}_{n,i}\Delta_{t_{j,n}}|^2 \right], \quad \text{for all } i \in \{3, 4, 5\}.$$

Proposition 28 and **Corollary 29** give two constants $c_2, \bar{c}_2 > 0$ satisfying (5.1) and (5.4) for $p = 2$ when c_p is replaced by c_2 and \bar{c}_2 , respectively. Thus, putting (5.5) and (5.6) together, we find that

$$\begin{aligned} \varphi_n(t) &\leq c_{2,4} |\mathbb{T}_n| (1 + E[\|{}_nY^r\|_\infty^2 + \|Y^r\|_\infty^2]) + (5 + c_{2,3}(T - r))\delta_{n,1} \\ &\quad + 5(\delta_{n,3}(t) + \delta_{n,4}(t) + \delta_{n,5}(t)) + c_{2,3} \int_r^t \delta_{n,2}(s) + \varphi_n(s) ds \end{aligned}$$

for given $t \in [r, T]$, where we have at first set $c_{2,3} := 10(c_{2,1} + c_{2,2})$ and then $c_{2,4} := 2(T - r)(1 + c_2 + \bar{c}_2)c_{2,3}$. Consequently,

$$\varphi_n(t)/c_2 \leq |\mathbb{T}_n| (1 + E[\|{}_nY^r\|_\infty^2 + \|Y^r\|_\infty^2]) + \delta_{n,1} + \sum_{i=2}^5 \delta_{n,i}(t)$$

with $c_2 := e^{(T-r)c_{2,3}}(5 + c_{2,4})$, by Gronwall’s inequality. This yields the claim. \square

Thanks to **Lemmas 19** and **21** and **Proposition 22**, only the last remainder in the estimation of **Proposition 30** requires further analysis, before we can prove (2.13). For this reason, we define a map $\Phi_{h,n} : [r, T] \times C([0, T], \mathbb{R}^m) \times C([0, T], \mathbb{R}^d) \rightarrow \mathbb{R}^m$ by

$$\begin{aligned} \Phi_{h,n}(s, y, w) &:= B_H(\underline{s}_n, y)(h(s) - h(\underline{s}_n)) + \overline{B}(\underline{s}_n, y)(L_n(w)(s) - L_n(w)(\underline{s}_n)) \\ &\quad + \Sigma(\underline{s}_n, y)(w(s) - w(\underline{s}_n)) \end{aligned}$$

for given $h \in H_r^1([0, T], \mathbb{R}^d)$ and $n \in \mathbb{N}$. If now ${}_nY$ is a solution to (2.9), then the following decomposition can be used to deal with the remainder in question:

$$\begin{aligned} &(\overline{B}(s, nY) - \overline{B}(\underline{s}_n, nY))_n \dot{W}_s - R(\underline{s}_n, nY)\gamma_n(s) \\ &= (\overline{B}(s, nY) - \overline{B}(\underline{s}_n, nY) - \partial_x \overline{B}(\underline{s}_n, nY)({}_nY_s - {}_nY_{\underline{s}_n}))_n \dot{W}_s \\ &\quad + \partial_x \overline{B}(\underline{s}_n, nY)({}_nY_s - {}_nY_{\underline{s}_n} - \Phi_{h,n}(s, nY, W))_n \dot{W}_s \\ &\quad + \partial_x \overline{B}(\underline{s}_n, nY)\Phi_{h,n}(s, nY, W)_n \dot{W}_s - R(\underline{s}_n, nY)\gamma_n(s) \end{aligned} \tag{5.7}$$

for any $s \in [r, T)$. In fact, in the next two sections we will consider each of these three terms to ensure that (2.13) follows.

5.2. Convergence of the first two remainders

To handle the first remainder term in (5.7), we will use the following estimation in combination with Lemma 20.

Proposition 31. *Let (C.6) be satisfied, $h \in H_r^1([0, T], \mathbb{R}^d)$ and F be a functional on $[r, T) \times C([0, T], \mathbb{R}^m)$ of class $\mathbb{C}^{1,2}$. Assume that \underline{B} and $\partial_x F$ are d_∞ -Lipschitz continuous and there are $c_0, \eta \geq 0$ such that*

$$|\partial_t F(t, x)| + |\partial_{xx} F(t, x)| \leq c_0(1 + \|x\|_\infty^\eta)$$

for all $t \in [r, T)$ and $x \in C([0, T], \mathbb{R}^m)$. Then for each $p \geq 2$ there exists $c_p > 0$ such that for each $n \in \mathbb{N}$ and any solution ${}_n Y$ to (2.9) it holds that

$$\begin{aligned} \sup_{s \in [r, T)} E \left[|F(s, {}_n Y) - F(\underline{s}_n, {}_n Y) - \partial_x F(\underline{s}_n, {}_n Y)({}_n Y_s - {}_n Y_{\underline{s}_n})|^p \right] \\ \leq c_p |\mathbb{T}_n|^p (1 + E[\|{}_n Y^r\|_\infty^{(\eta \vee 2)p}]). \end{aligned}$$

Proof. Let ${}_n \Delta : [\underline{s}_n, s] \times \Omega \rightarrow \mathbb{R}^{1 \times m}$ be defined by ${}_n \Delta_u := \partial_x F(u, {}_n Y) - \partial_x F(\underline{s}_n, {}_n Y)$ for $s \in [r, T)$. Then the functional Itô formula in [9] yields

$$\begin{aligned} F(s, {}_n Y) - F(\underline{s}_n, {}_n Y) - \partial_x F(\underline{s}_n, {}_n Y)({}_n Y_s - {}_n Y_{\underline{s}_n}) \\ = \int_{\underline{s}_n}^s \partial_u F(u, {}_n Y) + {}_n \Delta_u (\underline{B}(u, {}_n Y) + B_H(u, {}_n Y)\dot{h}(u) + \overline{B}(u, {}_n Y)_n \dot{W}_u) du \\ + \int_{\underline{s}_n}^s {}_n \Delta_u \Sigma(u, {}_n Y) dW_u + \frac{1}{2} \int_{\underline{s}_n}^s \text{tr}(\partial_{xx} F(u, {}_n Y)(\Sigma \Sigma')(u, {}_n Y)) du \quad \text{a.s.} \end{aligned} \tag{5.8}$$

By Proposition 28, for $\bar{\eta} := \eta \vee 2$ there is a constant $c_{\bar{\eta}p} > 0$ such that (5.1) holds when c_p and p are replaced by $c_{\bar{\eta}p}$ and $\bar{\eta}p$, respectively. So,

$$E \left[\left| \int_{\underline{s}_n}^s \partial_u F(u, {}_n Y) du \right|^p \right] \leq c_{p,1} |\mathbb{T}_n|^p (1 + E[\|{}_n Y^r\|_\infty^{\bar{\eta}p}])^{\eta/\bar{\eta}}$$

with $c_{p,1} := 2^{2p} c_0^p (1 + c_{\bar{\eta}p})^{\eta/\bar{\eta}}$. Next, let $\lambda_0 \geq 0$ denote a Lipschitz constant for $\partial_x F$, then we obtain that

$$(E[|{}_n \Delta_u|^2])^{1/2} \leq \bar{c}_p |\mathbb{T}_n|^{p/2} (1 + E[\|{}_n Y^r\|_\infty^{\bar{\eta}p}])^{1/\bar{\eta}}$$

for all $u \in [\underline{s}_n, s]$ with $\bar{c}_p := 2^{3p/2} \lambda_0^p (1 + c_{\bar{\eta}p})^{1/\bar{\eta}}$. Hence, for the second term in (5.8) the Cauchy–Schwarz inequality gives

$$E \left[\left| \int_{\underline{s}_n}^s {}_n \Delta_u \underline{B}(u, {}_n Y) du \right|^p \right] \leq c_{p,2} |\mathbb{T}_n|^p (s - \underline{s}_n)^{p/2} (1 + E[\|{}_n Y^r\|_\infty^{\bar{\eta}p}])^{2/\bar{\eta}}$$

with $c_{p,2} := 2^{3p/2} c^p (1 + c_{\bar{\eta}p})^{1/\bar{\eta}} \bar{c}_p$. Similarly, it follows that

$$E \left[\left| \int_{\underline{s}_n}^s {}_n \Delta_u B_H(u, {}_n Y) dh(u) \right|^p \right] \leq c_{p,3} |\mathbb{T}_n|^p (1 + E[\|{}_n Y^r\|_\infty^{\bar{\eta}p}])^{2/\bar{\eta}}$$

for $c_{p,3} := 2^{3p/2} \|h\|_{H,r}^p c^p (1 + \underline{c}_{\bar{\eta}p})^{1/\bar{\eta}} \bar{c}_p$. Lemma 17 yields the constant $\hat{w}_{p,2}$ such that (3.8) is valid when q is replaced by 2. Then

$$\begin{aligned} & E \left[\left| \int_{\underline{s}_n}^s {}_n\Delta_u \bar{B}(u, {}_nY) d_n W_u \right|^p \right] \\ & \leq c^p (s - \underline{s}_n)^{p-1} \int_{\underline{s}_n}^s (E[|{}_n\Delta_u|^2])^{1/2} \left(E \left[\left(\int_{\underline{s}_n}^s |{}_n\dot{W}_v|^2 dv \right)^p \right] \right)^{1/2} du \\ & \leq c_{p,4} |\mathbb{T}_n|^p (1 + E[\|{}_nY^r\|_\infty^p])^{1/\bar{\eta}}, \end{aligned}$$

by the Cauchy–Schwarz inequality, where $c_{p,4} := 2^p \hat{w}_{p,2}^{1/2} c^p \bar{c}_p$. Next, as Σ cannot exceed the constant c , we directly get that

$$E \left[\left| \int_{\underline{s}_n}^s {}_n\Delta_u \Sigma(u, {}_nY) dW_u \right|^p \right] \leq c_{p,5} |\mathbb{T}_n|^p (1 + E[\|{}_nY^r\|_\infty^p])^{1/\bar{\eta}}$$

for $c_{p,5} := 2^{p/2} w_p c^p \bar{c}_p$, where w_p satisfies (3.7). For the last term in (5.8) we define $c_{p,6} := 2^{2p} c^{2p} c_0^p (1 + \underline{c}_{\bar{\eta}p})^{\eta/\bar{\eta}}$ and readily compute that

$$E \left[\left| \int_{\underline{s}_n}^s \text{tr}(\partial_{xx} F(u, {}_nY)(\Sigma \Sigma')(u, {}_nY)) du \right|^p \right] \leq c_{p,6} |\mathbb{T}_n|^p (1 + E[\|{}_nY^r\|_\infty^p])^{\eta/\bar{\eta}}.$$

Thus, by setting $c_p := 6^{p-1} (c_{p,1} + (T - r)^{p/2} c_{p,2} + c_{p,3} + c_{p,4} + c_{p,5} + 2^{-p} c_{p,6})$, we obtain the asserted estimate. \square

We come to the second remainder term arising in (5.7). As before, we derive an estimation that is necessary to apply Lemma 20.

Lemma 32. *Let (C.6) and (C.7) hold and $h \in H_r^1([0, T], \mathbb{R}^d)$. Then for each $p \geq 2$ there is $c_p > 0$ such that for any $n \in \mathbb{N}$ and every solution ${}_nY$ to (2.9) we have*

$$\sup_{s \in [r, T]} [|{}_nY_s - {}_nY_{\underline{s}_n} - \Phi_{h,n}(s, {}_nY, W) |^p] \leq c_p |\mathbb{T}_n|^p (1 + E[\|{}_nY^r\|_\infty^{2p}])^{1/2}.$$

Proof. We apply Proposition 28 to get a constant $\underline{c}_{2p} > 0$ such that (5.1) is satisfied when c_p and p are replaced by \underline{c}_{2p} and $2p$, respectively. Then

$$E \left[\left| \int_{\underline{s}_n}^s \underline{B}(u, {}_nY) du \right|^p \right] \leq c_{p,1} |\mathbb{T}_n|^p (1 + E[\|{}_nY^r\|_\infty^{2p}])^{1/2}$$

for given $s \in [r, T]$ and $c_{p,1} := 2^{2p} c^p (1 + \underline{c}_{2p})^{1/2}$. Let $\lambda \geq 0$ denote a Lipschitz constant for B_H, \bar{B} and Σ , then the Cauchy–Schwarz inequality allows us to estimate that

$$E \left[\left| \int_{\underline{s}_n}^s B_H(u, {}_nY) - B_H(\underline{s}_n, {}_nY) dh(u) \right|^p \right] \leq c_{p,2} |\mathbb{T}_n|^p (1 + E[\|{}_nY^r\|_\infty^{2p}])^{1/2}$$

with $c_{p,2} := 2^{2p} \|h\|_{H,r}^p \lambda^p (1 + \underline{c}_{2p})^{1/2}$. We recall the constant $\hat{w}_{p,2}$ constructed in Lemma 17 such that (3.8) holds when q is replaced by 2 and compute that

$$\begin{aligned} & E \left[\left| \int_{\underline{s}_n}^s \bar{B}(u, {}_n Y) - \bar{B}(\underline{s}_n, {}_n Y) d_n W_u \right|^p \right] \\ & \leq \lambda^p (s - \underline{s}_n)^{p/2} E \left[\left((s - \underline{s}_n)^{1/2} + \|{}_n Y^s - {}_n Y^{\underline{s}_n}\|_\infty \right)^p \left(\int_{\underline{s}_n}^s |\dot{W}_v|^2 dv \right)^{p/2} \right] \\ & \leq c_{p,3} |\mathbb{T}_n|^p (1 + E[\|{}_n Y^r\|_\infty^{2p}])^{1/2}, \end{aligned}$$

by the Cauchy–Schwarz inequality, where $c_{p,3} := 2^{5p/2} \hat{w}_{p,2}^{1/2} \lambda^p (1 + \underline{c}_{2p})^{1/2}$. Finally, let us also recall the constant w_p in (3.7), then

$$E \left[\left| \int_{\underline{s}_n}^s \Sigma(u, {}_n Y) - \Sigma(\underline{s}_n, {}_n Y) dW_u \right|^p \right] \leq c_{p,4} |\mathbb{T}_n|^p (1 + E[\|{}_n Y^r\|_\infty^{2p}])^{1/2}$$

for $c_{p,4} := 2^{2p} w_p \lambda^p (1 + \underline{c}_{2p})^{1/2}$. So, the definition $c_p := 4^{p-1} (c_{p,1} + \dots + c_{p,4})$ concludes the proof. \square

5.3. Convergence of the third remainder

As preparation, we infer an estimate from Doob’s L^2 -maximal inequality. To this end, let for the moment $\tilde{d} \in \mathbb{N}$ and \mathbb{T} be a partition of $[r, T]$ that is of the form $\mathbb{T} = \{t_0, \dots, t_k\}$ with $k \in \mathbb{N}$ and $t_0, \dots, t_k \in [r, T]$ such that $r = t_0 < \dots < t_k = T$.

Lemma 33. For every $l \in \{1, \dots, \tilde{d}\}$ let $({}_l U_i)_{i \in \{1, \dots, k\}}$ and $({}_l V_i)_{i \in \{1, \dots, k\}}$ be two sequences of $\mathbb{R}^{1 \times \tilde{d}}$ -valued and $\mathbb{R}^{\tilde{d}}$ -valued random vectors, respectively, such that ${}_l U_i$ is $\mathcal{F}_{t_{i-1}}$ -measurable, ${}_l V_i$ is \mathcal{F}_{t_i} -measurable,

$$E[|{}_l U_i|^4 + |{}_l V_i|^4] < \infty \quad \text{and} \quad E[{}_l V_i | \mathcal{F}_{t_{i-1}}] = 0 \quad \text{a.s.}$$

for all $i \in \{1, \dots, k\}$. Then

$$E \left[\max_{j \in \{i_0, \dots, k\}} \left| \sum_{i=1}^{j-i_0} \sum_{l=1}^{\tilde{d}} {}_l U_i {}_l V_i \right|^2 \right] \leq 4 \sum_{i=1}^{k-i_0} \sum_{l_1, l_2=1}^{\tilde{d}} E[{}_l U_i {}_l V_i {}_l V'_i {}_l U'_i]$$

for every $i_0 \in \{0, \dots, k-1\}$.

Proof. We set $Y_i := \sum_{l=1}^{\tilde{d}} {}_l U_i {}_l V_i$ for each $i \in \{1, \dots, k-i_0\}$, then the sequence $(S_j)_{j \in \{i_0, \dots, k\}}$ of random variables given by $S_j := \sum_{i=1}^{j-i_0} Y_i$ is a square-integrable martingale with respect to $(\mathcal{F}_{t_{j-i_0}})_{j \in \{i_0, \dots, k\}}$. So,

$$E \left[\max_{j \in \{i_0, \dots, k\}} \left| \sum_{i=1}^{j-i_0} \sum_{l=1}^{\tilde{d}} {}_l U_i {}_l V_i \right|^2 \right] = E \left[\max_{j \in \{i_0, \dots, k\}} S_j^2 \right] \leq 4E[S_k^2],$$

by Doob’s L^2 -maximal inequality. Moreover, let $i, j \in \{1, \dots, k-i_0\}$ be such that $i \leq j$, then

$$E[Y_i Y_j] = \mathbb{1}_{\{i\}}(j) \sum_{l_1, l_2=1}^{\tilde{d}} E[{}_l U_i E[{}_l V_i {}_l V'_j | \mathcal{F}_{t_{j-1}}] {}_l U'_j].$$

In particular, Y_i and Y_j are uncorrelated for $i < j$. By Bienaymé’s identity, $E[S_k^2] = \sum_{i=1}^{k-i_0} E[Y_i^2]$, which yields the claim. \square

Proposition 34. Let (C.6) and (C.7) hold and $h \in H_r^1([0, T], \mathbb{R}^d)$. Then there is $c_2 > 0$ such that for each $n \in \mathbb{N}$ and any solution ${}_n Y$ to (2.9),

$$E \left[\max_{j \in \{0, \dots, k_n\}} \left| \int_r^{t_{j,n}} \partial_x \bar{B}(\underline{s}_n, {}_n Y) \Phi_{h,n}(s, {}_n Y, W)_n \dot{W}_s - R(\underline{s}_n, {}_n Y) \gamma_n(s) ds \right|^2 \right] \leq c_2 |\mathbb{T}_n| (1 + E[\|{}_n Y^r\|_\infty^2]).$$

Proof. First, let us recall the definition of R in (2.11) to write the k th coordinate of $\partial_x \bar{B}(\underline{s}_n, {}_n Y) \Phi_{h,n}(s, {}_n Y, W)_n \dot{W}_s - R(\underline{s}_n, {}_n Y) \gamma_n(s)$ in the form

$$\sum_{l=1}^d \partial_x \bar{B}_{k,l}(\underline{s}_n, {}_n Y) (\Phi_{h,n}(s, {}_n Y, W)_n \dot{W}_s^{(l)} - ((1/2)\bar{B} + \Sigma)(\underline{s}_n, {}_n Y) \gamma_n(s) e_l)$$

for all $k \in \{1, \dots, m\}$ and $s \in [r, T]$, where the l th coordinate of any \mathbb{R}^d -valued stochastic process X is denoted by $X^{(l)}$ for every $l \in \{1, \dots, d\}$. Moreover, we decompose that

$$\begin{aligned} & \Phi_{h,n}(s, {}_n Y, W)_n \dot{W}_s^{(l)} - ((1/2)\bar{B} + \Sigma)(\underline{s}_n, {}_n Y) \gamma_n(s) e_l \\ &= B_H(\underline{s}_n, {}_n Y)(h(s_n) - h(\underline{s}_n))_n \dot{W}_s^{(l)} + \bar{B}(\underline{s}_n, {}_n Y)({}_n W_{s_n} - {}_n W_{\underline{s}_n})_n \dot{W}_s^{(l)} \\ & \quad + \Sigma(\underline{s}_n, {}_n Y)(\Delta W_{s_n} \dot{W}_s^{(l)} - \gamma_n(s) e_l) + B_H(\underline{s}_n, {}_n Y)(h(s) - h(s_n))_n \dot{W}_s^{(l)} \\ & \quad + \bar{B}(\underline{s}_n, {}_n Y)(({}_n W_s - {}_n W_{s_n})_n \dot{W}_s^{(l)} - (1/2)\gamma_n(s) e_l) \\ & \quad + \Sigma(\underline{s}_n, {}_n Y)(W_s - W_{s_n})_n \dot{W}_s^{(l)} \end{aligned} \tag{5.9}$$

for any $l \in \{1, \dots, d\}$. We begin with the first term in this decomposition and use Lemma 15 to obtain that

$$\begin{aligned} & \int_r^{t_{j,n}} (\partial_x \bar{B}_{k,l} B_H)(\underline{s}_n, {}_n Y)(h(s_n) - h(\underline{s}_n))_n \dot{W}_s^{(l)} ds \\ &= \int_r^{t_{j-1,n}} (\partial_x \bar{B}_{k,l} B_H)(s_n, {}_n Y)(h(\bar{s}_n) - h(s_n)) dW_s^{(l)} \quad \text{a.s.} \end{aligned}$$

for each $j \in \{1, \dots, k_n\}$, $k \in \{1, \dots, m\}$ and $l \in \{1, \dots, d\}$. Proposition 28 provides $\underline{c}_2 > 0$ such that (5.1) is satisfied for $p = 2$ when the appearing constant c_p is replaced by \underline{c}_2 . Hence, condition (C.6) gives

$$\begin{aligned} & E \left[\max_{j \in \{0, \dots, k_n\}} \sum_{k=1}^m \left| \int_r^{t_{j,n}} \sum_{l=1}^d (\partial_x \bar{B}_{k,l} B_H)(\underline{s}_n, {}_n Y)(h(s_n) - h(\underline{s}_n))_n \dot{W}_s^{(l)} ds \right|^2 \right] \\ & \leq 2w_2 c^4 \int_r^T (1 + E[\|{}_n Y^{s_n}\|_\infty^{2\kappa}]) |h(\bar{s}_n) - h(s_n)|^2 ds \\ & \leq c_{2,1} |\mathbb{T}_n| (1 + E[\|{}_n Y^r\|_\infty^2]), \end{aligned}$$

where $c_{2,1} := 2^2 w_2 (T - r) \|h\|_{H,r}^2 c^4 (1 + \underline{c}_2)$ and w_2 satisfies (3.7) for $p = 2$. Similarly, another application of Lemma 15 gives us that

$$\begin{aligned} & \int_r^{t_{j,n}} (\partial_x \bar{B}_{k,l} \bar{B})(\underline{s}_n, {}_n Y) ({}_n W_{s_n} - {}_n W_{\underline{s}_n})_n \dot{W}_s^{(l)} ds \\ &= \int_r^{t_{j-1,n}} (\partial_x \bar{B}_{k,l} \bar{B})(s_n, {}_n Y) \Delta W_{s_n} dW_s^{(l)} \quad \text{a.s.} \end{aligned}$$

for all $j \in \{1, \dots, k_n\}$, $k \in \{1, \dots, m\}$ and $l \in \{1, \dots, d\}$. Thus, with the constant $c_{2,2} := w_2(T - r)dc^4$ we can estimate that

$$E \left[\max_{j \in \{0, \dots, k_n\}} \sum_{k=1}^m \left| \int_r^{t_{j,n}} \sum_{l=1}^d (\partial_x \bar{B}_{k,l} \bar{B})(\underline{s}_n, nY)_n W_{s_n} - {}_n W_{\underline{s}_n} \dot{W}_s^{(l)} ds \right|^2 \right] \leq w_2 c^4 \int_r^T E[|\Delta W_{s_n}|^2] ds \leq c_{2,2} |\mathbb{T}_n|.$$

Let us move on to the third expression in (5.9). First we define an \mathbb{R}^d -valued $\mathcal{F}_{t_{i,n}}$ -measurable random vector by

$${}_{l,n} V_i := \Delta W_{t_{i,n}} \Delta W_{t_{i,n}}^{(l)} - \Delta t_{i,n} e_l \tag{5.10}$$

for every $i \in \{1, \dots, k_n\}$ and $l \in \{1, \dots, d\}$, then ${}_{l,n} V_i$ is independent of $\mathcal{F}_{t_{i-1,n}}$ and satisfies $E[|{}_{l,n} V_i|^4] < \infty$ and $E[{}_{l,n} V_i] = 0$. Moreover, a case distinction shows that

$$E[{}_{l_1,n} V_i {}_{l_2,n} V_i'] = \mathbb{1}_{\{l_2\}}(l_1) (\Delta t_{i,n})^2 (\mathbb{I}_d + \mathbb{I}_{l_2, l_1}) \tag{5.11}$$

for each $i \in \{1, \dots, k_n\}$ and $l_1, l_2 \in \{1, \dots, d\}$, where $\mathbb{I}_{l_2, l_1} \in \mathbb{R}^{d \times d}$ denotes the matrix whose (l_2, l_1) -entry is 1 and whose all other entries are zero. We compute that

$$\begin{aligned} & \int_r^{t_{j,n}} (\partial_x \bar{B}_{k,l} \Sigma)(\underline{s}_n, nY) (\Delta W_{s_n} \dot{W}_s^{(l)} - \gamma_n(s) e_l) ds \\ &= \sum_{i=1}^{j-1} (\partial_x \bar{B}_{k,l} \Sigma)(t_{i-1,n}, nY) {}_{l,n} V_i \end{aligned}$$

for all $j \in \{1, \dots, k_n\}$, $k \in \{1, \dots, m\}$ and $l \in \{1, \dots, d\}$, since $\gamma_n(s) = 0$ for each $s \in [r, t_{1,n})$. Consequently, Lemma 33 and the representation (5.11) imply that

$$E \left[\max_{j \in \{0, \dots, k_n\}} \sum_{k=1}^m \left| \int_r^{t_{j,n}} \sum_{l=1}^d (\partial_x \bar{B}_{k,l} \Sigma)(\underline{s}_n, nY) (\Delta W_{s_n} \dot{W}_s^{(l)} - \gamma_n(s) e_l) ds \right|^2 \right] \leq 2^3 \sum_{i=1}^{k_n-1} (\Delta t_{i,n})^2 \sum_{k=1}^m \sum_{l=1}^d E[|(\partial_x \bar{B}_{k,l} \Sigma)(t_{i-1,n}, nY)|^2] \leq c_{2,3} |\mathbb{T}_n|$$

for $c_{2,3} := 2^3(T - r)c^4$, since we can use that $\bar{x}^t \mathbb{I}_{l_2, l_1} \bar{y} \leq (1/2)(\bar{x}_{l_2}^2 + \bar{y}_{l_1}^2)$ for all $l_1, l_2 \in \{1, \dots, d\}$ and $\bar{x}, \bar{y} \in \mathbb{R}^d$, by Young’s inequality. To deal with the fourth term in (5.9), let us note that

$$\begin{aligned} & \int_r^{t_{j,n}} (\partial_x \bar{B}_{k,l} B_H)(\underline{s}_n, nY) (h(s) - h(s_n))_n \dot{W}_s^{(l)} ds \\ &= \sum_{i=1}^{j-1} (\partial_x \bar{B}_{k,l} B_H)(t_{i-1,n}, nY) \frac{\Delta W_{t_{i,n}}^{(l)}}{\Delta t_{i+1,n}} \int_{t_{i,n}}^{t_{i+1,n}} h(s) - h(t_{i,n}) ds \\ &= \int_r^{t_{j,n}} (\partial_x \bar{B}_{k,l} B_H)(\underline{s}_n, nY) \Delta W_{s_n}^{(l)} \frac{(\bar{s}_n - s)}{\Delta \bar{s}_n} dh(s) \end{aligned} \tag{5.12}$$

for each $j \in \{1, \dots, k_n\}$, $k \in \{1, \dots, m\}$ and $l \in \{1, \dots, d\}$, as integration by parts yields that $\int_{t_{i,n}}^{t_{i+1,n}} h(s) - h(t_{i,n}) ds = \int_{t_{i,n}}^{t_{i+1,n}} t_{i+1,n} - s dh(s)$ for all $i \in \{0, \dots, k_n - 1\}$. So,

$$\begin{aligned} E \left[\max_{j \in \{0, \dots, k_n\}} \sum_{k=1}^m \left| \int_r^{t_{j,n}} \sum_{l=1}^d (\partial_x \bar{B}_{k,l} B_H)(\underline{s}_n, nY)(h(s) - h(s_n))_n \dot{W}_s^{(l)} ds \right|^2 \right] \\ \leq \|h\|_{H,r}^2 \sum_{k=1}^m \int_r^T E \left[\left| \sum_{l=1}^d (\partial_x \bar{B}_{k,l} B_H)(\underline{s}_n, nY) \Delta W_{s_n}^{(l)} \right|^2 \right] ds \\ \leq c_{2,4} |\mathbb{T}_n| (1 + E[\|nY^r\|_\infty^2]) \end{aligned}$$

with $c_{2,4} := 2^2(T - r)\|h\|_{H,r}^2 c^4(1 + c_2)$, by the Cauchy–Schwarz inequality and the fact that $\Delta W_{s_n}^{(1)}, \dots, \Delta W_{s_n}^{(d)}$ are pairwise independent and independent from \mathcal{F}_{s_n} for all $s \in [r, T]$.

To handle the fifth expression in (5.9), we define ${}_{l,n}U_s := ({}_nW_s - {}_nW_{s_n})_n \dot{W}_s^{(l)} - (1/2)\gamma_n(s)e_l$ for all $s \in [r, T]$ and note that

$$\int_r^{t_{j,n}} (\partial_x \bar{B}_{k,l} \bar{B})(\underline{s}_n, nY) {}_{l,n}U_s ds = \frac{1}{2} \sum_{i=1}^{j-1} (\partial_x \bar{B}_{k,l} \bar{B})(t_{i-1,n}, nY) {}_{l,n}V_i$$

for every $j \in \{1, \dots, k_n\}$, $k \in \{1, \dots, m\}$ and $l \in \{1, \dots, d\}$, where ${}_{l,n}V_i$ is given by (5.10) and we have used the fact that $\int_{t_{i,n}}^{t_{i+1,n}} s - t_{i,n} ds = (1/2)(\Delta t_{i+1,n})^2$ for each $i \in \{0, \dots, k_n - 1\}$. Consequently,

$$\begin{aligned} E \left[\max_{j \in \{0, \dots, k_n\}} \sum_{k=1}^m \left| \int_r^{t_{j,n}} \sum_{l=1}^d (\partial_x \bar{B}_{k,l} \bar{B})(\underline{s}_n, nY) {}_{l,n}U_s ds \right|^2 \right] \\ \leq 2 \sum_{i=1}^{k_n-1} (\Delta t_{i,n})^2 \sum_{k=1}^m \sum_{l=1}^d E[|(\partial_x \bar{B}_{k,l} \bar{B})(t_{i-1,n}, nY)|^2] \leq c_{2,5} |\mathbb{T}_n| \end{aligned}$$

for $c_{2,5} := 2(T - r)c^4$. We turn to the last term in (5.9) and proceed just as in (5.12) to get that

$$\begin{aligned} \int_r^{t_{j,n}} (\partial_x \bar{B}_{k,l} \Sigma)(\underline{s}_n, nY) (W_s - W_{s_n})_n \dot{W}_s^{(l)} ds \\ = \int_r^{t_{j,n}} (\partial_x \bar{B}_{k,l} \Sigma)(\underline{s}_n, nY) \Delta W_{s_n}^{(l)} \frac{(s_n - s)}{\Delta s_n} dW_s \quad \text{a.s.} \end{aligned}$$

for each $j \in \{1, \dots, k_n\}$, $k \in \{1, \dots, m\}$ and $l \in \{1, \dots, d\}$, as Itô’s formula gives $\int_{t_{i,n}}^{t_{i+1,n}} W_s - W_{t_{i,n}} ds = \int_{t_{i,n}}^{t_{i+1,n}} t_{i+1,n} - s dW_s$ a.s. for all $i \in \{0, \dots, k_n - 1\}$. Therefore,

$$\begin{aligned} E \left[\max_{j \in \{0, \dots, k_n\}} \sum_{k=1}^m \left| \int_r^{t_{j,n}} \sum_{l=1}^d (\partial_x \bar{B}_{k,l} \Sigma)(\underline{s}_n, nY) (W_s - W_{s_n})_n \dot{W}_s^{(l)} ds \right|^2 \right] \\ \leq w_2 \sum_{k=1}^m \int_r^T E \left[\left| \sum_{l=1}^d (\partial_x B_{k,l} \Sigma)(\underline{s}_n, nY) \Delta W_{s_n}^{(l)} \right|^2 \right] ds \leq c_{2,6} |\mathbb{T}_n| \end{aligned}$$

with $c_{2,6} := w_2(T - r)c^4$. As before, we used that $\Delta W_{s_n}^{(1)}, \dots, \Delta W_{s_n}^{(d)}$ are pairwise independent and independent of \mathcal{F}_{s_n} for every $s \in [r, T]$. Therefore, by setting $c_2 := 6(c_{2,1} + \dots + c_{2,6})$, the assertion follows. \square

5.4. Proofs of Theorems 7 and 1

At first, we consider a sufficient condition for a Doléans-Dade exponential to be a true martingale and let for the moment \mathbb{T} be a partition of $[r, T]$ of the form $\mathbb{T} = \{t_0, \dots, t_k\}$ with $k \in \mathbb{N}$ and $t_0, \dots, t_k \in [r, T]$ satisfying $r = t_0 < \dots < t_k = T$.

Lemma 35. *Let $f : [r, T] \rightarrow \mathbb{R}^{m \times d}$ be measurable such that $\int_r^T |f(s)|^2 ds < \infty$ and $(Y_i)_{i \in \{0, \dots, k-1\}}$ be an $(\mathcal{F}_{t_i})_{i \in \{0, \dots, k-1\}}$ -adapted sequence of $\mathbb{R}^{m \times d}$ -valued random matrices. Define the process $X : [r, T] \times \Omega \rightarrow \mathbb{R}^{1 \times d}$ coordinatewise via*

$$X_t^{(l)} := \sum_{i=0}^{k-1} \sum_{j=1}^m f_{j,l}(t) Y_i^{(j,l)} \mathbb{1}_{[t_i, t_{i+1})}(t),$$

then the continuous local martingale $Z \in \mathcal{C}([0, T], \mathbb{R})$ given by $Z^r = 1$ and

$$Z_t = \exp\left(\int_r^t X_s dW_s - \frac{1}{2} \int_r^t |X_s|^2 ds\right)$$

for all $t \in [r, T]$ a.s. is a martingale.

Proof. Since Z is a positive supermartingale, it suffices to show that $E[Z_T] = 1$. This in turn follows inductively if we can verify that $E[Z_{t_{i+1}} | \mathcal{F}_{t_i}] = Z_{t_i}$ a.s. for each $i \in \{0, \dots, k-1\}$.

In this regard, note that $\int_{t_i}^{t_{i+1}} X_s dW_s = \sum_{l=1}^d \sum_{j=1}^m Y_i^{(j,l)} \int_{t_i}^{t_{i+1}} f_{j,l}(s) dW_s^{(l)}$ a.s. Because $\int_{t_i}^{t_{i+1}} f_{j,1}(s) dW_s^{(1)}, \dots, \int_{t_i}^{t_{i+1}} f_{j,d}(s) dW_s^{(d)}$ are independent of \mathcal{F}_{t_i} for any $j \in \{1, \dots, m\}$, we have

$$E\left[\exp\left(\int_{t_i}^{t_{i+1}} X_s dW_s - \frac{1}{2} \int_{t_i}^{t_{i+1}} |X_s|^2 ds\right) \middle| \mathcal{F}_{t_i}\right] = \psi(Y_i) \quad \text{a.s.,}$$

where the Borel measurable function $\psi : \mathbb{R}^{m \times d} \rightarrow \mathbb{R}$ is given by

$$\psi(A) := E\left[\prod_{l=1}^d \exp\left(\int_{t_i}^{t_{i+1}} \sum_{j=1}^m A_{j,l} f_{j,l}(s) dW_s^{(l)} - \frac{1}{2} \int_{t_i}^{t_{i+1}} \left|\sum_{j=1}^m A_{j,l} f_{j,l}(s)\right|^2 ds\right)\right].$$

Moreover, as $\int_{t_i}^{t_{i+1}} \sum_{j=1}^m A_{j,l} f_{j,l}(s) dW_s^{(l)}$ is normally distributed with zero mean and variance given by $\int_{t_i}^{t_{i+1}} \left|\sum_{j=1}^m A_{j,l} f_{j,l}(s)\right|^2 ds$ for any $A \in \mathbb{R}^{m \times d}$ and $l \in \{1, \dots, d\}$, independence of the coordinates of W entails that $\psi = 1$, as desired. \square

Proof of Theorem 7. (i) If in condition (C.8) we have $\bar{b}_0 = 0$, then existence and uniqueness can be inferred from Proposition 3. Otherwise, we may let $\bar{b}_0 = 1$ and, by using Lemma 35, define a martingale ${}_n\bar{Z} \in \mathcal{C}([0, T], \mathbb{R})$ via ${}_n\bar{Z}^r = 1$ and

$${}_n\bar{Z}_t = \exp\left(-\int_r^t \bar{b}(s)_n \dot{W}'_s dW_s - \frac{1}{2} \int_r^t |\bar{b}(s)_n \dot{W}_s|^2 ds\right)$$

for all $t \in [r, T]$ a.s. Due to Girsanov’s theorem, the process ${}_n\bar{W} \in \mathcal{C}([0, T], \mathbb{R}^d)$ given by ${}_n\bar{W}_t := W_t + \int_r^{r \vee t} \bar{b}(s) d_n W_s$ is a d -dimensional $(\mathcal{F}_t)_{t \in [0, T]}$ -Brownian motion under the probability measure \bar{P}_n on (Ω, \mathcal{F}) defined by $\bar{P}_n(A) := E[\mathbb{1}_A {}_n\bar{Z}_T]$.

As this yields an equivalent probability measure, a process $Y \in \mathcal{C}([0, T], \mathbb{R}^m)$ solves (2.9) under P if and only if it is a solution to the SDE

$$dY_t = (\underline{B}(t, Y) + B_H(t, Y)\dot{h}(t)) dt + \Sigma(t, Y) d_n \bar{W}_t \quad \text{for } t \in [r, T]$$

under \bar{P}_n . For this reason, existence and uniqueness follow from Proposition 5 when $b = \underline{B} + B_H \dot{h}$ and $\sigma = \Sigma$. Further, independently of this case distinction, Propositions 28 and 12 imply the second claim.

(ii) This assertion is an immediate application of Proposition 5 in the case that $b = \underline{B} + R + B_H \dot{h}$ and $\sigma = \bar{B} + \Sigma$.

(iii) By Propositions 30, 28 and 22 and Lemmas 19 and 21, to establish (2.13), it suffices to show that there is $c_2 > 0$ such that

$$E \left[\max_{j \in \{0, \dots, k_n\}} \left| \int_r^{t_{j,n}} (\bar{B}(s, nY) - \bar{B}(\underline{s}_n, nY))_n \dot{W}_s - R(\underline{s}_n, nY) \gamma_n(s) ds \right|^2 \right] \leq c_2 |\mathbb{T}_n|$$

for each $n \in \mathbb{N}$. As $\partial_x \bar{B}$ is bounded, the existence of such a constant c_2 follows immediately from the decomposition (5.7), a combination of Proposition 31 and Lemma 32 with Lemma 20 and an application of Proposition 34. Moreover, since $\sup_{n \in \mathbb{N}} E[\|nY\|_{\beta,r}^p] + E[\|Y\|_{\beta,r}^p] < \infty$ for all $\beta \in [0, 1/2)$ and $p \geq 1$, the second assertion can now be inferred from Lemma 13. \square

To prove Theorem 1 we require the following basic result on the support of image probability measures.

Lemma 36. *Let $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$ be a probability space, (S, ν) be a metric space, $D \subset S$ and $Y : \tilde{\Omega} \rightarrow S$ be measurable such that $\tilde{P} \circ Y^{-1}$ is inner regular.*

- (i) *Let $(Y_n)_{n \in \mathbb{N}}$ be a sequence of S -valued measurable maps on $\tilde{\Omega}$ that converges in probability to Y . If $Y_n \in D$ a.s. for all $n \in \mathbb{N}$, then $\text{supp}(\tilde{P} \circ Y^{-1}) \subset \bar{D}$.*
- (ii) *Suppose that for each $y \in D$ there exists a sequence $(\tilde{P}_{y,n})_{n \in \mathbb{N}}$ of probability measures on $(\tilde{\Omega}, \tilde{\mathcal{F}})$ such that $\tilde{P}_{y,n} \ll \tilde{P}$ for all $n \in \mathbb{N}$ and*

$$\inf_{n \in \mathbb{N}} \tilde{P}_{y,n}(\nu(Y, y) \geq \varepsilon) < 1 \tag{5.13}$$

for each $\varepsilon > 0$. Then $\bar{D} \subset \text{supp}(\tilde{P} \circ Y^{-1})$.

Proof. (i) Let $y \in \text{supp}(\tilde{P} \circ Y^{-1})$ and $k \in \mathbb{N}$, then there exists $n_k \in \mathbb{N}$ such that $\tilde{P}(\nu(Y_n, Y) > 1/(2k)) < \tilde{P}(\nu(Y, y) < 1/(2k))$ for all $n \in \mathbb{N}$ with $n \geq n_k$. By the triangle inequality,

$$\tilde{P}(\nu(Y_n, y) \geq 1/k) \leq \tilde{P}(\nu(Y_n, Y) > 1/(2k)) + \tilde{P}(\nu(Y, y) \geq 1/(2k)) < 1$$

for any such $n \in \mathbb{N}$. So, there is $\tilde{\omega}_k \in \tilde{\Omega}$ such that $y_k := Y_{n_k}(\tilde{\omega}_k) \in D$ and $\nu(y_k, y) < 1/k$. As $k \in \mathbb{N}$ has been arbitrarily chosen, the resulting sequence $(y_k)_{k \in \mathbb{N}}$ converges to y , which gives the claim.

(ii) By way of contradiction, assume that there are $y \in \bar{D}$ and $\varepsilon > 0$ such that $\tilde{P}(\nu(Y, y) \geq \varepsilon) = 1$. Let $(y_n)_{n \in \mathbb{N}}$ be a sequence in D that converges to y and choose $n_\varepsilon \in \mathbb{N}$ such that $\nu(y_{n_\varepsilon}, y) < \varepsilon/2$. Then from

$$\tilde{P}(\nu(Y, y) \geq \varepsilon) \leq \tilde{P}(\nu(Y, y_{n_\varepsilon}) > \varepsilon/2) + \tilde{P}(\nu(y_{n_\varepsilon}, y) \geq \varepsilon/2)$$

and $\tilde{P}_{y_{n_\varepsilon}, n} \ll \tilde{P}$ it follows that $\tilde{P}_{y_{n_\varepsilon}, n}(\nu(Y, y_{n_\varepsilon}) \geq \varepsilon/2) = 1$ for each $n \in \mathbb{N}$. This, however, is a contradiction to (5.13). \square

Proof of Theorem 1. (i) Pathwise uniqueness follows from Lemma 27 and the other two assertions are direct consequences of Proposition 5.

(ii) For $h \in H_r^1([0, T], \mathbb{R}^d)$ we set $F_h := b - (1/2)\rho + \sigma \dot{h}$ and first check that F_h satisfies conditions (C.1) and (C.2). Since σ and $\partial_x \sigma$ are bounded, there is $c_0 \geq 0$ such that $|\sigma| \vee |\rho| \leq c_0$. Then

$$|F_h(t, x)| \leq c_1(1 + |\dot{h}(t)|)(1 + \|x\|_\infty)$$

for all $t \in [r, T]$ and $x \in C([0, T], \mathbb{R}^m)$ with $c_1 := 3 \max\{c, c_0\}$. Moreover, since σ and $\partial_x \sigma$ are d_∞ -Lipschitz continuous, so is the map ρ . Let $\lambda_0 \geq 0$ be a Lipschitz constant for ρ , then

$$|F_h(t, x) - F_h(t, y)| \leq \lambda_1(1 + |\dot{h}(t)|)\|x - y\|_\infty$$

for all $t \in [r, T]$ and $x, y \in C([0, T], \mathbb{R}^m)$ with $\lambda_1 := 2 \max\{\lambda, \lambda_0\}$. Hence, an application of Proposition 3 yields the first assertion.

Regarding the second claim, let us also choose $g \in H_r^1([0, T], \mathbb{R}^d)$ and define $c_2 := 2^2 \max\{c_0^2, \lambda_1^2\} \max\{1, T - r\}$. Then the above estimation shows that

$$\|x_g^t - x_h^t\|_{H,r}^2 \leq c_2 \int_r^t |\dot{g}(s) - \dot{h}(s)|^2 + (1 + |\dot{h}(s)|^2) \|x_g^s - x_h^s\|_{H,r}^2 ds$$

for given $t \in [r, T]$. By Gronwall’s inequality, $\|x_g - x_h\|_{H,r}^2 \leq c_3 e^{c_3 \|h\|_{H,r}^2} \|g - h\|_{H,r}^2$ with $c_3 := c_2 \exp((T - r)c_2)$, and the verification is complete.

(iii) Let N_α be the P -null set of all $\omega \in \Omega$ such that $X(\omega) \notin C_r^\alpha([0, T], \mathbb{R}^m)$, then $(N_\alpha^c, \mathcal{F} \cap N_\alpha^c, P|_{\mathcal{F} \cap N_\alpha^c})$ is a probability space and the probability measure

$$\mathcal{B}(C_r^\alpha([0, T], \mathbb{R}^m)) \rightarrow [0, 1], \quad B \mapsto P(\{X \in B\} \cap N_\alpha^c) \tag{5.14}$$

is inner regular and its support agrees with the support of $P \circ X^{-1}$ in $C_r^\alpha([0, T], \mathbb{R}^m)$. We note that the inner regularity can be inferred by using that X belongs a.s. to the separable closed linear subspace of all $x \in C_r^\alpha([0, T], \mathbb{R}^m)$ such that

$$\lim_{\delta \downarrow 0} \sup_{s,t \in [r,T]: 0 < |s-t| \leq \delta} \frac{|x(s) - x(t)|}{|s - t|^\alpha} = 0.$$

Since (2.6) follows from Theorem 7 by the choice $\underline{B} = b - (1/2)\rho$, $B_H = 0$, $\bar{B} = \sigma$ and $\Sigma = 0$, Lemma 36 entails that the support of (5.14) is included in the closure of $\{x_h \mid h \in H_r^1([0, T], \mathbb{R}^d)\}$ with respect to $\|\cdot\|_{\alpha,r}$.

Next, let $h \in H_r^1([0, T], \mathbb{R}^d)$ and for any $n \in \mathbb{N}$ we note that the non-anticipative product measurable map $[r, T] \times C([0, T], \mathbb{R}^d) \rightarrow \mathbb{R}^d$, $(t, x) \mapsto \dot{L}_n(x)(t)$ satisfies $|\dot{L}_n(x)(t)| \leq 2c_{\mathbb{T}} |\mathbb{T}_n|^{-1} \|x\|_\infty$ for all $t \in [r, T] \setminus \mathbb{T}_n$ and $x \in C([0, T], \mathbb{R}^d)$ and the map $C([0, T], \mathbb{R}^d) \rightarrow \mathbb{R}^d$, $x \mapsto \dot{L}_n(x)(t)$ is linear for each $t \in [r, T] \setminus \mathbb{T}_n$.

From these facts it follows that for every $x \in C([0, T], \mathbb{R}^d)$ there is a unique mild solution $y_{h,n,x} \in C([0, T], \mathbb{R}^d)$ to the following ordinary integral equation with running value condition:

$$y_{h,n,x}(t) = x(t) - \int_r^{r \vee t} \dot{h}(s) - \dot{L}_n(y_{h,n,x})(s) ds \quad \text{for } t \in [0, T].$$

Moreover, the map $C([0, T], \mathbb{R}^d) \rightarrow C([0, T], \mathbb{R}^d)$, $x \mapsto y_{h,n,x}$ is Lipschitz continuous on bounded sets and in particular, Borel measurable. These considerations allow us to define a process ${}_{h,n}W \in \mathcal{C}([0, T], \mathbb{R}^d)$ by ${}_{h,n}W_t := y_{h,n,W}(t)$.

Given this constructed process, it follows from Lemma 35 that we obtain a martingale ${}_{h,n}Z \in \mathcal{C}([0, T], \mathbb{R})$ by requiring that ${}_{h,n}Z^r = 1$ and

$${}_{h,n}Z_t = \exp\left(\int_r^t \dot{h}(s)' - \dot{L}_n({}_{h,n}W)(s)' dW_s - \frac{1}{2} \int_r^t |\dot{h}(s) - \dot{L}_n({}_{h,n}W)(s)|^2 ds\right)$$

for every $t \in [r, T]$ a.s. and we may define a probability measure $P_{h,n}$ on (Ω, \mathcal{F}) that is equivalent to P by $P_{h,n}(A) := E[\mathbb{1}_{A_{h,n}} Z_T]$. According to the second part of Lemma 36, if

$$\inf_{n \in \mathbb{N}} P_{h,n}(\{\|X - x_h\|_{\alpha,r} \geq \varepsilon\} \cap N_\alpha^c) < 1 \quad \text{for each } \varepsilon > 0, \quad (5.15)$$

then the closure of $\{x_g \mid g \in H_r^1([0, T], \mathbb{R}^d)\}$ with respect to $\|\cdot\|_{\alpha,r}$ is included in the support of (5.14). Now Girsanov's theorem implies that for each $n \in \mathbb{N}$ the process ${}_{h,n}W$ is a d -dimensional $(\mathcal{F}_t)_{t \in [0, T]}$ -Brownian motion under $P_{h,n}$ and X is a strong solution to (2.8) under $P_{h,n}$.

Hence, by using uniqueness in law, an application of Theorem 7 in the case that $\underline{B} = b$, $B_H = \sigma$, $\overline{B} = -\sigma$ and $\Sigma = \sigma$ gives (2.7). As this readily implies (5.15), the proof is complete. \square

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- [1] S. Aida, Support theorem for diffusion processes on Hilbert spaces, *Publ. Res. Inst. Math. Sci.* 26 (1990) 947–965.
- [2] S. Aida, S. Kusuoka, D. Stroock, On the support of Wiener functionals, in: *Asymptotic Problems in Probability Theory: Wiener Functionals and Asymptotics*. Proceedings of the 26th Taniguchi International Symposium, Sanda and Kyoto, Japan, August 31 - September 5, 1990, Longman Scientific & Technical, Wiley, Harlow, Essex, New York, 1990, pp. 3–34.
- [3] V. Bally, A. Millet, M. Sanz-Solé, Approximation and support theorem in Hölder norm for parabolic stochastic partial differential equations, *Ann. Probab.* 23 (1995) 178–222.
- [4] G. Ben Arous, M. Gradinaru, M. Ledoux, Hölder Norms and the support theorem for diffusions, *Ann. IHP - Probab. Stat.* 30 (1994) 415–436.
- [5] R. Cont, Functional Ito calculus and functional Kolmogorov equations, in: F. Utzet, J. Vives (Eds.), *Stochastic Integration by Parts and Functional Ito Calculus* (Barcelona Summer School in Stochastic Analysis, 2012), in: *Advanced Courses in Mathematics*, Birkhauser Basel, 2012, pp. 115–208.
- [6] R. Cont, P. Das, Quadratic variation and quadratic roughness, working paper, 2018.
- [7] R. Cont, D. Fournié, A functional extension of the Itô formula, *Comptes Rendus Mathématiques de l'Académie des Sciences Paris* 348 (2009) 57–61.
- [8] R. Cont, D.-A. Fournié, Change of variable formulas for non-anticipative functionals on path space, *J. Funct. Anal.* 259 (2010) 1043–1072.
- [9] R. Cont, D.-A. Fournié, Functional Itô calculus and stochastic integral representation of martingales, *Ann. Probab.* 41 (2013) 109–133.
- [10] R. Cont, Y. Lu, Weak approximation of martingale representations, *Stochastic Process. Appl.* 126 (2016) 857–882.
- [11] B. Dupire, Functional Itô calculus, *Quantitative Finance* 19 (5) (2019) 721–729.
- [12] P.K. Friz, N.B. Victoir, *Multidimensional Stochastic Processes as Rough Paths*, in: *Cambridge Studies in Advanced Mathematics*, 120, Cambridge University Press, Cambridge, 2010.
- [13] I. Gyöngy, T. Pröhle, On the approximation of stochastic differential equation and on stroock-varadhan's support theorem, *Comput. Math. Appl.* 19 (1990) 65–70.
- [14] M. Ledoux, Z. Qian, T. Zhang, Large deviations and support theorem for diffusion processes via rough paths, *Stochastic Process. Appl.* 102 (2002) 265–283.
- [15] X. Mao, *Stochastic Differential Equations and Applications*, second ed., Horwood Publishing Limited, Chichester, 2008.
- [16] A. Millet, M. Sanz-Solé, A simple proof of the support theorem for diffusion processes, in: *Séminaire de Probabilités, XXVIII*, in: *Lecture Notes in Math*, vol. 1583, Springer, Berlin, 1994, pp. 36–48.

- [17] M.S. Pakkanen, Stochastic integrals and conditional full support, *J. Appl. Probab.* 47 (2010) 650–667.
- [18] D. Revuz, M. Yor, Continuous Martingales and Brownian Motion, third ed., in: *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*, vol. 293, Springer-Verlag, Berlin, 1999.
- [19] D.W. Stroock, S.R.S. Varadhan, On the support of diffusion processes with applications to the strong maximum principle, in: *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability, Volume 3: Probability Theory*, University of California Press, Berkeley, Calif., 1972, pp. 333–359.
- [20] D.W. Stroock, S.R.S. Varadhan, *Multidimensional Diffusion Processes*, in: *Grundlehren der Mathematischen Wissenschaften, Vol. 233*, Springer-Verlag, Berlin, 1979.