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

Rama Cont, Alexander Kalinin

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

Rama Cont[*]   Alexander Kalinin[**]

June 2018

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 the solutions of 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 the Functional Ito calculus.

MSC2010 classification: 60H10 ; 28C20 ; 34K50.
Keywords: support theorem, stochastic differential equations, functional equations, semimartingales, Wiener space, functional Ito calculus.

[*]Laboratoire de Probabilités, Statistiques et Modélisation, CNRS-Sorbonne Université. Email: `Rama.Cont@math.cnrs.fr`

[**]Department of Mathematics, Imperial College, United Kingdom. Email: `a.kalinin@imperial.ac.uk`.

Contents

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 possible sample 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, with a focus on stochastic differential equations, under different function space topologies.

For diffusion processes, the support under the uniform norm was first described by Stroock and Varadhan [20, 21], a result known as the 'Stroock-Varadhan support theorem'. An extension to unbounded coefficients was given by Gyöngy [14]. The support of more general Wiener functionals and extensions to SDEs in Hilbert space are discussed in Aida et al [2] and [1]. These results were extended to the Hölder topology by Ben Arous et al. [5] and, using different techniques, by Millet and Sanz-Solé [16] ; Bally et al [4] use similar methods to derive a support theorem in Hölder norms for parabolic SPDEs. Support theorems in p-variation topology are discussed by Ledoux et al [15] using rough path techniques. Support theorems in Hölder and p-variation topologies are discussed in [13]. Pakkanen [18] gives conditions for a stochastic integral ito have full support.

In this work, we extend some of these results to stochastic differential equations with *path-dependent* coefficients. Let $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\in[0,T]}, P)$ be a filtered probability space on which there is a standard d-dimensional $(\mathscr{F}_t)_{t\in[0,T]}$-Brownian motion W. Consider the following stochastic differential equation

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

whose coefficients $b : [r, T] \times S \to \mathbb{R}^m$ and $\sigma : [r, T] \times S \to \mathbb{R}^{m\times d}$ are *non-anticipative* i.e. $b(t, X), \sigma(t, X)$ depend on the path $X^t = X(t \wedge .)$ of the solution up to t. Under Lipschitz conditions on the coefficients b, σ, this SDE has a unique solution X [17, 19] whose sample paths lie in some Hölder space $C^\alpha([0, T], \mathbb{R}^m)$. Our main result is a description of the support of the solution in the Hölder topology: we show that the support of the law of the solution is given by the image of the Cameron-Martin space H^1 under the flow associated with a system of functional differential equations.

1.2 Statement of main result

Let $T > r \geq 0$ and $d, m \in \mathbb{N}$. To keep notations simple, we denote by $|\cdot|$ the absolute value function, the Euclidean norm in $\mathbb{R}^d$ and $\mathbb{R}^m$ and the Frobenius norm in $\mathbb{R}^{m\times d}$. Denote by

$$S := C([0,T], \mathbb{R}^m) \tag{1.2}$$

the space of continuous $\mathbb{R}^m$-valued maps on $[0,T]$ equipped with the supremum norm $\|\cdot\|$ and by $C_r^\alpha([0,T],\mathbb{R}^m)$ the space of $x \in S$ that are α-Hölder-continuous on $[r,T]$ for $\alpha \in (0,1]$, endowed with the 'delayed Hölder norm'

$$\|x\|_{\alpha,r} := \|x\| + \sup_{s,t\in[r,T]:\, s\neq t} \frac{|x(s) - x(t)|}{|s-t|^\alpha}. \tag{1.3}$$

We set $C_0^r([0,T],\mathbb{R}^m) := S$ and $\|\cdot\|_{0,r} := \|\cdot\|$ by convention. Then, under the assumptions stated below, (1.1) admits a unique strong solution whose sample paths lie in Hölder space $C_r^\alpha([0,T],\mathbb{R}^m)$ for all $\alpha \in [0,1/2)$.

We denote $H_r^1([0,T],\mathbb{R}^m)$ the space of absolutely continuous functions on $[r,T]$ whose derivative $\dot{h}$ is square-integrable with respect to the Lebesgue measure. We equip this space with the norm

$$\|x\|_{H,r} := \|x\| + \left(\int_r^T |\dot{x}(s)|^2\, ds\right)^{1/2}. \tag{1.4}$$

Then $H_r^1([0,T],\mathbb{R}^m) \subsetneq C_r^{1/2}([0,T],\mathbb{R}^m)$ and every $x \in H_r^1([0,T],\mathbb{R}^m)$ satisfies $\|x\|_{1/2,r} \leq \|x\|_{H,r}$.

Using the concepts of horizontal and vertical differentiability for non-anticipative functionals [7, 12], we introduce in Section 2.1 regularity assumptions on the coefficient $\sigma \in \mathbb{C}^{1,2}([r,T)\times S, \mathbb{R}^{m\times d})$ and consider the map $\rho : [r,T)\times S \to \mathbb{R}^m$ given coordinatewise by

$$\rho_k(t,x) := \sum_{l=1}^d \partial_x \sigma_{k,l}(t,x)\sigma(t,x)e_l, \tag{1.5}$$

where $\{e_1,\dots,e_d\}$ is the canonical basis of $\mathbb{R}^d$ and $\partial_x\sigma_{k,l} : [r,T)\times S \to \mathbb{R}^{1\times m}$ denotes the vertical derivative [6,12] of the (k,l)-coordinate function of σ for any $k \in \{1,\dots,m\}$ and $l \in \{1,\dots,d\}$. Note that $\rho = \partial_x\sigma\cdot\sigma$ for $m = d = 1$.

In this context, the support of the unique strong solution to (1.1) may be characterized by studying the following path-dependent ordinary differential equation driven by an element $h \in H_r^1([0,T],\mathbb{R}^d)$:

$$\dot{x}_h(t) = (b - (1/2)\rho)(t, x_h^t) + \sigma(t, x_h^t)\,\dot{h}(t) \quad \text{for } t \in [r,T]. \tag{1.6}$$

Our main result may be stated as follows:

Theorem 1 (Support theorem for path-dependent SDEs).
Let $\hat{x} \in C([0,T],\mathbb{R}^m)$ *and* $\sigma \in \mathbb{C}^{1,2}([r,T) \times S, \mathbb{R}^{m\times d})$ *with horizontal (resp. vertical) derivative denoted by* $\partial_t \sigma$ *(resp.* $\partial_x \sigma$*). Assume* σ *and* $\partial_x \sigma$ *are bounded and there are constants* $c, \eta, \lambda \geq 0$ *and* $\kappa \in [0,1)$ *such that*

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

for all $s,t \in [r,T)$, $x,y \in S$, $k \in \{1,\dots,m\}$ *and* $l \in \{1,\dots,d\}$*. Then:*

(i) 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,* $E[\|X\|_{\alpha,r}^{2p}] < \infty$ *for all* $p \geq 1$ *and* $\alpha \in [0,1/2)$*.*

(ii) For any $h \in H_r^1([0,T],\mathbb{R}^d)$ *there is a unique mild solution* x_h *to* (1.6) *so 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* $H_r^1([0,T],\mathbb{R}^d) \to H_r^1([0,T],\mathbb{R}^m)$, $h \mapsto x_h$ *is Lipschitz continuous on bounded sets.*

(iii) For each $\alpha \in [0,1/2)$*, 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* (1.6)*, where* $h \in H_r^1([0,T],\mathbb{R}^d)$*. That is,*

$$\mathrm{supp}(P \circ X^{-1}) = \overline{\{x_h \,|\, h \in H_r^1([0,T],\mathbb{R}^d)\}} \quad \text{in } C_r^\alpha([0,T],\mathbb{R}^m). \quad (1.7)$$

This result extends previous results [2,5,16,20] on the support of diffusion processes to the case of path-dependent coefficients. In the diffusion case, we retrieve the results of [5,16] under weaker assumptions on σ.

Our proof adapts the approach used by Millet and Sanz-Solé [16] to the path-dependent case, using the tools of Functional Ito calculus [6,12]. We construct Hölder-continuous approximations of the solution using an adapted linear interpolation of Brownian motion and show that this approximation converges in probability to the solution in Hölder norm. A key ingredient is the use of functional estimates derived in [3] using the Functional Ito calculus, combined with interpolation error estimates in Hölder norm for stochastic processes.

Outline. The remainder of the paper is devoted to the proof of Theorem Support Theorem. Section 2 discusses the various building blocks of the proof. Section 2.1 recalls some functional calculus concepts from [6] and establishes several results useful in our setting. Section 2.2 gives conditions for existence and uniqueness of a (mild) solution to the path-dependent ODE (1.6); Section 2.3 gives conditions for the existence of a unique strong solution to (1.1); Section 2.4 dicusses the interpolation method used to characterize the support in Hölder topologies.

Section 3 discusses Hölder spaces for stochastic processes and the notion of convergence in probability in Hölder norm in more depth. Section 3.2 derives a variation on the Kolmogorov-Chentsov theorem with an estimate for the Hölder norm (Lemma 12) and an improved version of a statement from [16] (Proposition 14). Section 3.4 discusses adapted linear interpolation of Brownian motion, improving some results from [16]. Section 4 uses these ingredients to prove the existence and uniqueness of mild solutions to path-dependent ODEs (Sec. 4.1) and SDEs (Sec. 4.2). Finally, Section 5 combines these ingredients to give a proof of the main result.

2 Preliminaries

We shall denote $\mathbb{I}_d$ the $d \times d$ identity matrix; for a matrix A we denote by A' its transpose.

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 $\|\cdot\|$ and recall the following notions from [6,8]. A functional $F:[r,T]\times D([0,T],\mathbb{R}^m)\to\mathbb{R}$ is *non-anticipative* if

$$F(t,x) = F(t,x^t)$$

for all $t\in[r,T]$ and $x\in D([0,T],\mathbb{R}^m)$, where x^t is *path x stopped at time t*: $x^t(s) = x(s\wedge t)$ for each $s\in[0,T]$. F is called *boundedness-preserving* if for each $n\in\mathbb{N}$ there is $c_n \geq 0$ such that

$$|F(t,x)| \leq c_n$$

for every $t\in[r,T]$ and $x\in D([0,T],\mathbb{R}^m)$ satisfying $\|x\|\leq n$. In other words, F is ought to be bounded on bounded sets. We notice that the following pseudometric on $[r,T]\times D([0,T],\mathbb{R}^m)$ given by

$$d_\infty((t,x),(s,y)) := |t-s|^{1/2} + \|x^t - y^s\| \tag{2.1}$$

is complete and if F 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. A functional $F : [r, T) \times D([0,T], \mathbb{R}^m) \to \mathbb{R}$ is called *horizontally differentiable* if for each $t \in [r,T)$ and $x \in D([0,T], \mathbb{R}^m)$, the function

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

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

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

is differentiable at 0. In this case, its derivative there will be represented by $\partial_x F(t,x)$. We call F *partially vertically differentiable* if for all $k \in \{1, \ldots, m\}$, $t \in [r,T)$ and $x \in D([0,T], \mathbb{R}^m)$, the function

$$\mathbb{R} \to \mathbb{R}, \quad h \mapsto F(t, x + h\hat{e}_k \mathbb{1}_{[t,T]})$$

is differentiable at 0, where $\{\hat{e}_1, \ldots, \hat{e}_m\}$ is the canonical basis of $\mathbb{R}^m$. In this case, its derivative there will be denoted by $\partial_{x_i} F(t,x)$. By calculus, if F is vertically differentiable, then it is partially vertically differentiable and $\partial_x F = (\partial_{x_1} F, \ldots, \partial_{x_m} F)$.

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

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

It follows from Schwarz's Lemma that if F is twice vertically differentiable and $\partial_{xx} F$ is d_∞-continuous, then $\partial_{xx} F$ is symmetric: $\partial_{x_k x_l} F = \partial_{x_l x_k} F$ for each $k, l \in \{1, \ldots, m\}$.

A functional $G : [r,T) \times D([0,T], \mathbb{R}^m) \to \mathbb{R}$ is said to be *of class* $\mathbb{C}^{1,2}$ if it is once horizontally and twice vertically differentiable such that G itself and the derivatives $\partial_t G$, $\partial_x G$ and $\partial_{xx} G$ are boundedness-preserving and d_∞-continuous. By

$$\mathbb{C}^{1,2}([r,T) \times S) \tag{2.2}$$

we denote the space of functionals $F : [r,T) \times S \to \mathbb{R}^{m \times d}$ that admit an extension $G : [r,T) \times D([0,T], \mathbb{R}^m) \to \mathbb{R}$ of class $\mathbb{C}^{1,2}$, where S is given by (1.2). Then it follows from [6, Theorem 5.4.1] that

$$\partial_t F := \partial_t G, \quad \partial_x F := \partial_x G \quad \text{and} \quad \partial_{xx} F := \partial_{xx} G \quad \text{on } [r,T) \times S$$

are independent of the choice of the extension G. Note that (2.2) allows us to use the functional Itô formula [9] in the proof of Proposition 34, which gives one of the main arguments to establish (1.7). To conclude, we write $\mathbb{C}([r,T) \times S, \mathbb{R}^{m \times d})$ for the linear space of all maps $F : [r,T) \times S \to \mathbb{R}^{m \times d}$ satisfying $F_{k,l} \in \mathbb{C}([r,T) \times S)$ for each $k \in \{1, \ldots, m\}$ and $l \in \{1, \ldots, d\}$.

2.2 Mild solutions to path-dependent ODEs

We show in this section a unique mild solution to the ODE (1.6), which belongs to the delayed Cameron-Martin space $H_r^1([0,T],\mathbb{R}^m)$. To this end, let us consider the general path-dependent ordinary differential equation

$$\dot{x}(t) = F(t, x^t) \quad \text{for } t \in [r,T], \tag{2.3}$$

where $F : [r,T] \times S \to \mathbb{R}^m$ denotes a non-anticipative product measurable map. Then for each $h \in H_r^1([0,T],\mathbb{R}^d)$ the choice $F = b - (1/2)\rho + \sigma\dot{h}$, where ρ is given by (1.5), yields the support characterizing ODE (1.6).

As for $x \in S$ the map $[r,T] \to \mathbb{R}^m$, $t \mapsto F(t,x^t)$ may fail to be continuous, one may in general not expect to derive classical solutions. So, we recall the concept of a *mild solution* to (2.3), which is a path $x \in S$ satisfying

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

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 Borel measurable map $[r,T] \to \mathbb{R}^m$, $s \mapsto F(s,x^s)$ 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.

(O.i) There exists a measurable function $c_F : [r,T] \to [0,\infty)$ satisfying $\int_r^T c_F(s)^2\,ds < \infty$ and

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

for all $t \in [r,T)$ and $x \in S$ that is absolutely continuous on $[r,T]$.

(O.ii) For each $n \in \mathbb{N}$ there is a measurable function $\lambda_{F,n} : [r,T] \to [0,\infty)$ such that $\int_r^T \lambda_{F,n}(s)^2\,ds < \infty$ and

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

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

Under the above growth condition and Lipschitz smoothness on bounded sets, we obtain a unique mild solution that can be approximated by a Picard iteration in the delayed Cameron-Martin norm $\|\cdot\|_{H,r}$ given by (1.4).

Proposition 2. *Let* (O.i) *and* (O.ii) *hold and* $\hat{x} \in C([0,T],\mathbb{R}^m)$*, then the ODE* (2.3) *admits a unique mild solution* y_F *satisfying*

$$y_F(s) = \hat{x}(s) \quad \textit{for all } s \in [0,r]$$

and it holds that $y_F \in H_r^1([0,T],\mathbb{R}^m)$*. Moreover, the sequence* $(x_n)_{n\in\mathbb{N}_0}$ *in* $H_r^1([0,T],\mathbb{R}^m)$*, recursively defined via* $x_0(t) := \hat{x}(r \wedge t)$ *and*

$$x_{n+1}(t) := x_0(t) + \int_r^{r\vee t} F(s, x_n^s)\, ds \tag{2.4}$$

for all $n \in \mathbb{N}_0$*, converges in the delayed Cameron-Martin norm* $\|\cdot\|_{H,r}$ *to* y_F*.*

2.3 Strong solutions for path-dependent SDEs

We turn to the derivation of a unique strong solution to (1.1), for which a.e. path lies in $C_r^\alpha([0,T],\mathbb{R}^m)$ for any $\alpha \in [0,1/2)$. We consider the stochastic differential equation with path-dependent coefficients

$$dX_t = B(t, X^t)\, dt + \Sigma(t, X^t)\, dW_t \quad \text{for } t \in [r,T], \tag{2.5}$$

where $B : [r,T]\times S \to \mathbb{R}^m$ and $\Sigma : [r,T]\times S \to \mathbb{R}^{m\times d}$ are two non-anticipative product measurable maps.

A strong solution to (2.5) is an $(\mathscr{F}_t)_{t\in[0,T]}$-adapted right-continuous process $X : [0,T] \times \Omega \to \mathbb{R}^m$ with a.s. continuous paths satisfying

$$\begin{aligned} &\int_r^T |B(s, X^s)|\, ds + \int_r^T |\Sigma(s, X^s)|^2\, ds < \infty \quad \text{a.s. and} \\ &X_t = X_r + \int_r^t B(s, X^s)\, ds + \int_r^t \Sigma(s, X^s)\, dW_s \quad \text{for all } t \in [r,T] \text{ a.s.} \end{aligned}$$

Remark 3. The fact that we do not have to assume the usual conditions is clarified in Section 3.1 and irrespective how $B(s,y)$ and $\Sigma(s,y)$ are extended for $s \in [r,T]$ and any right-continuous map $y : [0,T] \to \mathbb{R}^m$ that is not continuous, the above integrals remain unchanged up to indistinguishability.

We now state the assumptions on the coefficients, valid in the setting of Theorem 1 for the choice $B = b$ and $\Sigma = \sigma$.

(S.i) There are a measurable function $c_B : [r,T] \to [0,\infty)$ and a constant $c_\Sigma \geq 0$ such that $\int_r^T c_B(s)^2\, ds < \infty$ and for all $(t,x) \in [r,T) \times S$,

$$|B(t,x)| \leq c_B(t)(1 + \|x\|) \quad \text{and} \quad |\Sigma(t,x)| \leq c_\Sigma(1 + \|x\|).$$

(S.ii) There are $\alpha_0 \in [0,1/2)$, a measurable function $\lambda_B : [r,T] \to [0,\infty)$ and a constant $\lambda_\Sigma \geq 0$ such that $\int_r^T \lambda_B(s)^2\,ds < \infty$ and

$$|B(t,x) - B(t,y)| \leq \lambda_B(t)\|x-y\|_{\alpha_0,r},$$
$$|\Sigma(t,x) - \Sigma(t,y)| \leq \lambda_\Sigma\|x-y\|_{\alpha_0,r}$$

for all $t \in [r,T)$ and $x,y \in C_r^{\alpha_0}([0,T],\mathbb{R}^m)$, where $\|\cdot\|_{\alpha_0,r}$ equals $\|\cdot\|$ if $\alpha_0 = 0$ and otherwise is given by (1.3) when α is replaced by α_0.

Remark 4. If condition (S.ii) holds, then it is also true if α_0 is replaced by any $\alpha \in (\alpha_0, 1/2)$. Thus, it is strongest in the case that $\alpha_0 = 0$.

For $p \geq 1$ and $\alpha \in [0,1]$ we let $\mathscr{C}_{r,p}^\alpha([0,T],\mathbb{R}^m)$ denote the space of all $(\mathscr{F}_t)_{t\in[0,T]}$-adapted right-continuous processes $X : [0,T]\times\Omega \to \mathbb{R}^m$ satisfying $E[\|X\|_{\alpha,r}^p] < \infty$, equipped with the intrinsic seminorm

$$\mathscr{C}_{r,p}^\alpha([0,T],\mathbb{R}^m) \to [0,\infty), \quad X \mapsto \Big(E\Big[\|X\|_{\alpha,r}^p\Big]\Big)^{1/p}, \tag{2.6}$$

which is complete, by Proposition 11. Moreover, if a sequence $({}_nX)_{n\in\mathbb{N}}$ in this linear space converges with respect to the above seminorm, then it also converges in the delayed Hölder norm $\|\cdot\|_{\alpha,r}$ in probability. Finally, we set $\mathscr{C}_\infty([0,T],\mathbb{R}^m) := \cap_{p\geq 1}\mathscr{C}_{r,p}^0([0,T],\mathbb{R}^m)$ and let

$$\mathscr{C}_{r,\infty}^{1/2-}([0,T],\mathbb{R}^m)$$

denote the intersection of the spaces $\mathscr{C}_{r,p}^\alpha([0,T],\mathbb{R}^m)$ over all $p \geq 1$ and $\alpha \in [0,1/2)$, which yields a completely pseudometrizable topological space.

Proposition 5. *Assume* (S.i)-(S.ii) *and let $\hat{X} \in \mathscr{C}_\infty([0,T],\mathbb{R}^m)$. Then up to indistinguishability there is a unique strong solution X to* (2.5) *such that*

$$X_s = \hat{X}_s \quad \textit{for all } s \in [0,r] \textit{ a.s.}$$

we have that $X \in \mathscr{C}_{r,\infty}^{1/2-}([0,T],\mathbb{R}^m)$. Furthermore, the sequence $({}_nX)_{n\in\mathbb{N}_0}$ in $\mathscr{C}_{r,\infty}^{1/2-}([0,T],\mathbb{R}^m)$, recursively given by ${}_0X_t := \hat{X}_{r\wedge t}$ and

$${}_{n+1}X_t = {}_0X_t + \int_r^{r\vee t} B(s,{}_nX^s)\,ds + \int_r^{r\vee t}\Sigma(s,{}_nX^s)\,dW_s \tag{2.7}$$

for all $n \in \mathbb{N}_0$, converges in the seminorm (2.6) *to X for each $p \geq 2$ and $\alpha \in [0,1/2)$. In particular,* $\lim_{n\uparrow\infty} P(\|{}_nX - X\|_{\alpha,r} \geq \varepsilon) = 0$ *for all $\varepsilon > 0$.*

Remark 6. Pathwise uniqueness is shown in Lemma 30, requiring only the following Lipschitz condition on bounded sets, which follows from (S.ii) in the strongest case $\alpha_0 = 0$:

(S.iii) For each $n \in \mathbb{N}$ there is a measurable function $\lambda_n : [r,T] \to [0,\infty)$ satisfying $\int_r^T \lambda_n(s)^2\,ds < \infty$ and

$$|B(t,x) - B(t,y)| + |\Sigma(t,x) - \Sigma(t,y)| \leq \lambda_n(t)\|x - y\|$$

for all $t \in [r,T]$ and $x, y \in S$ with $\|x\| \vee \|y\| \leq n$.

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 shall characterize the support (1.7). For $n \in \mathbb{N}$ let $\mathbb{T}_n$ be a partition of $[r,T]$ that we write in the form

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

for some $k_n \in \mathbb{N}$ and $t_{0,n}, \ldots, t_{k_n,n} \in [r,T]$ so that $r = t_{0,n} < \cdots < t_{k_n,n} = T$ and we denote its mesh by $|\mathbb{T}_n| = \max_{i\in\{0,\ldots,k_n-1\}}(t_{i+1,n} - t_{i,n})$. We assume that $\lim_{n\uparrow\infty} |\mathbb{T}_n| = 0$ and that the sequence of partitions is *well-balanced* in the sense of [11], that is, there is $c_{\mathbb{T}} \geq 1$ such that

$$|\mathbb{T}_n| \leq c_{\mathbb{T}} \inf_{i\in\{0,\ldots,k_n-1\}} (t_{i+1,n} - t_{i,n}) \quad \text{for all } n \in \mathbb{N}. \tag{2.8}$$

Moreover, for $n \in \mathbb{N}$ we define an $(\mathscr{F}_t)_{t\in[0,T]}$-adapted right-continuous process ${}_nW : [0,T] \times \Omega \to \mathbb{R}^d$ by setting ${}_nW_t := W_{r\wedge t}$ for $t \in [0, t_{1,n})$,

$${}_nW_t := W_{t_{i-1,n}} + (t - t_{i,n})\frac{W_{t_{i,n}} - W_{t_{i-1,n}}}{t_{i+1,n} - t_{i,n}} \tag{2.9}$$

for $t \in [t_{i,n}, t_{i+1,n})$ with $i \in \{1, \ldots, k_n - 1\}$ and ${}_nW_T := W_{t_{k_{n-1},n}}$. Then ${}_nW$ can be regarded as adapted linear interpolation of the d-dimensional $(\mathscr{F}_t)_{t\in[0,T]}$-Brownian motion W on $[r,T]$ and almost each of its paths belongs to $H_r^1([0,T], \mathbb{R}^d)$.

Thus, let us suppose that the assumptions and the first two claims of Theorem 1 hold. To establish $\mathrm{supp}(P \circ X^{-1}) \subseteq \overline{\{x_h \,|\, h \in H_r^1([0,T], \mathbb{R}^d)\}}$ 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

$$\lim_{n\uparrow\infty} P(\|x_{{}_nW} - X\|_{\alpha,r} \geq \varepsilon) = 0 \quad \text{for all } \varepsilon > 0. \tag{2.10}$$

By definition of a mild solution to (2.3), we see for each $n \in \mathbb{N}$ that $x_{{}_nW}$ is a strong solution to the degenerate path-dependent SDE

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

with initial condition ${}_nY^r = \hat{x}^r$ a.s. For each $h \in H^1_r([0,T],\mathbb{R}^d)$ and $n \in \mathbb{N}$, we introduce an a.s. continuous local martingale ${}_{h,n}Z : [0,T] \times \Omega \to (0,\infty)$ by requiring that ${}_{h,n}Z^r = 1$ a.s. and

$$ {}_{h,n}Z_t = \exp\left(\int_r^t \dot{h}(s) - {}_n\dot{W}_s\, dW_s - \frac{1}{2}\int_r^t |\dot{h}(s) - {}_n\dot{W}_s|^2\, ds\right) \tag{2.11} $$

for all $t \in [r,T]$ a.s. In fact, ${}_{h,n}Z$ is a martingale, as clarified in Lemma 39. Hence, $P_{h,n} : \mathscr{F} \to [0,1]$ given by $P_{h,n}(A) := E[{}_{h,n}Z_T \mathbb{1}_A]$ is a probability measure equivalent to P. By using this fact, we will show that the converse inclusion in (1.7) follows once we have proven that

$$ \lim_{n\uparrow\infty} P_{h,n}(\|X - x_h\|_{\alpha,r} \geq \varepsilon) = 0 \quad \text{for all } \varepsilon > 0. \tag{2.12} $$

By Girsanov's theorem, the process ${}_{h,n}W : [0,T] \times \Omega \to \mathbb{R}^d$ defined via ${}_{h,n}W_t := W_t - \int_r^{r\vee t} \dot{h}(s) - {}_n\dot{W}_s\, ds$ is a d-dimensional $(\mathscr{F}_t)_{t\in[0,T]}$-Brownian motion under $P_{h,n}$ and X is a strong solution to the path-dependent SDE

$$ d_nY_t = \Big(b(t,{}_nY^t) + \sigma(t,{}_nY^t)(\dot{h}(t) - {}_n\dot{W}_t)\Big)\, dt + \sigma(t,{}_nY^t)\, d_{h,n}W_t \tag{2.13} $$

for $t \in [r,T]$ under $P_{h,n}$ with initial condition ${}_nY^r = \hat{x}^r$ a.s. Hence, to prove (2.10) and (2.12) at the same time, we consider the following general framework.

Namely, we let $\underline{B} : [r,T] \times S \to \mathbb{R}^m$ and $B_H : [r,T] \times S \to \mathbb{R}^{m\times d}$ be two non-anticipative product measurable maps and $\overline{B} \in \mathbb{C}^{1,2}([r,T) \times S, \mathbb{R}^{m\times d})$. Then for each $n \in \mathbb{N}$ we introduce the path-dependent SDE

$$ \begin{aligned} d_nY_t = {} & \Big(\underline{B}(t,{}_nY^t) + B_H(t,{}_nY^t)\dot{h}(t) + \overline{B}(t,{}_nY^t){}_n\dot{W}_t\Big)\, dt \\ & + \Sigma(t,{}_nY^t)\, dW_t \quad \text{for } t \in [r,T], \end{aligned} \tag{2.14} $$

where $\Sigma : [r,T] \times S \to \mathbb{R}^{m\times d}$ is a non-anticipative product measurable map, as considered in Section 2.3. In addition, we introduce the path-dependent SDE

$$ dZ_t = \Big((\underline{B} + R)(t,Z^t) + B_H(t,Z^t)\dot{h}(t)\Big)\, dt + (\overline{B} + \Sigma)(t,Z^t)\, dW_t \tag{2.15} $$

for $t \in [r,T]$, where we require the non-anticipative product measurable map $R : [r,T) \times S \to \mathbb{R}^m$ given coordinatewise by

$$ R_k(s,y) := \sum_{l=1}^d \partial_x \overline{B}_{k,l}(s,y)((1/2)\overline{B} + \Sigma)(s,y)e_l. \tag{2.16} $$

In Theorem 8 below, we in particular show that whenever ${}_nY$ and Z are strong solutions to (2.14) and (2.15), respectively, such that ${}_nY^r = Z^r = \hat{x}^r$ a.s. for each $n \in \mathbb{N}$, then

$$\lim_{n\uparrow\infty} P(\|{}_nY - Z\|_{\alpha,r} \geq \varepsilon) = 0 \quad \text{for all } \varepsilon > 0. \tag{2.17}$$

Then the choice $\underline{B} = b - (1/2)\rho$, $B_H = 0$, $\overline{B} = \sigma$ and $\Sigma = 0$ yields (2.10), since $R = (1/2)\rho$ in this case. Moreover, by choosing $\underline{B} = b$, $B_H = \sigma$, $\overline{B} = -\sigma$ and $\Sigma = \sigma$ instead, (2.12) follows. Since these are the two desired results, we consider the following regularity conditions:

(C.i) $\overline{B} \in \mathbb{C}^{1,2}([r,T) \times S, \mathbb{R}^{m\times d})$ and there are $c, \eta \geq 0$ and $\kappa \in [0,1)$ such that for all $(t,x) \in [r,T) \times S$, $|\underline{B}(t,x)| + |B_H(t,x)| \leq c(1 + \|x\|^\kappa)$,

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

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

(C.ii) $\underline{B}$ is Lipschitz continuous in $x \in S$, uniformly in $t \in [r,T)$, and B_H, $\overline{B}$, $\partial_x\overline{B}$ and Σ are Lipschitz continuous with respect to d_∞ given by (2.1).

(C.iii) There is a measurable function $\overline{b} : [r,T] \to \mathbb{R}$ so that $\int_r^T |\overline{b}(s)|^2\, ds < \infty$ and $\overline{B}(s,y) = \Sigma(s,y)\overline{b}(s)$ for all $(s,y) \in [r,T) \times S$.

Remark 7. Condition (C.iii) allows us to perform a change of measures to get a unique strong solution to (2.14). However, when deriving (2.17) in Sections 5.1, 5.2 and 5.3 we merely assume that (C.i) and (C.ii) hold.

Theorem 8. *Let* (C.i)-(C.iii) *hold and* $h \in H_r^1([0,T],\mathbb{R}^d)$. *Assume that* $\hat{X} \in \mathscr{C}_\infty([0,T],\mathbb{R}^m)$ *and* $({}_n\hat{X})_{n\in\mathbb{N}}$ *is a sequence in* $\mathscr{C}_\infty([0,T],\mathbb{R}^m)$ *satisfying*

$$\sup_{n\in\mathbb{N}} E\Big[\|{}_n\hat{X}^r\|^{2p}\Big] < \infty \quad \textit{for each } p \geq 1.$$

Then the following three assertions hold:

(i) For any $n \in \mathbb{N}$ *there is a unique strong solution* ${}_nY$ *to* (2.14) *satisfying* ${}_nY^r = {}_n\hat{X}^r$ *a.s. Moreover,* $\sup_{n\in\mathbb{N}} E[\|{}_nY\|_{\alpha,r}^{2p}] < \infty$ *for all* $p \geq 1$ *and* $\alpha \in [0,1/2)$.

(ii) There is a unique strong solution Z *to* (2.15) *such that* $Z^r = \hat{X}^r$ *a.s. and we have* $E[\|Z\|_{\alpha,r}^{2p}] < \infty$ *for all* $p \geq 1$ *and* $\alpha \in [0,1/2)$.

(iii) Let $\alpha \in [0,1/2)$ and $\lim_{n\uparrow\infty} E[\|_n\hat{X}^r - \hat{X}^r\|^2]/|\mathbb{T}_n|^{2\alpha} = 0$, then it follows that

$$\lim_{n\uparrow\infty} E\Big[\max_{j\in\{0,\dots,k_n\}} |_nY_{t_{j,n}} - Z_{t_{j,n}}|^2\Big]/|\mathbb{T}_n|^{2\alpha} = 0. \tag{2.18}$$

In particular, (2.17) *holds, that is, $(_nY)_{n\in\mathbb{N}}$ converges in the delayed Hölder norm $\|\cdot\|_{\alpha,r}$ in probability to Z.*

3 Convergence in probability in Hölder norm

3.1 Hölder spaces for stochastic processes

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

$$\begin{aligned} &\mathscr{C}_r^\alpha([0,T],\mathbb{R}^m) \times \mathscr{C}_r^\alpha([0,T],\mathbb{R}^m) \to [0,\infty),\\ &(X,Y) \mapsto E[\|X-Y\|_{\alpha,r} \wedge 1]. \end{aligned} \tag{3.1}$$

We notice that a sequence $(_nX)_{n\in\mathbb{N}}$ in this pseudometric space converges to some $X \in \mathscr{C}_r^\alpha([0,T],\mathbb{R}^m)$ if and only if it converges to this process in the delayed Hölder norm $\|\cdot\|_{\alpha,r}$ in probability. Put differently, $(\|_nX - X\|_{\alpha,r})_{n\in\mathbb{N}}$ converges to zero in probability. Further, $(_nX)_{n\in\mathbb{N}}$ is Cauchy if and only if it is Cauchy in the norm $\|\cdot\|_{\alpha,r}$ in probability in the sense that

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

Next, we set $\mathscr{C}([0,T],\mathbb{R}^m) := \mathscr{C}_r^0([0,T],\mathbb{R}^m)$, which is the linear space of all $\mathbb{R}^m$-valued adapted right-continuous a.s. continuous processes. Despite the fact that we do not assume the usual conditions, $\mathscr{C}([0,T],\mathbb{R}^m)$ is complete, which yields the following result.

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

Proof. Let $(_nX)_{n\in\mathbb{N}}$ be a Cauchy sequence in $\mathscr{C}_r^\alpha([0,T],\mathbb{R}^m)$. By 4.3.3 Lemma in [21], there is $X \in \mathscr{C}([0,T],\mathbb{R}^m)$ 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

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

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 $P(\|_{k_l}X - X\| \geq (\varepsilon/4)\delta_l^\alpha) < \eta/2$.

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 all $n \in \mathbb{N}$ 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. □

For $p \geq 1$ we recall that $\mathscr{C}^\alpha_{r,p}([0,T],\mathbb{R}^m)$ denotes the space of all processes $X \in \mathscr{C}^\alpha_r([0,T],\mathbb{R}^m)$ satisfying $E[\|X\|^p_{\alpha,r}] < \infty$, endowed with the seminorm (2.6). A sequence $({}_nX)_{n\in\mathbb{N}}$ in this space is called p-fold uniformly integrable if $(\|{}_nX\|_{\alpha,r})_{n\in\mathbb{N}}$ satisfies this property in the usual sense.

Lemma 10. *Any sequence $({}_nX)_{n\in\mathbb{N}}$ in $\mathscr{C}^\alpha_{r,p}([0,T],\mathbb{R}^m)$ that is Cauchy with respect to the seminorm* (2.6) *is p-fold uniformly integrable.*

Proof. Let $\varepsilon > 0$, then there is $n_0 \in \mathbb{N}$ so that $E[\|{}_kX - {}_nX\|^p_{\alpha,r}] < \varepsilon/2^p$ 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

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

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

We conclude with the following convergence characterization.

Proposition 11. *A sequence $({}_nX)_{n\in\mathbb{N}}$ in $\mathscr{C}^\alpha_{r,p}([0,T],\mathbb{R}^m)$ converges in the seminorm* (2.6) *if and only if it is p-fold uniformly integrable and there is $X \in \mathscr{C}^\alpha_r([0,T],\mathbb{R}^m)$ such that*

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

In the latter case, we have $E[\|X\|^p_{\alpha,r}] < \infty$ and $\lim_{n\uparrow\infty} E[\|{}_nX - X\|^p_{\alpha,r}] = 0$. Moreover, $\mathscr{C}^\alpha_{r,p}([0,T],\mathbb{R}^m)$ equipped with (2.6) *is complete.*

Proof. By Lemmas 9 and 10, it suffices to show the if-direction of the first claim. To this end, 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 to zero a.s., then

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

by Fatou's Lemma. Now let $\varepsilon > 0$, then there exist $\delta > 0$ and $n_0 \in \mathbb{N}$ such that $\sup_{n\in\mathbb{N}} E[\|_n X\|_{\alpha,r}^p \mathbb{1}_A] < (\varepsilon/3)^p$ for each $A \in \mathscr{F}$ with $P(A) < \delta$ and $P(\|_n X - X\|_{\alpha,r} \geq \varepsilon/3) < \delta$ for all $n \in \mathbb{N}$ with $n \geq n_0$. Thus,

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

for every such $n \in \mathbb{N}$, since similar reasoning as above gives $E[\|X\|_{\alpha,r}^p \mathbb{1}_A] \leq \sup_{n\in\mathbb{N}} E[\|_n X\|_{\alpha,r}^p \mathbb{1}_A]$ for all $A \in \mathscr{F}$. This completes the proof. □

3.2 A general Kolmogorov-Chentsov estimate

In this section we revisit the proof of the Kolmogorov-Chentsov Theorem to allow for processes that are merely right-continuous and to obtain a quantitative estimate of the Hölder norm. Let

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

for $p \geq 1/2$, $q > 0$ and $\alpha \in (0, q/(2p))$ and note that $\lim_{\alpha\uparrow q/(2p)} k_{\alpha,q,p} = \infty$. Then we have the following result.

Proposition 12. *Let $\mathscr{H}$ be a set of $\mathbb{R}^m$-valued right-continuous processes and $s, t \in [r, T]$ be such that $s < t$. Asume that there are $c_0 \geq 0$, $p \geq 1/2$ and $q > 0$ such that*

$$\sup_{U\in\mathscr{H}} E\Big[|U_u - U_v|^{2p}\Big] \leq c_0|u - v|^{1+q} \tag{3.3}$$

for all $u, v \in [s, t]$. Then for each $\alpha \in (0, q/(2p))$ it holds that

$$\sup_{U\in\mathscr{H}} E\Bigg[\bigg(\sup_{u,v\in[s,t]:\, u\neq v} \frac{|U_u - U_v|}{|u - v|^\alpha}\bigg)^{2p}\Bigg] \leq k_{\alpha,q,p}c_0(t - s)^{1+q-2\alpha p}.$$

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

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

then it is readily seen that there are 2^n tuples $(u, v) \in \Delta_n$ satisfying $u < v$. For fixed $U \in \mathscr{H}$ we set $V_n := \sup_{(u,v)\in\Delta_n} |U_u - U_v|$, then (3.3) gives

$$E\Big[V_n^{2p}\Big] \leq \sum_{(u,v)\in\Delta_n:\, u<v} E\Big[|U_u - U_v|^{2p}\Big] \leq 2^{-nq}c_0(t - s)^{1+q}. \tag{3.4}$$

We now set $\mathbb{D} := \bigcup_{n\in\mathbb{N}_0}\mathbb{D}_n$ and let $u,v\in\mathbb{D}$ satisfy $0 < v-u < 2^{-n}(t-s)$ 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 $d_{i_k-1,k}\leq u < d_{i_k,k}$, and

$$d_{j_k-1,k}\leq v < d_{j_k,k}\quad\text{for } v<t\quad\text{and}\quad d_{j_k,k}=t\quad\text{for } v=t,$$

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

$$U_u - U_v = U_{d_{i_n,n}} - U_{d_{j_n,n}} + \sum_{k=n}^{\infty} U_{d_{i_{k+1},k+1}} - U_{d_{i_k,k}} + \sum_{k=n}^{\infty} U_{d_{j_{k+1},k+1}} - U_{d_{j_k,k}}.$$

We note that either $i_n = j_n$ or instead $n\geq 1$, $j_n\geq 2$ and $i_n = j_n - 1$, as $0 < v-u < 2^{-n}(t-s)$. In both cases, $(d_{i_n,n}, d_{j_n,n})\in\Delta_n$. 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,

$$|U_u - U_v| \leq V_n + 2\sum_{k=n}^{\infty} V_{k+1} \leq 2\sum_{k=n}^{\infty} V_k. \tag{3.5}$$

Next, pick $u,v\in\mathbb{D}$ such that $0 < v-u < t-s$, then there is a unique $n\in\mathbb{N}_0$ such that $2^{-n-1}(t-s)\leq v-u < 2^{-n}(t-s)$. By (3.5),

$$(v-u)^{-\alpha}|U_u - U_v| \leq 2^{1+\alpha}(t-s)^{-\alpha}\sum_{k=n}^{\infty} 2^{\alpha k}V_k$$

because $2^{\alpha n}\leq 2^{\alpha k}$ for all $k\in\mathbb{N}_0$ with $k\geq n$. Clearly, $u,v\in\mathbb{D}$ satisfy $v-u = t-s$ if and only if $u=s$ and $v=t$. In this case, $|U_s - U_t|/(t-s)^{\alpha}$ $\leq (t-s)^{-\alpha}Z_0$. Thus, we have shown that

$$\sup_{u,v\in[s,t]:\, u\neq v}\frac{|U_u - U_v|}{|u-v|^{\alpha}} \leq 2^{1+\alpha}(t-s)^{-\alpha}\sum_{k=0}^{\infty} 2^{\alpha k}V_k, \tag{3.6}$$

as $\mathbb{D}$ is a countable dense set in $[s,t]$ containing t and U is right-continuous. Hence, (3.6), the triangle inequality, monotone convergence and (3.4) yield that

$$\begin{aligned}\left(E\left[\left(\sup_{u,v\in[s,t]:\, u\neq v}\frac{|U_u - U_v|}{|u-v|^{\alpha}}\right)^{2p}\right]\right)^{\frac{1}{2p}} &\leq 2^{1+\alpha}(t-s)^{-\alpha}\sum_{k=0}^{\infty} 2^{\alpha k}\left(E\left[V_k^{2p}\right]\right)^{\frac{1}{2p}}\\ &\leq 2^{1+\alpha}c_0^{\frac{1}{2p}}(t-s)^{\frac{1+q}{2p}-\alpha}\sum_{k=0}^{\infty} 2^{(\alpha-\frac{q}{2p})k}.\end{aligned}$$

The power series on the right-hand side converges absolutely to the inverse of $1 - 2^{\alpha - q/(2p)}$, since $\alpha < q/(2p)$. For this reason, the proposition follows. □

3.3 Convergence along a sequence of partitions

In this section, we state a sufficient criterion for a sequence of processes to converge in probability in the delayed Hölder norm $\|\cdot\|_{\alpha,r}$ where $\alpha \in [0,1]$. For this purpose, we require the following estimate.

Lemma 13. *Let $\mathbb{T}$ be a partition of $[r,T]$ of the form $\mathbb{T} = \{t_0,\ldots,t_k\}$ for some $k \in \mathbb{N}$ and $t_0,\ldots,t_k \in [r,T]$ such that $r = t_0 < \cdots < t_k = T$. Then*

$$\sup_{s,t\in[r,T]:\, s\neq t} \frac{|x(s)-x(t)|}{|s-t|^\alpha} \leq 2 \max_{j\in\{0,\ldots,k-1\}} \sup_{u,v\in[t_j,t_{j+1}]:\, u\neq v} \frac{|x(u)-x(v)|}{|u-v|^\alpha} + \max_{i,j\in\{0,\ldots,k\}:\, i\neq j} \frac{|x(t_i)-x(t_j)|}{|t_i-t_j|^\alpha}$$

for each map $x : [r,T] \to \mathbb{R}^m$.

Proof. First, assume that $i,j \in \{0,\ldots,k-1\}$ are such that $i < j-1$ and $s \in (t_i,t_{i+1})$, then

$$\frac{|x(s)-x(t_j)|}{|s-t|^\alpha} \leq \frac{|x(s)-x(t_{i+1})|}{|s-t_{i+1}|^\alpha} + \frac{|x(t_{i+1})-x(t_j)|}{|t_{i+1}-t_j|^\alpha},$$

because $|s-t| > |s-t_{i+1}| \vee |t_{i+1}-t_j|$. Now suppose that $i,j \in \{0,\ldots,k-1\}$ satisfy $i < j$, $s \in (t_i,t_{i+1})$ and $t \in (t_j,t_{j+1})$. In this case,

$$\frac{|x(s)-x(t)|}{|s-t|^\alpha} \leq \frac{|x(s)-x(t_{i+1})|}{|s-t_{i+1}|^\alpha} + \frac{|x(t_{i+1})-x(t_j)|}{|t_{i+1}-t_j|^\alpha} + \frac{|x(t_j)-x(t)|}{|t_j-t|^\alpha},$$

since $|s-t| > |s-t_{i+1}| \vee |t_{i+1}-t_j| \vee |t_j-t|$. Now the assertion follows. □

This yields the before mentioned criterion.

Proposition 14. *Let $({}_nU)_{n\in\mathbb{N}}$ be a sequence of $\mathbb{R}^m$-valued right-continuous processes and assume there are $p,q > 0$ so that for each $\beta \in (0,q/(2p))$ there is $c_\beta \geq 0$ satisfying*

$$P\left(\max_{j\in\{0,\ldots,k_n-1\}} \sup_{u,v\in[t_{j,n},t_{j+1,n}]} \frac{|{}_nU_u - {}_nU_v|}{|u-v|^\beta} \geq \lambda\right) \leq c_\beta \lambda^{-2p} \tag{3.7}$$

for all $n \in \mathbb{N}$ and $\lambda > 0$. If $(\|{}_nU^r\|)_{n\in\mathbb{N}}$ and $(\max_{j\in\{0\ldots,k_n\}} |{}_nU_{t_{j,n}}|/|\mathbb{T}_n|^\alpha)_{n\in\mathbb{N}}$ converge in probability to zero, then so does the sequence $(\|{}_nU\|_{\alpha,r})_{n\in\mathbb{N}}$ for all $\alpha \in [0,q/(2p))$.

Proof. Let $n \in \mathbb{N}$ and fix $\beta \in (\alpha, q/(2p))$, then it follows that

$$\|_nU\| \leq \max_{j\in\{0,\dots,k_n-1\}} \sup_{t\in(t_{j,n},t_{j+1,n})} \frac{|_nU_t - {}_nU_{t_{j,n}}|}{|t-t_{j,n}|^\beta}|\mathbb{T}_n|^\beta + \max_{j\in\{0,\dots,k_n-1\}} |_nU_{t_{j,n}}|,$$

since $|_nU_t| \leq (|_nU_t - {}_nU_{t_{j,n}}|/|t-t_{j,n}|^\beta)|\mathbb{T}_n|^\beta + |_nU_{t_{j,n}}|$ for all $j \in \{0,\dots,k_n-1\}$ and $t \in (t_{j,n}, t_{j+1,n})$. Consequently, from (3.7) we obtain that

$$\mathbb{P}\Big(\|_nU\| \geq \varepsilon\Big) \leq c_\beta(2/\varepsilon)^{2p}|\mathbb{T}_n|^{2\beta p} + \mathbb{P}\Big(\max_{j\in\{0,\dots,k_n\}} |_nU_{t_{j,n}}| \geq \varepsilon/2\Big)$$

for all $\varepsilon > 0$, which directly entails that $(\|_nU\|)_{n\in\mathbb{N}}$ converges in probability to zero. Next, for fixed $n \in \mathbb{N}$ Lemma 13 gives us that

$$\sup_{s,t\in[r,T]:\, s\neq t} \frac{|_nU_s - {}_nU_t|}{|s-t|^\alpha} \leq 2 \max_{j\in\{0,\dots,k_n-1\}} \sup_{u,v\in[t_{j,n},t_{j+1,n}]:\, u\neq v} \frac{|_nU_u - {}_nU_v|}{|u-v|^\alpha} + \max_{i,j\in\{0,\dots,k_n\}:\, i\neq j} \frac{|_nU_{t_{i,n}} - {}_nU_{t_{j,n}}|}{|t_{i,n}-t_{j,n}|^\alpha}.$$

By using the facts that $|u-v|^{\beta-\alpha} \leq |\mathbb{T}_n|^{\beta-\alpha}$ 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 $u, v \in [t_{j,n}, t_{j+1,n}]$, we see that

$$P\left(\sup_{s,t\in[r,T]:\, s\neq t} \frac{|_nU_s - {}_nU_t|}{|s-t|^\alpha} \geq \varepsilon\right) \leq c_\beta(4/\varepsilon)^{2p}|\mathbb{T}_n|^{(\beta-\alpha)2p} + P\left(2c_\mathbb{T}^\alpha \max_{j\in\{0,\dots,k_n\}} |_nU_{t_{j,n}}|/|\mathbb{T}_n|^\alpha \geq \varepsilon/2\right)$$

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

Remark 15. If $p \geq 1/2$ and there is $c_0 \geq 0$ such that $E[|_nU_u - {}_nU_v|^{2p}] \leq c_0|u-v|^{1+q}$ for all $n \in \mathbb{N}$ and $u, v \in [r,T]$, then Proposition 12 and Chebyshev's inequality assure that condition (3.7) is satisfied.

3.4 Adapted linear interpolation of Brownian motion

We study the sequence $(_nW)_{n\in\mathbb{N}}$ of adapted linear interpolations of W given by (2.9) and for which a.e. path lies 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

$$\underline{t}_n := t_{(i-1)\vee 0,n}, \quad t_n := t_{i,n} \quad \text{and} \quad \bar{t}_n := t_{i+1,n}. \tag{3.8}$$

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 $\overline{T}_n := T$. In addition, we use the following abbreviations:

$$\Delta t_{i,n} := (t_{i,n} - t_{i-1,n}) \quad \text{and} \quad \Delta W_{t_{i,n}} := W_{t_{i,n}} - W_{t_{i-1,n}} \tag{3.9}$$

for each $i \in \{1, \ldots, k_n - 1\}$. After these preparations, let us begin with a general integral representation.

Lemma 16. *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 square-integrable process Y satisfies*

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

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

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

if $i, j \in \{1, \ldots, k_n - 1\}$ satisfy $i < j$, $s \in [t_{i,n}, t_{i+1,n}]$ and $t \in [t_{j,n}, t_{j+1,n}]$. In particular, for all $j \in \{1, \ldots, k_n\}$ we have

$$\int_r^{t_{j,n}} Y_{\underline{u}_n} \, d_n W_u = \int_r^{t_{j-1,n}} Y_{u_n} \, dW_u \quad a.s.$$

Proof. As ${}_nW_u = 0$ for each $u \in [r, t_{1,n}]$, the second claim follows from the first, by choosing $s = t_{1,n}$ and $t = t_{j,n}$ for $j \in \{1, \ldots, k_n\}$.

To check the first claim, suppose initially that $s, t \in [t_{i,n}, t_{i+1,n}]$ for some $i \in \{1, \ldots, k_n - 1\}$, then

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

Now assume instead that there are $i, j \in \{1, \ldots, k_n - 1\}$ such that $i < j$, $s \in [t_{i,n}, t_{i+1,n}]$ and $t \in [t_{j,n}, t_{j+1,n}]$. Then the a.s. decomposition

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

and the case considered above imply the asserted representation. □

Let us recall that for each $p \geq 1$ there is a constant $w_p > 0$ depending only on p such that for every $\mathbb{R}^{m\times d}$-valued progressively measurable process Y it holds that

$$E\bigg[\sup_{v\in[s,t]}\bigg|\int_s^v Y_u\,dW_u\bigg|^{2p}\bigg] \leq w_p E\bigg[\bigg(\int_s^t |Y_u|^2\,du\bigg)^p\bigg] \tag{M}$$

for all $s,t \in [r,T]$ with $s \leq t$. In fact, $w_p = 2^p p^{3p}/(p-1/2)^p$ and the dimensions m and d do not alter w_p. We derive a corresponding result for the sequence $({}_nW)_{n\in\mathbb{N}}$ of adapted linear interpolations of W.

Proposition 17. *For each $p \geq 1$ there is a constant $\hat{w}_p > 0$ such that every $\mathbb{R}^{m\times d}$-valued progressively measurable process Y satisfies*

$$E\bigg[\sup_{v\in[s,t]}\bigg|\int_s^v Y_{\underline{u}_n} d_nW_u\bigg|^{2p}\bigg] \leq \hat{w}_p \max_{j\in\{0,\dots,k_n\}:\, t_{j,n}\in[\underline{s}_n,\underline{t}_n]} E\Big[|Y_{t_{j,n}}|^{2p}\Big](t-s)^p \tag{3.10}$$

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

Proof. First, if $t \leq t_{1,n}$, then $\int_s^v Y_{\underline{u}_n}\,d_nW_u = 0$ for each $v \in [s,t]$ a.s. For $s < t_{1,n}$ and $t \geq t_{1,n}$ we have

$$\int_s^v Y_{\underline{u}_n}\,d_nW_u = \int_{t_{1,n}}^v Y_{\underline{u}_n}\,d_nW_u \quad \text{for all } v \in [t_{1,n},t] \text{ a.s.}$$

Thus, let us use Lemma 16 and assume that $s,t \in [t_{i,n},t_{i+1,n}]$ for some $i \in \{1,\dots,k_n-1\}$, then (M) yields that

$$\begin{aligned}E\bigg[\sup_{v\in[s,t]}\bigg|\int_s^v Y_{\underline{u}_n}\,d_nW_u\bigg|^{2p}\bigg] &= \frac{(t-s)^{2p}}{(\Delta t_{i+1,n})^{2p}}E\bigg[\bigg|\int_{t_{i-1,n}}^{t_{i,n}} Y_{u_n}\,dW_u\bigg|^{2p}\bigg]\\ &\leq w_p c_{\mathbb{T}}^p E\Big[|Y_{t_{i-1,n}}|^{2p}\Big](t-s)^p,\end{aligned}$$

where $c_{\mathbb{T}}$ is the constant appearing in (2.8). Next, suppose that there are $i,j \in \{1,\dots,k_n-1\}$ such that $i<j$, $s \in [t_{i,n},t_{i+1,n}]$ and $t \in [t_{j,n},t_{j+1,n}]$. Then

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

This is due to Lemma 16, which asserts that the process $[s,t]\times\Omega \to \mathbb{R}^m$, $(v,\omega) \mapsto \int_s^v Y_{\underline{u}_n}(\omega)\,d_nW_u(\omega)$ is piecewise linear. Hence, we obtain that

$$E\bigg[\sup_{v\in[s,t]}\bigg|\int_s^v Y_{\underline{u}_n}\,d_nW_u\bigg|^{2p}\bigg] \leq \hat{w}_p \max_{h\in\{i-1,\dots,j-1\}} E\Big[|Y_{t_{h,n}}|^{2p}\Big](t-s)^p$$

for $\hat{w}_p := 3^{2p-1}w_p c_{\mathbb{T}}^p$, which yields the claim. □

Lemma 18. *For each $p, q \geq 1$ there exists a constant $\hat{w}_{p,q} > 0$ satisfying*

$$E\bigg[\bigg(\int_s^t |_n\dot{W}_u|^q\, du\bigg)^p\bigg] \leq \hat{w}_{p,q}|\mathbb{T}_n|^{-qp/2}(t-s)^p \tag{3.11}$$

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

Proof. Clearly, if $t \leq t_{1,n}$, then $\int_s^t |_n\dot{W}_u|^q\, du = 0$. For $s < t_{1,n}$ and $t \geq t_{1,n}$ we have

$$\int_s^t |_n\dot{W}_u|^q\, du = \int_{t_{1,n}}^t |_n\dot{W}_u|^q\, du.$$

So, let now $s, t \in [t_{i,n}, t_{i+1,n}]$ 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

$$E\bigg[\bigg(\int_s^t |_n\dot{W}_u|^q\, du\bigg)^p\bigg] = E\Big[|Z|^{qp}\Big]\frac{(\Delta t_{i,n})^{qp/2}}{(\Delta t_{i+1,n})^{qp}}(t-s)^p \leq c|\mathbb{T}_n|^{-qp/2}(t-s)^p,$$

where $\hat{w}_{p,q} := E[|Z|^{qp}]c_{\mathbb{T}}^{qp}$ and $c_{\mathbb{T}}$ is the constant in (2.8). Next, let instead $i, j \in \{1, \dots, k_n - 1\}$ 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

$$\begin{aligned}\bigg(E\bigg[\bigg(\int_s^t |_n\dot{W}_u|^q\, du\bigg)^p\bigg]\bigg)^{1/p} &\leq \sum_{h=i}^{j}\bigg(E\bigg[\bigg(\int_{s\vee t_{h,n}}^{t\wedge t_{h+1,n}} |_n\dot{W}_u|^q\, du\bigg)^p\bigg]\bigg)^{1/p} \\ &\leq \hat{w}_{p,q}^{1/p}|\mathbb{T}_n|^{-q/2}(t-s),\end{aligned}$$

by what we have just shown. Therefore, the claim holds. ☐

3.5 Auxiliary convergence results

Lemma 19. *Let $({}_nU)_{n\in\mathbb{N}}$ be a sequence of non-negative measurable processes for which there are $p > 1$ and $c_p > 0$ such that $E[{}_nU_s^{2p}] \leq c_p|\mathbb{T}_n|^{2p}$ for each $s \in [r, T)$ and $n \in \mathbb{N}$. Then there is $c_1 > 0$ satisfying*

$$E\bigg[\bigg(\int_r^T {}_nU_s|_n\dot{W}_s|\, ds\bigg)^2\bigg] \leq c_1|\mathbb{T}_n| \quad \textit{for all } n \in \mathbb{N}.$$

Proof. Let $q > 1$ be such that $1/p + 1/q = 1$, then Lemma 18 gives a constant $\hat{w}_{q,2} > 0$ that is independent of n such that

$$E\bigg[\bigg(\int_r^T |_n\dot{W}_s|^2\, ds\bigg)^q\bigg] \leq \hat{w}_{q,2}|\mathbb{T}_n|^{-q}(T-r)^q.$$

Thus, we define $c_{1,1} := \hat{w}_{q,2}^{1/q}(T-r)$ and $c_1 := c_p^{1/p}(T-r)c_{1,1}$, then it follows from Cauchy-Schwarz's and Hölder's inequalities that

$$E\bigg[\bigg(\int_r^T {}_nU_s|_n\dot{W}_s|\,ds\bigg)^2\bigg] \leq \bigg(E\bigg[\bigg(\int_r^T {}_nU_s^2\,ds\bigg)^p\bigg]\bigg)^{1/p} c_{1,1}|\mathbb{T}_n|^{-1} \leq c_1|\mathbb{T}_n|.$$

□

Lemma 20. *Let $n \in \mathbb{N}$ and for every right-continuous map $x : [0,T] \to \mathbb{R}^m$ set $L_n(x)(t) := x(r \wedge t)$ for all $t \in [0, t_{1,n})$,*

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

for $t \in [t_{i,n}, t_{i+1,n})$ with $i \in \{1, \ldots, k_n - 1\}$ and $L_n(x)(T) := x(t_{k_n-1,n})$. Then it holds that $\|L_n(x)^t\| \leq \|x^r\| \vee \max_{j\in\{1,\ldots,k_n-1\}:\, t_{j,n}\leq t} |x(t_{j,n})|$ and

$$\|L_n(x)^t - x^t\| \leq 2 \max_{j\in\{0,\ldots,k_n-1\}:\, t_{j,n}\leq t}\ \sup_{s\in[t_{j,n},t_{j+1,n}]} |x^t(s) - x(t_{(j-1)\vee 0,n})|$$

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

Proof. Fix $s \in [t_{1,n}, t]$ and let $i \in \{1, \ldots, k_n - 1\}$ be so 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 is linear on $[t_{i,n}, t_{i+1,n}]$. In addition, we immediately obtain that

$$\begin{aligned}|L_n(x)(s) - x(s)| &\leq |x(s) - x(t_{i-1,n})| + |x(t_{i,n}) - x(t_{i-1,n})|\\ &\leq 2 \sup_{v\in[t_{i,n},t_{i+1,n}]} |x^t(v) - x(t_{i-1,n})|\end{aligned}$$

and the assertions follow. □

Lemma 21. *Let $({}_nU)_{n\in\mathbb{N}}$ denote a sequence of $\mathbb{R}^m$-valued right-continuous processes for which there are $c_0 \geq 0$, $p \geq 1/2$ and $q > 0$ such that*

$$E\Big[|{}_nU_s - {}_nU_t|^{2p}\Big] \leq c_0|s-t|^{1+q}$$

for all $n \in \mathbb{N}$ and $s, t \in [r, T]$. Then for every $\alpha \in [0, q/(2p))$,

$$\lim_{n\uparrow\infty} E\Big[\|{}_nU - L_n({}_nU)\|^{2p}\Big]/|\mathbb{T}_n|^{2\alpha p} = 0.$$

Proof. Fix $\beta \in (\alpha, q/(2p))$. From Lemma 20 we obtain that

$$\|{}_nU - L_n({}_nU)\| \leq 2^{1+\beta}|\mathbb{T}_n|^\beta \sup_{s,t\in[r,T]:\, s\neq t} \frac{|{}_nU_s - {}_nU_t|}{|s-t|^\beta}$$

for every $n \in \mathbb{N}$. Consequently, Proposition 12 implies that the constant $c_\beta := 2^{2(1+\beta)p} k_{\beta,q,p} c_0 (T-r)^{1+q-2\beta p}$ satisfies

$$E\Big[\|{}_nU - L_n({}_nU)\|^{2p}\Big] \leq c_\beta |\mathbb{T}_n|^{2\beta p} \quad \text{for all } n \in \mathbb{N},$$

where $k_{\beta,q,p}$ is given by (3.2) when α is replaced by β. Now the claim follows, since $\alpha < \beta$. □

Lemma 22. *Let $G : [r,T] \times S \to \mathbb{R}^m$ be d_∞-Lipschitz continuous, $({}_nU)_{n\in\mathbb{N}}$ be a sequence in $\mathscr{C}([0,T],\mathbb{R}^m)$ and $c_0 \geq 0$ be such that $|G(t,x)| \leq c_0(1+\|x\|)$ and*

$$E\Big[\|{}_nU\|^2\Big] + E\Big[\|{}_nU^s - {}_nU^t\|^2\Big]/|s-t| \leq c_0\Big(1 + E\Big[\|{}_nU^r\|^2\Big]\Big)$$

for all $n \in \mathbb{N}$, $s,t \in [r,T]$ with $s \neq t$ and $x \in S$. Then there is $c_1 > 0$ satisfying for each $n \in \mathbb{N}$,

$$E\bigg[\max_{j\in\{0,\dots,k_n\}} \bigg|\int_r^{t_{j,n}} G(\underline{s}_n, {}_nU^{\underline{s}_n})\bigg(\frac{\Delta s_n}{\Delta \overline{s}_n} - 1\bigg)\,ds\bigg|^2\bigg] \\ \leq c_1|\mathbb{T}_n|\Big(1 + E\Big[\|{}_nU^r\|^2\Big]\Big).$$

Proof. We assume that $E[\|{}_nU^r\|^2] < \infty$, as otherwise there is nothing to show. By decomposing the integral, we can rewrite that

$$\int_r^{t_{j,n}} G(\underline{s}_n, {}_nU^{\underline{s}_n})\frac{\Delta s_n}{\Delta \overline{s}_n}\,ds = \int_r^{t_{j-1,n}} G(s_n, {}_nU^{s_n})\,ds$$

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

$$E\bigg[\max_{j\in\{1,\dots,k_n\}} \bigg|\int_r^{t_{j-1,n}} G(s_n, {}_nU^{s_n}) - G(\underline{s}_n, {}_nU^{\underline{s}_n})\,ds\bigg|^2\bigg] \\ \leq c_{1,1}|\mathbb{T}_n|\Big(1 + E\Big[\|{}_nU^r\|^2\Big]\Big)$$

with $c_{1,1} := 2\lambda^2(T-r)^2(1+c_0)$. In addition, we estimate that

$$E\bigg[\max_{j\in\{1,\dots,k_n\}} \bigg|\int_{t_{j-1,n}}^{t_{j,n}} G(\underline{s}_n, {}_nU^{\underline{s}_n})\,ds\bigg|^2\bigg] \leq c_{2,2}|\mathbb{T}_n|^2\Big(1 + E\Big[\|{}_nU^r\|^2\Big]\Big),$$

where $c_{2,2} := 2c_0^2(1+c_0)$. So, the constant $c_1 := 2(c_{1,1} + c_{1,2}(T-r))$ yields the claim. □

Proposition 23. *Let $G : [r,T] \times S \to \mathbb{R}^{m\times d}$ be d_∞-Lipschitz continuous and $({}_nU)_{n\in\mathbb{N}}$ be a sequence in $\mathscr{C}([0,T],\mathbb{R}^m)$. Suppose there are $c_0 \geq 0$ and $p > 1$ such that $|G(t,x)| \leq c_0(1+\|x\|)$ and*

$$E\Big[\|{}_nU\|^{2p}\Big] + \frac{E\Big[\|{}_nU^s - {}_nU^t\|^{2p}\Big]}{|s-t|^p} \leq c_0$$

for each $n \in \mathbb{N}$, $s, t \in [r, T]$ such that $s \neq t$ and $x \in S$. Then for every $\alpha \in [0, 1/2 - 1/(2p))$,

$$\lim_{n\uparrow\infty} |\mathbb{T}_n|^{-2\alpha} E\bigg[\max_{j\in\{0,\ldots,k_n\}} \bigg|\int_r^{t_{j,n}} G(\underline{s}_n, {}_nU^{\underline{s}_n})\, d({}_nW_s - W_s)\bigg|^2\bigg] = 0.$$

Proof. We fix $n \in \mathbb{N}$ and once again decompose the integral to get that

$$\int_r^{t_{j,n}} G(\underline{s}_n, {}_nU^{\underline{s}_n})\, d_nW_s = \int_r^{t_{j-1,n}} G(s_n, {}_nU^{s_n})\, dW_s \quad \text{a.s.}$$

for all $j \in \{1, \ldots, k_n\}$. Let $\lambda \geq 0$ denote a Lipschitz constant for G, then

$$E\bigg[\max_{j\in\{0,\ldots,k_n\}} \bigg|\int_r^{t_{j-1,n}} G(s_n, {}_nU^{s_n}) - G(\underline{s}_n, {}_nU^{\underline{s}_n})\, dW_s\bigg|^2\bigg]$$
$$\leq 2w_1\lambda^2 \int_r^T (s_n - \underline{s}_n) + E\Big[\|{}_nU^{s_n} - {}_nU^{\underline{s}_n}\|^2\Big]\, ds \leq c_1|\mathbb{T}_n|$$

with $c_1 := 2w_1\lambda^2(T-r)(1+c_0^{1/p})$, where w_1 is the constant satisfying (M) for $p = 1$. Next, we let ${}_nM \in \mathscr{C}([0, T], \mathbb{R}^m)$ be a square-integrable martingale satisfying

$${}_nM_t = \int_r^t G(\underline{s}_n, {}_nU^{\underline{s}_n})\, dW_s$$

for all $t \in [r, T]$ a.s., then ${}_nM_{t_{j,n}} - {}_nM_{t_{j-1,n}} = \int_{t_{j-1,n}}^{t_{j,n}} G(\underline{s}_n, {}_nU^{\underline{s}_n})\, dW_s$ a.s. for each $j \in \{1, \ldots, k_n\}$. Furthermore,

$$E\Big[|{}_nM_s - {}_nM_t|^{2p}\Big] \leq 2^{2p-1}w_pc_0^{2p}(t-s)^p\Big(1 + E\Big[\|{}_nU\|^{2p}\Big]\Big) \leq c_2(t-s)^p$$

for all $s, t \in [r, T]$ with $s \leq t$, where $c_2 := 2^{2p-1}w_pc_0^{2p}(1 + c_0)$. Thus, let $\beta \in (\alpha, 1/2 - 1/(2p))$, then it follows from Proposition 12 that

$$E\bigg[\bigg(\sup_{s,t\in[r,T]:\, s\neq t} \frac{|{}_nM_s - {}_nM_t|}{|s-t|^\beta}\bigg)^{2p}\bigg] \leq k_{\beta,p-1,p}c_2(T-r)^{(1-2\beta)p},$$

since $p > 1$, by assumption. Finally, we set $c_\beta := (k_{\beta,p-1,p}c_2)^{1/p}(T-r)^{1-2\beta}$, then

$$E\bigg[\max_{j\in\{1,\ldots,k_n\}} \bigg|\int_{t_{j-1,n}}^{t_{j,n}} G(\underline{s}_n, {}_nU^{\underline{s}_n})\, dW_s\bigg|^2\bigg] \leq c_\beta|\mathbb{T}_n|^{2\beta},$$

by Hölder's inequality, and the assertion follows. □

4 Path-dependent ODEs and SDEs: proofs

In this section, we give the proof for

- the existence and uniqueness of mild solutions to path-dependent ODEs in Section 2.2 and
- the existence and uniqueness of strong solutions to path-dependent SDEs in Section 2.3.

4.1 Proof of Proposition 2

We first derive a global estimate for any mild solution. This allows us to use (O.ii), the Lipschitz condition on bounded sets, to derive existence and uniqueness results.

Lemma 24. *Under* (O.i)*, there is $c_H > 0$ depending only on $T - r$ such that any mild solution x to the ODE* (2.3) *satisfies for all $t \in [r, T]$,*

$$\|x^t\|_{H,r}^2 \leq c_H e^{c_H \int_r^t c_F(s)^2\, ds}\bigg(\|x^r\|^2 + \int_r^t c_F(s)^2\, ds\bigg). \tag{4.1}$$

Proof. By estimating $\|x^t\| + \int_r^t |\dot{x}(s)|\, ds$ for given $t \in [r, T]$, it follows readily from Gronwall's inequality that

$$\|x^t\| + \int_r^t |\dot{x}(s)|\, ds \leq e^{2\int_r^t c_F(s)\, ds}\bigg(\|x^r\| + 2\int_r^t c_F(s)\, ds\bigg).$$

Moreover, for $c_1 := 2^2(T - r + 1)$ we have

$$\|x^t\|_{H,r}^2 \leq 2\|x^r\|^2 + c_1 \int_r^t c_F(s)^2\bigg(1 + \|x^s\| + \int_r^s |\dot{x}(u)|\, du\bigg)^2 ds. \tag{4.2}$$

Thus, we set $c_H := 2e^2 c_1$, then from $2\int_r^t c_F(s)\, ds \leq 1 + (T - r)\int_r^t c_F(s)^2\, ds$ we infer that

$$\bigg(\|x^t\| + \int_r^t |\dot{x}(s)|\, ds\bigg)^2 \leq c_H e^{c_H \int_r^t c_F(s)^2\, ds}\bigg(\|x^r\|^2 + \int_r^t c_F(s)^2\, ds\bigg).$$

The claim follows from (4.2), the fundamental theorem of calculus for Riemann-Stieltjes integrals and the transitivity of absolutely continuous measures. □

We now show uniqueness of mild solutions, which implies uniqueness for classical solutions.

Lemma 25. *Assume that* (O.i) *and* (O.ii) *hold, then any two mild solutions* x *and* y *to the ODE* (2.3) *that satisfy* $x^r = y^r$ *must coincide.*

Proof. By Lemma 24, there is $n \in \mathbb{N}$ such that $\|x\|_{H,r} \vee \|y\|_{H,r} \leq n$. Thus,

$$\|x^t - y^t\|_{H,r}^2 \leq 2(T - r + 1) \int_r^t \lambda_{F,n}^2(s) \|x^s - y^s\|_{H,r}^2 \, ds$$

for all $t \in [r, T]$. Gronwall's inequality implies that $x = y$. □

Proof of Proposition 2. As the uniqueness claim follows from Lemma 25, we directly turn to the existence assertion. To this end, let $\mathscr{H}$ be the set of all $x \in H_r^1([0,T], \mathbb{R}^m)$ satisfying $x(s) = \hat{x}(s)$ for every $s \in [0, r]$ and the estimate (4.1), where c_H is chosen largely enough so that

$$c_H \geq 2^4 (T - r + 1)^2. \tag{4.3}$$

By Lemma 24, a map $x \in S$ is a mild solution to the ODE (2.3) such that $x(s) = \hat{x}(s)$ for all $s \in [0, r]$ if and only if $x \in \mathscr{H}$ and it is a fixed-point of the operator $\Psi : \mathscr{H} \to H_r^1([0,T], \mathbb{R}^m)$ given by

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

We remark that condition (4.3) assures that Ψ maps $\mathscr{H}$ into itself. Indeed, this follows by inserting (4.1) into the inequality

$$\|\Psi(x)^t\|_{H,r}^2 \leq c_H \|x_0\|^2 + c_H \int_r^t c_F(s)^2 (1 + \|x^s\|_{H,r}^2) \, ds,$$

valid for all $x \in \mathscr{H}$ and $t \in [r, T]$. As $x_0 \in \mathscr{H}$ and $x_n = \Psi(x_{n-1})$ for each $n \in \mathbb{N}$, by definition (2.4), we now know that $(x_n)_{n \in \mathbb{N}_0}$ is a sequence in $\mathscr{H}$.

Next, let us choose $l \in \mathbb{N}$ satisfying $\|x\|_{H,r} \leq l$ for all $x \in \mathscr{H}$ and set $c_1 := 2(T - r + 1)$. Then we obtain that

$$\|\Psi(x)^t - \Psi(y)^t\|_{H,r}^2 \leq c_1 \int_r^t \lambda_{F,l}(s)^2 \|x^s - y^s\|_{H,r}^2 \, ds$$

for each $x, y \in \mathscr{H}$ and $t \in [r, T]$, which in particular shows that Ψ must be $\|\cdot\|_{H,r}$-Lipschitz continuous. Moreover, it follows inductively that

$$\|x_{n+1}^t - x_n^t\|_{H,r}^2 \leq \frac{\delta^2}{n!} \left(c_1 \int_r^t \lambda_{F,l}(s)^2 \, ds \right)^n$$

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

$$\|x_n - x_k\|_{H,r} \leq \delta \sum_{i=k}^{n-1} \left(\frac{1}{i!} \right)^{1/2} \left(c_1 \int_r^T \lambda_{F,l}(s)^2 \, ds \right)^{i/2}$$

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

As $\mathscr{H}$ is closed with respect to the complete norm $\|\cdot\|_{H,r}$, there exists a unique map $y_F \in \mathscr{H}$ such that $\lim_{n\uparrow\infty}\|x_n - y_F\|_{H,r} = 0$. Lipschitz continuity of Ψ implies $\lim_{n\uparrow\infty}\|x_{n+1} - \Psi(y_F)\|_{H,r} = 0$. For this reason, $y_F = \Psi(y_F)$ and the proposition is proven. □

4.2 Proof of Proposition 5

Lemma 26. *Let X be an $\mathbb{R}^m$-valued adapted right-continuous process and $B \subset \mathbb{R}^m$ be closed, then*

$$\tau := \inf\{t \in [0,T] \,|\, \overline{\{X_s \,|\, s \in [0,t]\}} \cap B \neq \emptyset\}$$

is a stopping time satisfying $\tau = \inf\{t \in [0,T] \,|\, X_t \in B\}$ on $\{X \in S\}$.

Proof. First, we check that $\{\tau \leq t\} = \{\overline{\{X_s \,|\, s \in [0,t]\}} \cap B \neq \emptyset\}$ for fixed $t \in [0,T]$. To this end, it suffices to show that if $t < T$ and $\omega \in \Omega$ satisfies $\tau(\omega) = t$, then $\overline{\{X_s(\omega) \,|\, s \in [0,t]\}} \cap B \neq \emptyset$.

In this case, for each $n \in \mathbb{N}$ there are $s_n \in [0, t + (T-t)/n)$ and $y_n \in B$ satisfying $|y_n - X_{s_n}(\omega)| < 1/n$. So, we choose a strictly increasing sequence $(\nu_n)_{n\in\mathbb{N}}$ in $\mathbb{N}$ such that $(s_{\nu_n})_{n\in\mathbb{N}}$ converges to some $s \in [0,t]$, then it follows that $X_s(\omega) = \lim_{n\uparrow\infty} X_{s_{\nu_n}}(\omega) \in B$, which yields the intermediate claim.

Next, we set $B_n := \{x \in B \,|\, |x| \leq n\}$ for all $n \in \mathbb{N}$ and use the notation $\operatorname{dist}(x, C) = \inf_{y\in C}|x - y|$ for all $x \in \mathbb{R}^m$ and $C \subset \mathbb{R}^m$. Let $t \in [0,T]$ and D be a countable dense set in $[0,t]$ containing t, then

$$\{\tau \leq t\} = \bigcup_{n\in\mathbb{N}}\{\inf_{s\in D}\operatorname{dist}(X_s, B_n) = 0\} = \bigcup_{n\in\mathbb{N}}\bigcap_{k\in\mathbb{N}}\bigcup_{s\in D}\{\operatorname{dist}(X_s, B_n) < 1/k\} \in \mathscr{F}_t,$$

by the above representation of $\{\tau \leq t\}$. Finally, if $\omega \in \Omega$ satisfies $X(\omega) \in S$, then $\{X_s(\omega) \,|\, s \in [0,t]\}$ is compact and in particular closed. For this reason, $\tau(\omega) < \inf\{t \in [0,T] \,|\, X_t(\omega) \in B\}$ would violate the definition of $\tau(\omega)$. □

Example 27. Let $X \in \mathscr{C}([0,T], \mathbb{R}^m)$ and $n \in \mathbb{N}$, then the above lemma gives a stopping time $\tau_n \geq r$ such that $\tau_n = \inf\{t \in [0,T] \,|\, |X_t| \geq n\} \vee r$ a.s. This ensures that

$$\|X^{t\wedge\tau_n}\| \leq \|X^r\| \vee n \quad \text{a.s.}$$

for all $t \in [r,T]$, since we have that $\|X^{t\wedge\tau_n}\| \leq n$ a.s. on $\{\|X^r\| \leq n\}$ and $\tau_n = r$ a.s. on $\{\|X^r\| > n\}$.

In this section, whenever $p \geq 1$ and condition (S.i) is satisfied, we set

$$m_p := \left(\int_r^T c_B(s)^2\,ds\right)^p + c_\Sigma^{2p} w_p,$$

where w_p is the constant appearing in (M).

Lemma 28. *Under* (S.i)*, for each* $p > 2$ *and* $\alpha \in (0, 1/2 - 1/p)$ *there is* $c_{\alpha,p} > 0$ *depending only on* α*,* p *and* $T - r$ *such that any strong solution* X *to* (2.5) *satisfies*

$$E\Big[\|X^t\|_{\alpha,r}^{2p}\Big] \leq c_{\alpha,p} e^{c_{\alpha,p} m_p (t-r)}\Big(E\Big[\|X^r\|^{2p}\Big] + m_p(t-r)\Big) \tag{4.4}$$

for all $t \in [r, T]$*.*

Proof. Assume that $E[\|X^r\|^{2p}] < \infty$ and let $n \in \mathbb{N}$. Then Example 27 yields a stopping time $\tau_n \geq r$ such that $\|X^{\tau_n}\| \leq \|X^r\| \vee n$ a.s. First,

$$\|X^{t\wedge\tau_n}\| \leq \|X^r\| + \int_r^{t\wedge\tau_n} |B(s, X^s)|\,ds + \sup_{v\in[r,t]}\left|\int_r^{v\wedge\tau_n} \Sigma(u, X^u)\,dW_u\right| \quad \text{a.s.}$$

for fixed $t \in [r, T]$. Thus, from Jensen's and Cauchy-Schwarz's inequality we obtain that

$$E\Big[\|X^{t\wedge\tau_n}\|^{2p}\Big] \leq 3^{2p-1} E\Big[\|X^r\|^{2p}\Big] + 6^{2p-1} m_p (t-r)^{p-1}\int_r^t 1 + E\Big[\|X^{s\wedge\tau_n}\|^{2p}\Big]\,ds.$$

Moreover, a similar computation shows that

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

for all $u, v \in [r, t]$ with $u < v$. Therefore, Proposition 12 yields that

$$E\left[\left(\sup_{u,v\in[r,t]:\,u\neq v}\frac{|X_u^{\tau_n} - X_v^{\tau_n}|}{|u-v|^\alpha}\right)^{2p}\right] \leq k_{\alpha,p-2,p} 4^{2p-1} m_p (t-r)^{p-1-2\alpha p}\int_r^t 1 + E\Big[\|X^{s\wedge\tau_n}\|^{2p}\Big]\,ds,$$

where the constant $k_{\alpha,p-2,p}$ is given by (3.2) for $q = p - 2$. Thus,

$$E\Big[\|X^{t\wedge\tau_n}\|_{\alpha,r}^2\Big]/c_{\alpha,p} \leq E\Big[\|X^r\|^{2p}\Big] + m_p\int_r^t 1 + E\Big[\|X^{s\wedge\tau_n}\|^{2p}\Big]\,ds$$

for $c_{\alpha,p} := 12^{2p-1}(1 + k_{\alpha,p-2,p})(T - r + 1)^{p-1}$. By Gronwall's inequality and Fatou's lemma,

$$E\Big[\|X^t\|_{\alpha,r}^{2p}\Big] \leq \liminf_{n\uparrow\infty} E\Big[\|X^{t\wedge\tau_n}\|_{\alpha,r}^{2p}\Big] \leq c_{\alpha,p} e^{c_{\alpha,p} m_p (t-r)}\Big(E\Big[\|X^r\|^{2p}\Big] + m_p(t-r)\Big),$$

which is the claim. □

Remark 29. If $p \geq 1$ and $\alpha \in (0, 1/2)$ are such that $\alpha < 1/2 - 1/p$ fails, then, under (S.i), we still have that $E[\|X^t\|_{\alpha,r}^{2p}] \leq (E[\|X^t\|_{\alpha,r}^{2q}])^{p/q} < \infty$ for any $q > p$ such that $\alpha < 1/2 - 1/q$, by Hölder's inequality.

Lemma 30. *Under* (S.iii)*, pathwise uniqueness holds for* (2.5)*.*

Proof. Let X and $\tilde{X}$ be two weak solutions to (2.5) defined on a common filtered probability space $(\tilde{\Omega}, \tilde{\mathscr{F}}, (\tilde{\mathscr{F}}_t)_{t\in[0,T]}, \tilde{P})$ on which there is a standard d-dimensional $(\tilde{\mathscr{F}}_t)_{t\in[0,T]}$-Brownian motion $\tilde{W}$ such that $X^r = \tilde{X}^r$ a.s.

We fix $n \in \mathbb{N}$, then it follows from Example 27 that there is a stopping time $\tau_n \geq r$ such that $\tau_n = \inf\{t \in [0,T] \,|\, |X_t| \geq n \text{ or } |\tilde{X}_t| \geq n\} \vee r$ a.s. Clearly,

$$\|X^{t\wedge\tau_n} - \tilde{X}^{t\wedge\tau_n}\| \leq \int_r^{t\wedge\tau_n} |B(s, X^s) - B(s, \tilde{X}^s)| \, ds + \sup_{v\in[r,t]} \left| \int_r^{v\wedge\tau_n} \Sigma(u, X^u) - \Sigma(u, \tilde{X}^u) \, d\tilde{W}_u \right| \quad \text{a.s.}$$

for given $t \in [r, T]$. We set $c_1 := ((T - r) + w_1)$, where w_1 is the constant in (M) for $p = 1$, then

$$\tilde{E}\Big[\|X^{t\wedge\tau_n} - \tilde{X}^{t\wedge\tau_n}\|^2\Big] \leq c_1 \int_r^t \lambda_n(s)^2 \tilde{E}\Big[\|X^{s\wedge\tau_n} - \tilde{X}^{s\wedge\tau_n}\|^2\Big] \, ds.$$

So, $X^{\tau_n} = \tilde{X}^{\tau_n}$ a.s., by Gronwall's inequality. As $\tau_n \leq \tau_{n+1}$ a.s. for all $n \in \mathbb{N}$ and $\sup_{n\in\mathbb{N}} \tau_n = \infty$ a.s., 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 implies that $X = \tilde{X}$ a.s. □

Proof of Proposition 5. We define $\mathscr{H}$ be the set of all $X \in \mathscr{C}_r([0,T], \mathbb{R}^m)$ satisfying $X_s = \hat{X}_s$ for all $s \in [0, r]$ a.s. and the estimate (4.4) for any $p > 2$ and $\alpha \in (0, 1/2 - 1/p)$, where the constant $c_{\alpha,p}$ is chosen largely enough so that

$$c_{\alpha,p} \geq 12^{2p-1}(1 + k_{\alpha,p-2,p})(T - r + 1)^{p-1}. \tag{4.5}$$

By Lemma 28 and Remark 29, we have that $\mathscr{H} \subset \mathscr{C}_{r,\infty}^{1/2-}([0,T], \mathbb{R}^m)$ and a process $X \in \mathscr{C}([0,T], \mathbb{R}^m)$ is a solution to (2.5) satisfying $X_s = \hat{X}_s$ for all $s \in [0, r]$ a.s. if and only if $X \in \mathscr{H}$ and it is a fixed point of the operator $\Psi : \mathscr{H} \to \mathscr{C}([0,T], \mathbb{R}^m)$ specified by requiring that

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

for all $t \in [0, T]$ a.s. We stress the fact that, due to Proposition 12 and condition (4.5), for every $X \in \mathscr{H}$, $p > 2$ and $\alpha \in (0, 1/2 - 1/p)$ it follows that

$$E\Big[\|\Psi(X)^t\|_{\alpha,r}^{2p}\Big] \leq c_{\alpha,p} E\Big[\|{}_0X\|^{2p}\Big] + c_{\alpha,p} m_p \int_r^t 1 + E\Big[\|X^s\|^{2p}\Big] \, ds$$

for all $t \in [r,T]$. Thus, $\Psi(\mathscr{H}) \subset \mathscr{H}$ follows from plugging (4.4) into the above inequality. Since ${}_0X \in \mathscr{H}$ and ${}_nX = \Psi({}_{n-1}X)$ a.s. for all $n \in \mathbb{N}$, by (2.7), we have shown that $({}_nX)_{n\in\mathbb{N}_0}$ is a sequence in $\mathscr{H}$.

Next, choose $p > 2$ and $\alpha \in (0,1/2)$ such that $\alpha_0 \leq \alpha < 1/2 - 1/p$, where α_0 is the constant in the Lipschitz condition (S.ii). Further, we set $l_p := (\int_r^T \lambda_B(s)^2\,ds)^p + \lambda_\Sigma^{2p} w_p$, then

$$E\Big[\|\Psi(X)^t - \Psi(Y)^t\|^{2p}\Big] \leq 2^{2p-1} l_p (t-r)^{p-1} \int_r^t E\Big[\|X^s - Y^s\|_{\alpha_0,r}^{2p}\Big]\,ds$$

for all given $X, Y \in \mathscr{H}$ and $t \in [r,T]$. After applying Proposition 12 and using $\|x\|_{\alpha_0,r} \leq (T-r+1)^{\alpha-\alpha_0}\|x\|_{\alpha,r}$ for all $x \in C_r^\alpha([0,T],\mathbb{R}^m)$, we get

$$E\Big[\|\Psi(X)^t - \Psi(Y)^t\|_{\alpha,r}^{2p}\Big] \leq \overline{c}_{\alpha,p} l_p \int_r^t E\Big[\|X^s - Y^s\|_{\alpha,r}^{2p}\Big]\,ds$$

with $\overline{c}_{\alpha,p} := 4^{2p-1}(1 + k_{\alpha,p-2,p})(T-r+1)^{2p-1}$. Hence, Gronwall's inequality entails that there is at most a unique solution X to (2.5) satisfying $X_s = \hat{X}_s$ for all $s \in [0,r]$ a.s.

We also infer from the above inequality that Ψ is Lipschitz continuous with respect to the seminorm (2.6), where p is replaced by $2p$. In addition,

$$E\Big[\|{}_{n+1}X^t - {}_nX^t\|_{\alpha,r}^{2p}\Big] \leq \frac{\delta^{2p}}{n!}(\overline{c}_{\alpha,p} l_p)^n (t-r)^n$$

for each $n \in \mathbb{N}_0$, by induction with $\delta := (E[\|\Psi({}_0X) - {}_0X\|_{\alpha,r}^{2p}])^{1/(2p)}$. Hence, the triangle inequality gives

$$\Big(E\Big[\|{}_nX - {}_kX\|_{\alpha,r}^{2p}\Big]\Big)^{\frac{1}{2p}} \leq \delta \sum_{i=k}^{n-1} \left(\frac{1}{i!}\right)^{\frac{1}{2p}} (\overline{c}_{\alpha,p} l_p)^{\frac{i}{2p}} (T-r)^{\frac{i}{2p}}$$

for each $k, n \in \mathbb{N}_0$ with $k < n$. The ratio test implies that the series $\sum_{i=0}^\infty (1/i!)^{1/(2p)} x^{i/(2p)}$ converges absolutely for each $x \geq 0$. So,

$$\lim_{k\uparrow\infty} \sup_{n\in\mathbb{N}:\, n\geq k} E\Big[\|{}_nX - {}_kX\|_{\alpha,r}^{2p}\Big] = 0.$$

Due to Proposition 11, because $\mathscr{H}$ is closed with respect to the complete seminorm (2.6), where p is replaced by $2p$, there exists a process $X \in \mathscr{H}$ that is unique up to indistinguishability such that

$$\lim_{n\uparrow\infty} E\Big[\|{}_nX - X\|_{\alpha,r}^{2p}\Big] = 0. \tag{4.6}$$

Lipschitz continuity of Ψ implies that $\lim_{n\uparrow\infty} E[\|_{n+1}X - \Psi(X)\|_{\alpha,r}^{2p}] = 0$. For this reason, $X = \Psi(X)$ a.s. Finally, assume $p \geq 1$ and $\alpha \in (0, 1/2)$ are such that $\alpha_0 \leq \alpha < 1/2 - 1/p$ fails. If $\alpha < \alpha_0$, then

$$E\Big[\|_nX - X\|_{\alpha,r}^{2p}\Big] \leq (T - r + 1)^{(\alpha_0-\alpha)2p} E\Big[\|_nX - X\|_{\alpha_0,r}^{2p}\Big]$$

for all $n \in \mathbb{N}$, which implies (4.6). For $\alpha \geq 1/2 - 1/p$ we take $q > p$ so that $\alpha < 1/2 - 1/q$ and use that $E[\|_nX - X\|_{\alpha,r}^{2p}] \leq (E[\|_nX - X\|_{\alpha,r}^{2q}])^{p/q}$ for all $n \in \mathbb{N}$. As this also gives (4.6), the proof is complete. □

5 Proof of main result

5.1 Decomposition into remainder terms

Proposition 31. *Let* (C.i) *hold,* $h \in H_r^1([0,T], \mathbb{R}^d)$ *and* $\overline{B}$ *be* d_∞*-Lipschitz continuous. Then for each* $p \geq 1$ *there is* $c_p > 0$ *such that any* $n \in \mathbb{N}$ *and any strong solution* $_nY$ *to* (2.14) *satisfy*

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

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

Proof. We let $l \in \mathbb{N}$ and use Example 27 to define a stopping time $\tau_{l,n} \geq r$ such that $\|_nY^{\tau_{l,n}}\| \leq \|_nY^r\| \vee l$ a.s. Further, we estimate that

$$\begin{aligned}\|_nY^{s\wedge\tau_{l,n}} - {}_nY^{t\wedge\tau_{l,n}}\| \leq &\int_s^{t\wedge\tau_{l,n}} |\underline{B}(u, {}_nY^u)| + |B_H(u, {}_nY^u)\dot{h}(u)|\,du \\ &+ \sup_{v\in[s,t]}\left|\int_s^{v\wedge\tau_{l,n}} \overline{B}(u, {}_nY^u)_n\dot{W}_u\,du\right| \\ &+ \sup_{v\in[s,t]}\left|\int_s^{v\wedge\tau_{l,n}} \Sigma(u, {}_nY^u)\,dW_u\right| \quad \text{a.s.}\end{aligned}$$

for fixed $s, t \in [r, T]$ with $s \leq t$. Thus, the triangle inequality and the inequalities of Cauchy-Schwarz and Jensen yield that

$$\begin{aligned}&\Big(E\Big[\|_nY^{s\wedge\tau_{l,n}} - {}_nY^{t\wedge\tau_{l,n}}\|^{2p}\Big]\Big)^{\frac{1}{2p}} \\ &\leq \left(c_{p,1}(t-s)^{p-1}\int_s^t 1 + E\Big[\|_nY^{u\wedge\tau_{l,n}}\|^{2\kappa p}\Big]\,du\right)^{\frac{1}{2p}} \\ &\quad + \left(E\left[\sup_{v\in[s,t]}\left|\int_s^{v\wedge\tau_{l,n}} \overline{B}(u, {}_nY^u)_n\dot{W}_u\,du\right|^{2p}\right]\right)^{\frac{1}{2p}},\end{aligned} \tag{5.2}$$

where we have set $c_{p,1} := (6c)^{2p}((T-r)^p + (\int_r^T |\dot{h}(u)|^2\,du)^p + w_p)$ and w_p is the constant appearing in (M).

Since $\kappa < 1$, we can pick $\gamma \in (1, \kappa^{-1})$, then Lemma 18 provides a constant $c_{p,2} > 0$ such that (3.11) holds when q and p are replaced by 2 and $p/(1-\gamma\kappa)$, respectively. This yields that

$$
\begin{aligned}
&E\bigg[\bigg(\int_s^{t\wedge\tau_{l,n}} \Big|\Big(\overline{B}(u, {}_nY^u) - \overline{B}(\underline{u}_n, {}_nY^{\underline{u}_n})\Big)_n\dot{W}_u\Big|\,du\bigg)^{2p}\bigg] \\
&\leq (2\lambda)^{2p}/2(t-s)^{p-1}\bigg(\int_s^t (u-\underline{u}_n)^p\,du\bigg) c_{p,2}^{1-\gamma\kappa}|\mathbb{T}_n|^{-p}(t-s)^p \\
&\quad + (2\lambda)^{2p}/2(t-s)^{p-1}\int_s^t E\bigg[\|{}_nY^{u\wedge\tau_{l,n}} - {}_nY^{\underline{u}_n\wedge\tau_{l,n}}\|^{2p}\bigg(\int_s^t |{}_n\dot{W}_v|^2\,dv\bigg)^p\bigg]\,du \\
&\leq c_{p,3}(t-s)^{p-1}\int_s^t 1 + |\mathbb{T}_n|^{-p}\Big(E\Big[\|{}_nY^{u\wedge\tau_{l,n}} - {}_nY^{\underline{u}_n\wedge\tau_{l,n}}\|^{\frac{2p}{\gamma\kappa}}\Big]\Big)^{\gamma\kappa}\,du,
\end{aligned}
$$

by the inequalities of Cauchy-Schwarz, Jensen and Hölder , where $\lambda \geq 0$ denotes a Lipschitz constant for $\overline{B}$ and $c_{p,3} := 2^{3p}\lambda^{2p}(T-r)^p c_{p,2}^{1-\gamma\kappa}$. Note here that the choice of $c_{p,2}$ entails that

$$
\bigg(E\bigg[\int_s^t |{}_n\dot{W}_v|^2\,dv\bigg)^{\frac{p}{1-\gamma\kappa}}\bigg]\bigg)^{1-\gamma\kappa} \leq c_{p,2}^{1-\gamma\kappa}|\mathbb{T}_n|^{-p}(t-s)^p.
$$

Next, let $c_{p,4} > 0$ be a constant satisfying (3.11) when q and p are replaced by 2 and $p/((\gamma-1)\kappa)$, respectively. Then Cauchy-Schwarz's, Jensen's and Hölder's inequality imply that

$$
\begin{aligned}
&E\bigg[\bigg(\int_{\underline{u}_n}^{u\wedge\tau_{l,n}} |\overline{B}(v, {}_nY^v)_n\dot{W}_v|\,dv\bigg)^{\frac{2p}{\gamma\kappa}}\bigg] \\
&\leq (2c)^{\frac{2p}{\gamma\kappa}}2^{-1}(u-\underline{u}_n)^{\frac{p}{\gamma\kappa}-1}\bigg(\int_{\underline{u}_n}^u 1\,dv\bigg) c_{p,4}^{\frac{\gamma-1}{\gamma}}|\mathbb{T}_n|^{-\frac{p}{\gamma\kappa}}(u-\underline{u}_n)^{\frac{p}{\gamma\kappa}} \\
&\quad + (2c)^{\frac{2p}{\gamma\kappa}}2^{-1}(u-\underline{u}_n)^{\frac{p}{\gamma\kappa}-1}\int_{\underline{u}_n}^u E\bigg[\|{}_nY^{v\wedge\tau_{l,n}}\|^{\frac{2p}{\gamma}}\bigg(\int_{\underline{u}_n}^u |{}_n\dot{W}_w|^2\,dw\bigg)^{\frac{p}{\gamma\kappa}}\bigg]\,dv \\
&\leq c_{p,5}|\mathbb{T}_n|^{\frac{p}{\gamma\kappa}}\Big(1 + \Big(E\Big[\|{}_nY^{u\wedge\tau_{l,n}}\|^{2p}\Big]\Big)^{\frac{1}{\gamma}}\Big)
\end{aligned}
$$

for any given $u \in [s,T]$, where we have set $c_{p,5} := (4c)^{(2p)/(\gamma\kappa)}c_{p,4}^{(\gamma-1)/\gamma}$, since by the choice of $c_{p,4}$ we can utilize that

$$
\bigg(E\bigg[\bigg(\int_{\underline{u}_n}^u |{}_n\dot{W}_v|^2\,dv\bigg)^{\frac{p}{(\gamma-1)\kappa}}\bigg]\bigg)^{\frac{\gamma-1}{\gamma}} \leq c_{p,4}^{\frac{\gamma-1}{\gamma}}|\mathbb{T}_n|^{-\frac{p}{\gamma\kappa}}(u-\underline{u}_n)^{\frac{p}{\gamma\kappa}}.
$$

Thus, by the virtue of (5.2), we may set $c_{p,6} := (6c)^{(2p)/(\gamma\kappa)}((T-r)^{p/(\gamma\kappa)} + (\int_r^T |\dot{h}(u)|^2\,du)^{p/(\gamma\kappa)} + c_{p/(\gamma\kappa),M})$, then

$$\Big(E\Big[\|{}_nY^{u\wedge\tau_{l,n}} - {}_nY^{\underline{u}_n\wedge\tau_{l,n}}\|^{\frac{2p}{\gamma\kappa}}\Big]\Big)^{\gamma\kappa} \leq c_{p,7}|\mathbb{T}_n|^p\Big(1 + E\Big[\|{}_nY^{u\wedge\tau_{l,n}}\|^{2p}\Big]\Big)$$

with $c_{p,7} := 2^{2p}(2^p c_{p,6} + c_{p,5}^{\gamma\kappa})$, since $x^\beta \leq (1+x)^\beta \leq 1+x$ for all $x \geq 0$ and $\beta \in [0,1]$. Thus, we have established that

$$E\Bigg[\bigg(\int_s^{t\wedge\tau_{l,n}} \Big|\big(\overline{B}(u,{}_nY^u) - \overline{B}(\underline{u}_n,{}_nY^{\underline{u}_n})\big)_n\dot{W}_u\Big|\,du\bigg)^{2p}\Bigg] \leq c_{p,8}(t-s)^{p-1}\int_s^t 1 + E\Big[\|{}_nY^{u\wedge\tau_{l,n}}\|^{2p}\Big]\,du,$$

where $c_{p,8} := c_{p,3}(1+c_{p,7})$. Next, Proposition 17 gives a constant $\hat{w}_p > 0$ satisfying (3.10), which directly yields that

$$E\Bigg[\bigg|\int_s^{t\wedge\tau_{l,n}} \overline{B}(\underline{u}_n,{}_nY^{\underline{u}_n})_n\dot{W}_u\,du\bigg|^{2p}\Bigg] \leq \hat{w}_p c^{2p}(t-s)^p.$$

Hence, from (5.2) we in total obtain that

$$E\Big[\|{}_nY^{s\wedge\tau_{l,n}} - {}_nY^{t\wedge\tau_{l,n}}\|^{2p}\Big] \leq c_{p,9}(t-s)^{p-1}\int_s^t 1 + E\Big[\|{}_nY^{u\wedge\tau_{l,n}}\|^{2p}\Big]\,du$$

with $c_{p,9} := 4^{2p-1}(c_{p,1}+c_{p,8}+\hat{w}_p c^{2p})$. Thus, Gronwall's inequality and Fatou's lemma entail that

$$E\Big[\|{}_nY^t\|^{2p}\Big] \leq \liminf_{m\uparrow\infty} E\Big[\|{}_nY^{t\wedge\tau_{l,n}}\|^{2p}\Big] \leq c_{14}\Big(1 + E\Big[\|{}_nY^r\|^{2p}\Big]\Big)$$

with $c_{p,10} := e^{2^{2p}(1+c_{p,9}(T-r)^p)}$. Thus, by setting $c_p := (1+c_{p,9})(1+c_{p,10})$, the claim follows from an application of Fatou's lemma. □

Corollary 32. *Let* (C.i) *hold and* $h \in H_r^1([0,T],\mathbb{R}^d)$*. Then for each* $p \geq 1$ *there is* $c_p > 0$ *such that any strong solution* Z *to* (2.15) *satisfies*

$$E\Big[\|Z\|^{2p}\Big] + E\Big[\|Z^s - Z^t\|^{2p}\Big]/|s-t|^p \leq c_p\Big(1 + E\Big[\|Z^r\|^{2p}\Big]\Big) \tag{5.3}$$

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

Proof. Because the map R defined via (2.16) is bounded, we may apply Proposition 31 in the case that $\underline{B}$ is replaced by $\overline{B}+R$, $\overline{B}$ is replaced by 0 and Σ is replaced by $\overline{B}+\Sigma$. From this the claim follows immediately. □

Proposition 33. *Let* (C.i) *and* (C.ii) *be valid and* $h \in H_r^1([0,T],\mathbb{R}^d)$*. Then there is* $c_1 > 0$ *such that for any* $n \in \mathbb{N}$ *and any strong solutions* ${}_nY$ *and* Z *to* (2.14) *and* (2.15)*, respectively,*

$$\begin{aligned} E\Big[\max_{j\in\{0,\dots,k_n\}} |{}_nY_{t_{j,n}} - Z_{t_{j,n}}|^2\Big]/c_1 &\leq |\mathbb{T}_n|\Big(1 + E\Big[\|{}_nY^r\|^2 + \|Z^r\|^2\Big]\Big) \\ &+ E\Big[\|{}_nY^r - Z^r\|^2\Big] + E\Big[\|{}_nY - L_n({}_nY)\|^2\Big] + E\Big[\|Z - L_n(Z)\|^2\Big] \\ &+ E\Bigg[\max_{j\in\{0,\dots,k_n\}}\bigg|\int_r^{t_{j,n}} \overline{B}(\underline{s}_n, {}_nY^{\underline{s}_n})\, d({}_nW_s - W_s)\bigg|^2\Bigg] \\ &+ E\Bigg[\max_{j\in\{0,\dots,k_n\}}\bigg|\int_r^{t_{j,n}} \Big(\overline{B}(s,{}_nY^s) - \overline{B}(\underline{s}_n, {}_nY^{\underline{s}_n})\Big){}_n\dot{W}_s - R(\underline{s}_n, {}_nY^{\underline{s}_n})\, ds\bigg|^2\Bigg]. \end{aligned}$$

Proof. We define an increasing function $\varphi_n : [r,T] \to \mathbb{R}_+$ by

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

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

$$\begin{aligned} {}_nY_t - Z_t = \int_r^t &\underline{B}(s,{}_nY^s) - \underline{B}(s,Z^s) + \Big(B_H(s,{}_nY^s) - B_H(s,Z^s)\Big)\dot{h}(s)\, ds \\ &+ {}_nY_r - Z_r + {}_n\Gamma_t + \int_r^t \Sigma(s,{}_nY^s) - \Sigma(s,Z^s)\, dW_s \end{aligned}$$

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

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

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

$$\begin{aligned} \varphi_n(t)^{1/2} \leq &\bigg(c_{1,1}\int_r^{t_n} \beta_n + \zeta_n(s) + \eta_n(s) + \varphi_n(s)\, ds\bigg)^{1/2} \\ &+ \beta_n^{1/2} + \gamma_n(t)^{1/2} \end{aligned} \tag{5.4}$$

for all $t \in [r,T]$, where we have set $c_{1,1} := 15\lambda^2(T - r + \int_r^T |\dot{h}(s)|^2\, ds + w_1)$ and $\beta_n := E[\|{}_nY^r - Z^r\|^2]$ and the functions $\gamma_n, \zeta_n, \eta_n : [r,T] \to [0,\infty)$, which are readily seen to be measurable, are defined via

$$\begin{aligned} \gamma_n(t) &:= E\Big[\max_{j\in\{0,\dots,k_n\}:\, t_{j,n}\leq t} |{}_n\Gamma_{t_{j,n}}|^2\Big], \\ \zeta_n(s) &:= E\Big[\|Y^s - Y^{\underline{s}_n}\|^2 + \|Z^s - Z^{\underline{s}_n}\|^2\Big] \quad \text{and} \end{aligned}$$

$$\eta_n(s) := E\Big[\|Y^{\underline{s}_n} - L_n({}_nY)^{\underline{s}_n}\|^2 + \|Z^{\underline{s}_n} - L_n(Z)^{\underline{s}_n}\|^2\Big].$$

In deriving (5.4), we have used that $E[\|L_n({}_nY)^{\underline{s}_n} - L_n(Z)^{\underline{s}_n}\|^2] \leq \beta_n \vee \varphi_n(s)$ for all $s \in [r,T]$, which follows from Lemma 20, since L_n is linear.

The next step of the proof is to estimate the function γ_n. For this purpose, let us choose two processes ${}_n\Delta, {}_n\Theta \in \mathscr{C}([0,T],\mathbb{R}^m)$ such that

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

and ${}_n\Theta_t = \int_r^t \overline{B}(\underline{s}_n, {}_nY^{\underline{s}_n})\, d({}_nW_s - W_s)$ for each $t \in [r,T]$ a.s. Then ${}_n\Gamma$ can be rewritten in the following way:

$$\begin{aligned}{}_n\Gamma_t = {}_n\Delta_t + {}_n\Theta_t &+ \int_r^t \overline{B}(\underline{s}_n, {}_nY^{\underline{s}_n}) - \overline{B}(s, Z^s)\, dW_s \\ &+ \int_r^t R(\underline{s}_n, {}_nY^{\underline{s}_n}) - R(s, Z^s)\, ds\end{aligned}$$

for all $t \in [r,T]$ a.s. Thus, we set $c_{1,2} := 10\lambda^2(T - r + w_1)$, then it follows readily that

$$\begin{aligned}\gamma_n(t)^{1/2} &\leq \delta_n(t)^{1/2} + \theta_n(t)^{1/2} \\ &\quad + \bigg(c_{1,2}\int_r^{t_n} \beta_n + (s - \underline{s}_n) + \zeta_n(s) + \eta_n(s) + \varphi_n(s)\, ds\bigg)^{1/2}\end{aligned} \tag{5.5}$$

for all $t \in [r,T]$, where the increasing functions $\delta_n, \theta_n : [r,T] \to [0,\infty)$ are given by

$$\delta_n(t) := E\Big[\max_{j\in\{0,\dots,k_n\}:\, t_{j,n}\leq t} |{}_n\Delta_{t_{j,n}}|^2\Big], \quad \theta_n(t) := E\Big[\max_{j\in\{0,\dots,k_n\}:\, t_{j,n}\leq t} |{}_n\Theta_{t_{j,n}}|^2\Big].$$

Proposition 31 and Lemma 32 give constants $l_1, m_1 > 0$ satisfying (5.1) and (5.3) for $p = 1$ when c_p is replaced by l_1 and m_1, respectively. Thus, putting (5.4) and (5.5) together, we find that

$$\begin{aligned}\varphi_n(t) &\leq c_{1,4}|\mathbb{T}_n|\Big(1 + E\Big[\|{}_nY^r\|^2 + \|Z^r\|^2\Big]\Big) + 4\Big(\beta_n + \delta_n(t) + \theta_n(t)\Big) \\ &\quad + c_{1,3}\int_r^{t_n} \beta_n + \eta_n(s) + \varphi_n(s)\, ds\end{aligned}$$

for given $t \in [r,T]$, where we have first set $c_{1,3} := 8(c_{1,1} + c_{1,2})$ and then $c_{1,4} := 2c_{1,3}(T-r)(1 + l_1 + m_1)$. Consequently,

$$\varphi_n(t) \leq c_1\Big(|\mathbb{T}_n|\Big(1 + E\Big[\|{}_nY^r\|^2 + \|Z^r\|^2\Big]\Big) + \beta_n + \delta_n(t) + \theta_n(t) + \eta_n(t)\Big)$$

for $c_1 := e^{c_{1,3}(T-r)}(c_{1,4}+4)$, by Gronwall's inequality. This gives the claim. □

A look at Lemma 21 and Proposition 23 shows us that only the last remainder in the estimation of Proposition 33 requires further analysis, before we can prove (2.18). So, for each $h \in H_r^1([0,T],\mathbb{R}^d)$ and $n \in \mathbb{N}$, we define a map $\Phi_{h,n} : [r,T] \times S \times C([0,T],\mathbb{R}^d) \to \mathbb{R}^m$ by

$$\Phi_{h,n}(s,y,w) := B_H(\underline{s}_n, y)(h(s) - h(\underline{s}_n)) + \Sigma(\underline{s}_n, y)(w(s) - w(\underline{s}_n)) \\ + \overline{B}(\underline{s}_n, y)(s - \underline{s}_n)\Big(\tilde{L}_n(w)(s) - \tilde{L}_n(w)(\underline{s}_n)\Big),$$

where, just as in Lemma 20, for each right-continuous map $w : [0,T] \to \mathbb{R}^d$ we have set $\tilde{L}_n(w)(t) := w(r \wedge t)$ for $t \in [0, t_{1,n})$,

$$\tilde{L}_n(w)(t) := w(t_{i-1,n}) + (t - t_{i,n})\frac{w(t_{i,n}) - w(t_{i-1,n})}{\Delta t_{i+1,n}}$$

for $t \in [t_{i,n}, t_{i+1,n})$ with $i \in \{1, \ldots, k_n - 1\}$ and $\tilde{L}_n(w)(T) := w(t_{k_{n-1},n})$. If now ${}_nY$ and Z are strong solutions to (2.14) and (2.15), respectively, then the following decomposition can be used to deal with the remainder:

$$\begin{aligned} &\Big(\overline{B}(s, {}_nY^s) - \overline{B}(\underline{s}_n, {}_nY^{\underline{s}_n})\Big)_n\dot{W}_s - R(\underline{s}_n, {}_nY^{\underline{s}_n}) \\ &= \Big(\overline{B}(s, {}_nY^s) - \overline{B}(\underline{s}_n, {}_nY^{\underline{s}_n}) - \partial_x\overline{B}(\underline{s}_n, {}_nY^{\underline{s}_n})({}_nY_s - {}_nY_{\underline{s}_n})\Big)_n\dot{W}_s \\ &\quad + \partial_x\overline{B}(\underline{s}_n, {}_nY^{\underline{s}_n})({}_nY_s - {}_nY_{\underline{s}_n} - \Phi_{h,n}(s, {}_nY^s, W^s))_n\dot{W}_s \\ &\quad + \partial_x\overline{B}(\underline{s}_n, {}_nY^{\underline{s}_n})\Phi_{h,n}(s, {}_nY^s, W^s)_n\dot{W}_s - R(\underline{s}_n, {}_nY^{\underline{s}_n}) \end{aligned} \tag{5.6}$$

for any $s \in [r,T)$. In fact, in the next two sections, we will deal with the three terms on the right-hand side to ensure that (2.18) follows.

5.2 Convergence of the first two remainders

To deal with the first remainder term in (5.6), we will use the following estimation in combination with Lemma 19.

Proposition 34. *Assume* (C.i) *and let* $G \in \mathbb{C}^{1,2}([r,T) \times S)$. *Further, let* $\overline{B}$ *and* $\partial_x G$ *be* d_∞*-Lipschitz continuous and suppose there are* $c_0, \eta \geq 0$ *so that*

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

for all $(t,x) \in [r,T) \times S$. *Then for each* $p \geq 1$ *there is* $c_p > 0$ *such that for any* $n \in \mathbb{N}$ *and any strong solution* ${}_nY$ *to* (2.14) *it holds that*

$$\sup_{s\in[r,T)} E\Big[\Big|G(s, {}_nY^s) - G(\underline{s}_n, {}_nY^{\underline{s}_n}) - \partial_x G(\underline{s}_n, {}_nY^{\underline{s}_n})({}_nY_s - {}_nY_{\underline{s}_n})\Big|^{2p}\Big] \\ \leq c_p|\mathbb{T}_n|^{2p}\Big(1 + E\Big[\|{}_nY^r\|^{2(\eta\vee 2)p}\Big]\Big).$$

Proof. Let $s \in [r,T)$ and ${}_n\Delta \in \mathscr{C}([0,T],\mathbb{R}^m)$ be given by ${}_n\Delta_u := 0$ for $u \in [0,T]\backslash[\underline{s}_n, s]$ and ${}_n\Delta_u := \partial_x G'(u, {}_nY^u) - \partial_x G'(\underline{s}_n, {}_nY^{\underline{s}_n})$ for $u \in [\underline{s}_n, s]$. Then the functional Itô formula [9] yields that

$$\begin{aligned} &G(s, {}_nY^s) - G(\underline{s}_n, {}_nY^{\underline{s}_n}) - \partial_x G(\underline{s}_n, {}_nY^{\underline{s}_n})({}_nY_s - {}_nY_{\underline{s}_n}) \\ &= \int_{\underline{s}_n}^s \partial_u G(u, {}_nY^u)\,du + \int_{\underline{s}_n}^s {}_n\Delta_u'\Sigma(u, {}_nY^u)\,dW_u \\ &+ \int_{\underline{s}_n}^s {}_n\Delta_u'\Big(\underline{B}(u, {}_nY^u) + B_H(u, {}_nY^u)\dot{h}(u) + \overline{B}(u, {}_nY^u){}_n\dot{W}_u\Big)\,du \\ &+ \frac{1}{2}\int_{\underline{s}_n}^s \mathrm{tr}(\partial_{xx}G(u, {}_nY^u)\Sigma(u, {}_nY^u)\Sigma(u, {}_nY^u)')\,du \quad \text{a.s.} \end{aligned} \tag{5.7}$$

We set $\overline{\eta} := \eta \vee 2$, then Proposition 31 gives a constant $l_{\overline{\eta}p} > 0$ such that (5.1) holds when c_p and p are replaced by $l_{\overline{\eta}p}$ and $\overline{\eta}p$, respectively. So,

$$E\bigg[\Big|\int_{\underline{s}_n}^s \partial_u G(u, {}_nY^u)\,du\Big|^{2p}\bigg] \leq c_{p,1}|\mathbb{T}_n|^{2p}\Big(1 + E\Big[\|{}_nY^r\|^{2\overline{\eta}p}\Big]\Big)$$

for $c_{p,1} := (4c_0)^{2p}(1 + l_{\overline{\eta}p})^{\eta/\overline{\eta}}$, by Hölder's inequality. Next, let $\lambda_0 \geq 0$ denote a Lipschitz constant for $\partial_x G$, then we obtain that

$$\Big(E\Big[|{}_n\Delta_u|^{4p}\Big]\Big)^{\frac{1}{2}} \leq m_p|\mathbb{T}_n|^p\Big(1 + E\Big[\|{}_nY^r\|^{2\overline{\eta}p}\Big]\Big)^{1/\overline{\eta}} \tag{5.8}$$

for all $u \in [\underline{s}_n, s]$ with $m_p := 2^{3p}\lambda_0^{2p}(1 + l_{\overline{\eta}p})^{1/\overline{\eta}}$. Thus, for the second term in (5.7) Cauchy-Schwarz's inequality gives

$$\begin{aligned} &E\bigg[\Big|\int_{\underline{s}_n}^s {}_n\Delta_u'\underline{B}(u, {}_nY^u)\,du\Big|^{2p}\bigg] \\ &\quad \leq (2c)^{2p}(s - \underline{s}_n)^{2p-1}\int_{\underline{s}_n}^s \Big(E\Big[|{}_n\Delta_u|^{4p}\Big]\Big)^{1/2}\Big(1 + \Big(E\Big[\|{}_nY^u\|^{4p}\Big]\Big)^{1/2}\Big)\,du \\ &\quad \leq c_{p,2}|\mathbb{T}_n|^{2p}\Big(1 + E\Big[\|{}_nY^r\|^{2\overline{\eta}p}\Big]\Big) \end{aligned}$$

with $c_{p,2} := 2^{3p}c^{2p}(T-r)^p m_p(1 + l_{\overline{\eta}p}^{1/\overline{\eta}})$. Similarly, it follows from (5.8) that

$$E\bigg[\Big|\int_{\underline{s}_n}^s {}_n\Delta_u' B_H(u, {}_nY^u)\dot{h}(u)\,du\Big|^{2p}\bigg] \leq c_{p,3}|\mathbb{T}_n|^{2p}\Big(1 + E\Big[\|{}_nY^r\|^{2\overline{\eta}p}\Big]\Big)$$

for $c_{p,3} := 2^{3p}c^{2p}(\int_r^T |\dot{h}(u)|^2\,du)^p m_p(1 + l_{\overline{\eta}p}^{1/\overline{\eta}})$. Lemma 18 yields $\hat{w}_{2p,2} > 0$ satisfying (3.11) when p and q are replaced by $2p$ and 2, respectively. Then

$$E\bigg[\Big|\int_{\underline{s}_n}^s {}_n\Delta_u'\overline{B}(u, {}_nY^u){}_n\dot{W}_u\,du\Big|^{2p}\bigg]$$

$$\leq c^{2p}(s-\underline{s}_n)^{p-1}\int_{\underline{s}_n}^{s}\Big(E\Big[|{}_n\Delta_u|^{4p}\Big]\Big)^{1/2}\bigg(E\bigg[\bigg(\int_{\underline{s}_n}^{s}|{}_n\dot{W}_v|^2\,dv\bigg)^{2p}\bigg]\bigg)^{1/2}du$$
$$\leq c_{p,4}|\mathbb{T}_n|^{2p}\Big(1+E\Big[\|{}_nY^r\|^{2\overline{\eta}p}\Big]\Big),$$

by Cauchy-Schwarz's inequality, where we have set $c_{p,4} := (2c)^{2p}m_p\hat{w}_{2p,2}^{1/2}$. Another estimation shows us that

$$E\bigg[\bigg|\int_{\underline{s}_n}^{s}{}_n\Delta_u'\Sigma(u,{}_nY^u)\,dW_u\bigg|^{2p}\bigg]\leq c_{p,5}|\mathbb{T}_n|^{2p}\Big(1+E\Big[\|{}_nY^r\|^{2\overline{\eta}p}\Big]\Big)$$

for $c_{p,5} := 2^pc^{2p}w_pm_p$, where w_p is the constant satisfying (M). We move on to the last term arising in (5.7). Here, we readily compute that

$$E\bigg[\bigg|\frac{1}{2}\int_{\underline{s}_n}^{s}\mathrm{tr}(\partial_{xx}G(u,{}_nY^u)\Sigma(u,{}_nY^u)\Sigma(u,{}_nY^u)')\,du\bigg|^{2p}\bigg]$$
$$\leq c_{p,6}|\mathbb{T}_n|^{2p}\Big(1+E\Big[\|{}_nY^r\|^{2\overline{\eta}p}\Big]\Big)$$

with $c_{p,6} := (2c_0)^{2p}c^{4p}(1+l_{\overline{\eta}p})^{\eta/\overline{\eta}}$. Thus, by setting $c_p := 6^{2p-1}(c_{p,1}+\cdots+c_{p,6})$, we obtain the asserted estimate. □

We come to the second remainder term arising in (5.6). As before, we will derive an estimation that is necessary to apply Lemma 19.

Lemma 35. *Let* (C.i) *and* (C.ii) *hold and* $h\in H_r^1([0,T],\mathbb{R}^d)$. *Then for each* $p\geq 1$ *there is* $c_p>0$ *such that for any* $n\in\mathbb{N}$ *and any strong solution* ${}_nY$ *to* (2.14) *we have*

$$\sup_{s\in[r,T]}\Big[|{}_nY_s-{}_nY_{\underline{s}_n}-\Phi_{h,n}(s,{}_nY^s,W^s)|^{2p}\Big]\leq c_p|\mathbb{T}_n|^{2p}\Big(1+E\Big[\|{}_nY^r\|^{4p}\Big]\Big)^{1/2}.$$

Proof. We apply Proposition 31 to get a constant $l_{2p}>0$ such that (5.1) is satisfied when c_p and p are replaced by l_{2p} and $2p$, respectively. Let us pick $s\in[r,T]$, then

$$E\bigg[\bigg|\int_{\underline{s}_n}^{s}\underline{B}(u,{}_nY^u)\,du\bigg|^{2p}\bigg]\leq c_{p,1}|\mathbb{T}_n|^{2p}\Big(1+E\Big[\|{}_nY^r\|^{4p}\Big]\Big)^{1/2}$$

for $c_{p,1} := (4c)^{2p}(1+l_{2p}^{1/2})$. Let $\lambda\geq 0$ denote a Lipschitz constant for $B_H,\overline{B}$ and Σ, then Cauchy-Schwarz's inequality allows us to estimate that

$$E\bigg[\bigg|\int_{\underline{s}_n}^{s}\Big(B_H(u,{}_nY^u)-B_H(\underline{s}_n,{}_nY^{\underline{s}_n})\Big)\dot{h}(u)\,du\bigg|^{2p}\bigg]$$

$$\leq 2^{3p-1}\lambda^{2p}\bigg(\int_r^T |\dot{h}(u)|^2\,du\bigg)^p |\mathbb{T}_n|^p\Big((s-\underline{s}_n)^p + E\Big[\|{}_nY^s - {}_nY^{\underline{s}_n}\|^{2p}\Big]\Big)$$
$$\leq c_{p,2}|\mathbb{T}_n|^{2p}\Big(1+E\Big[\|{}_nY^r\|^{4p}\Big]\Big)^{1/2}$$

with $c_{p,2} := (4\lambda)^{2p}(\int_r^T |\dot{h}(u)|^2\,du)^p(1+l_{2p})^{1/2}$. We recall the constant $\hat{w}_{2p,2}$ constructed in Lemma 18 such that (3.11) is valid when p and q are replaced by $2p$ and 2, respectively. Then

$$E\bigg[\bigg|\int_{\underline{s}_n}^s \Big(\overline{B}(u,{}_nY^u) - \overline{B}(\underline{s}_n,{}_nY^{\underline{s}_n})\Big)\,{}_n\dot{W}_u\,du\bigg|^{2p}\bigg]$$
$$\leq \lambda^{2p}(s-\underline{s}_n)^p E\bigg[\Big((s-\underline{s}_n)^{1/2} + \|{}_nY^s - {}_nY^{\underline{s}_n}\|\Big)^{2p}\bigg(\int_{\underline{s}_n}^s |{}_n\dot{W}_v|^2\,dv\bigg)^p\bigg]$$
$$\leq c_{p,3}|\mathbb{T}_n|^{2p}\Big(1+E\Big[\|{}_nY^r\|^{4p}\Big]\Big)^{1/2},$$

by Cauchy-Schwarz's inequality, where $c_{p,3} := 2^{5p}\lambda^{2p}(1+l_{2p})^{1/2}\hat{w}_{2p,2}^{1/2}$. Finally, let also recall the constant w_p appearing in (M), then

$$E\bigg[\bigg|\int_{\underline{s}_n}^s \Sigma(u,{}_nY^u) - \Sigma(\underline{s}_n,{}_nY^{\underline{s}_n})\,dW_u\bigg|^{2p}\bigg] \leq c_{p,4}|\mathbb{T}_n|^{2p}\Big(1+E\Big[\|{}_nY^r\|^{4p}\Big]\Big)^{1/2}$$

for $c_{p,4} := (4\lambda)^{2p}w_p(1+l_{2p})^{1/2}$. So, the definition $c_p := 4^{2p-1}(c_{p,1}+\cdots+c_{p,4})$ concludes the proof. □

5.3 Convergence of the third remainder

As preparation, we require the following application of Doob's L^2-martingale inequality.

Lemma 36. *For each $l \in \{1,\ldots,d\}$ and $n \in \mathbb{N}$, let $({}_{l,n}U_i)_{i\in\{1,\ldots,k_n\}}$ be an $(\mathscr{F}_{t_{i,n}})_{i\in\{1,\ldots,k_n\}}$-predictable sequence of $\mathbb{R}^{1\times d}$-valued random vectors and $({}_{l,n}V_i)_{i\in\{1,\ldots,k_n\}}$ be an $(\mathscr{F}_{t_{i,n}})_{i\in\{1,\ldots,k_n\}}$-adapted sequence of $\mathbb{R}^d$-valued random vectors such that*

$$E[|{}_{l,n}U_i|^4] + E[|{}_{l,n}V_i|^4] < \infty \quad \textit{and} \quad E[{}_{l,n}V_i|\mathscr{F}_{t_{i-1,n}}] = 0 \quad \textit{a.s.}$$

for all $i \in \{1,\ldots,k_n\}$. Then

$$E\bigg[\max_{j\in\{i_0,\ldots,k_n\}}\bigg|\sum_{i=1}^{j-i_0}\sum_{l=1}^d {}_{l,n}U_i\,{}_{l,n}V_i\bigg|^2\bigg]$$
$$\leq 4\sum_{i=1}^{k_n-i_0}\sum_{l_1,l_2=1}^d E[{}_{l_1,n}U_i E[{}_{l_1,n}V_i\,{}_{l_2,n}V_i'|\mathscr{F}_{t_{i-,1n}}]\,{}_{l_2,n}U_i']$$

for all $i_0 \in \{0,\ldots,k_n-1\}$ and $n \in \mathbb{N}$.

Proof. Let us set ${}_nY_i := \sum_{l=1}^d {}_{l,n}U_i\, {}_{l,n}V_i$ for each $i \in \{1, \dots, k_n - i_0\}$, then ${}_nY_i$ is $\mathscr{F}_{t_{i,n}}$-measurable, $E[|{}_nY_i|^2] < \infty$ and $E[{}_nY_i|\mathscr{F}_{t_{i-1,n}}] = 0$ a.s. Hence, the discrete-time process ${}_nS : \{i_0, \dots, k_n\} \times \Omega \to \mathbb{R}$ defined via

$$ {}_nS_j := \sum_{i=1}^{j-i_0} {}_nY_i $$

is a square-integrable martingale with respect to $(\mathscr{F}_{t_{j-i_0,n}})_{j\in\{i_0,\dots,k_n\}}$. For this reason, Doob's $\mathscr{L}^2$-martingale inequality implies that

$$ E\bigg[\max_{j\in\{i_0,\dots,k_n\}}\bigg|\sum_{i=1}^{j-i_0}\sum_{l=1}^{d} {}_{l,n}U_i\, {}_{l,n}V_i\bigg|^2\bigg] = E\Big[\max_{j\in\{i_0,\dots,k_n\}} {}_nS_j^2\Big] \le 4E[{}_nS_{k_n}^2]. $$

Moreover, let $i, j \in \{1, \dots, k_n - i_0\}$ be such that $i \le j$, then we observe that $E[{}_{l_1,n}V_i\, {}_{l_2,n}V_j'|\mathscr{F}_{t_{j-1,n}}] = \mathbb{1}_{\{i\}}(j)E[{}_{l_1,n}V_i\, {}_{l_2,n}V_i'|\mathscr{F}_{t_{i-1,n}}]$ a.s. and

$$ E[{}_nY_i\, {}_nY_j] = \sum_{l_1,l_2=1}^{d} E[{}_{l_1,n}U_i E[{}_{l_1,n}V_i\, {}_{l_2,n}V_j'|\mathscr{F}_{t_{j-1,n}}]\, {}_{l_2,n}U_j']. $$

In particular, ${}_nY_i$ and ${}_nY_j$ are uncorrelated for $i < j$. By Bienaymé's identity, $E[{}_nS_{k_n}^2] = \sum_{i=1}^{k_n-i_0} E[{}_nY_i^2]$, which yields the claim. □

Proposition 37. *Let* (C.i) *and* (C.ii) *hold and* $h \in H_r^1([0,T], \mathbb{R}^d)$. *Then there is* $c_1 > 0$ *such that for any* $n \in \mathbb{N}$ *and any strong solution* ${}_nY$ *to* (2.14) *it follows that*

$$ E\bigg[\max_{j\in\{0,\dots,k_n\}}\bigg|\int_r^{t_{j,n}} \partial_x\overline{B}(\underline{s}_n, {}_nY^{\underline{s}_n})\Phi_{h,n}(s, {}_nY^s, W^s)_n\dot{W}_s - R(\underline{s}_n, {}_nY^{\underline{s}_n})\, ds\bigg|^2\bigg] \le c_1|\mathbb{T}_n|\Big(1 + E\Big[\|{}_nY^r\|^2\Big]\Big). $$

Proof. Due to Proposition 31, we may apply Lemma 22, which provides a constant $c_{1,0} > 0$ that is independent of n such that

$$ E\bigg[\max_{j\in\{0,\dots,k_n\}}\bigg|\int_r^{t_{j,n}} R(\underline{s}_n, {}_nY^{\underline{s}_n})(\delta_n(s) - 1)\, ds\bigg|^2\bigg] \le c_{1,0}|\mathbb{T}_n|\Big(1 + E\Big[\|{}_nY^r\|^2\Big]\Big), $$

where $\delta_n(s) := \Delta s_n/\Delta\overline{s}_n$ for all $s \in [r, T]$. We recall the definition of R in (2.16) to write the k-th coordinate of $\partial_x\overline{B}(\underline{s}_n, {}_nY^{\underline{s}_n})\Phi_{h,n}(s, {}_nY^s, W^s)\, {}_n\dot{W}_s - R(\underline{s}_n, {}_nY^{\underline{s}_n})\delta_n(s)$ in the form

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

for all $k \in \{1,\dots,m\}$ and $s \in [r,T)$. Moreover, we decompose that

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

where $l \in \{1,\dots,d\}$. We begin with the first term in this decomposition and use Lemma 16 to obtain that

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

for each $j \in \{1,\dots,k_n\}$, $k \in \{1,\dots,m\}$ and $l \in \{1,\dots,d\}$. Proposition 31 gives $l_1 > 0$ such that (5.1) is satisfied for $p = 1$ when the appearing constant c_p is replaced by l_1. Hence, condition (C.ii) gives

$$
\begin{aligned}
&E\bigg[\max_{j\in\{0,\dots,k_n\}} \sum_{k=1}^m \bigg|\int_r^{t_{j,n}} \sum_{l=1}^d (\partial_x\overline{B}_{k,l}B_H)(\underline{s}_n, {}_nY^{\underline{s}_n})(h(s_n) - h(\underline{s}_n))_n\dot{W}_s^{(l)}\,ds\bigg|^2\bigg] \\
&\qquad \leq 2c^4 w_1 \int_r^{t_{k_n-1,n}} \Big(1 + E\Big[\|{}_nY^{s_n}\|^2\Big]\Big)|h(\overline{s}_n) - h(s_n)|^2\,ds \\
&\qquad \leq c_{1,1}|\mathbb{T}_n|\Big(1 + E\Big[\|{}_nY^r\|^2\Big]\Big),
\end{aligned}
$$

where w_1 satisfies (M) for $p = 1$ and $c_{1,1} := 2c^4 w_1(1 + l_1)(T - r)\int_r^T |\dot{h}(s)|^2\,ds$. Similarly, another application of Lemma 16 gives us that

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

for all $j \in \{1,\dots,k_n\}$, $k \in \{1,\dots,m\}$ and $l \in \{1,\dots,d\}$. Thus, we define $c_{1,2} := c^4 w_1 d(T - r)$, then we can estimate that

$$
E\bigg[\max_{j\in\{0,\dots,k_n\}} \sum_{k=1}^m \bigg|\int_r^{t_{j,n}} \sum_{l=1}^d (\partial_x\overline{B}_{k,l}\overline{B})(\underline{s}_n, {}_nY^{\underline{s}_n})({}_nW_{s_n} - {}_nW_{\underline{s}_n})_n\dot{W}_s^{(l)}\,ds\bigg|^2\bigg]
$$

$$\leq c^4 w_1 \int_r^{t_{k_n-1,n}} E\Big[|\Delta W_{s_n}|^2\Big]\, ds \leq c_{1,2}|\mathbb{T}_n|.$$

Let us move on to the third expression in (5.9). First of all, we define an $\mathbb{R}^d$-valued $\mathscr{F}_{t_{i,n}}$-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 $l \in \{1, \ldots, d\}$ and $i \in \{1, \ldots, k_n\}$, then $_{l,n}V_i$ is independent of $\mathscr{F}_{t_{i-1,n}}$ and satisfies $E[|_{l,n}V_i|^4] < \infty$ and $E[_{l,n}V_i] = 0$. Moreover, a case distinction shows that

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

for all $l_1, l_2 \in \{1, \ldots, d\}$ and $i \in \{1, \ldots, k_n\}$, 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

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

for all $j \in \{1, \ldots, k_n\}$, $k \in \{1, \ldots, m\}$ and $l \in \{1, \ldots, d\}$, since $\delta_n(s) = 0$ for each $s \in [r, t_{1,n}]$. Consequently, Lemma 36 and the representation (5.11) imply that

$$\begin{aligned}
&E\Bigg[\max_{j\in\{0,\ldots,k_n\}} \sum_{k=1}^m \bigg| \int_r^{t_{j,n}} \sum_{l=1}^d (\partial_x \overline{B}_{k,l}\Sigma)(\underline{s}_n, {}_nY^{\underline{s}_n})\Big(\Delta W_{s_n n}\dot{W}_s^{(l)} - \delta_n(s)e_l\Big)\, ds\bigg|^2\Bigg] \\
&\qquad \leq 8 \sum_{i=1}^{k_n-1} (\Delta t_{i,n})^2 \sum_{k=1}^m \sum_{l=1}^d E[|(\partial_x \overline{B}_{k,l}\Sigma)(t_{i-1,n}, {}_nY^{t_{i-1,n}})|^2] \leq c_{1,3}|\mathbb{T}_n|
\end{aligned}$$

for $c_{1,3} := 8c^4(T - r)$, since we can use that $x^t \mathbb{I}_{l_2,l_1} y \leq (1/2)(x_{l_2}^2 + y_{l_1}^2)$ for all $x, y \in \mathbb{R}^d$, by Young's inequality. To deal with the fourth term in (5.9), let us note that

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

for each $j \in \{1, \ldots, k_n\}$, $k \in \{1, \ldots, m\}$ and $l \in \{1, \ldots, d\}$, 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 $i \in \{0, \ldots, k_n - 1\}$. So,

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

with $c_{1,4} := 2c^4(1 + l_1)(T - r) \int_r^T |\dot{h}(s)|^2\, ds$, by Cauchy-Schwarz's inequality and the facts that $\Delta W_{s_n}^{(1)}, \ldots, \Delta W_{s_n}^{(d)}$ are not only pairwise independent but also independent of $\mathscr{F}_{\underline{s}_n}$ for every $s \in [r, T]$.

Next, to handle the fifth expression in (5.9), we proceed similarly as with the third expression. We define ${}_{l,n}U_s := ({}_nW_s - {}_nW_{s_n})_n \dot{W}_s^{(l)} - (1/2)\delta_n(s) e_l$ for all $s \in [r, T]$ and note that

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

for every $j \in \{1, \ldots, k_n\}$, $k \in \{1, \ldots, m\}$ and $l \in \{1, \ldots, d\}$, where ${}_{l,n}V_i$ is defined via (5.10) and we have used that $\int_{t_{i,n}}^{t_{i+1,n}} s - t_{i,n}\, ds = (1/2)(\Delta t_{i+1,n})^2$ for all $i \in \{1, \ldots, k_n\}$. Consequently,

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

for $c_{1,5} := 2c^4(T - r)$. We turn to the last term in (5.9) and proceed just as in (5.12) to get that

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

for each $j \in \{1, \ldots, k_n\}$, $k \in \{1, \ldots, m\}$ and $l \in \{1, \ldots, d\}$, 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 $i \in \{0, \ldots, k_n - 1\}$.

Therefore,

$$E\bigg[\max_{j\in\{0,\dots,k_n\}}\sum_{k=1}^m\bigg|\int_r^{t_{j,n}}\sum_{l=1}^d(\partial_x\overline{B}_{k,l}\Sigma)(\underline{s}_n,{}_nY^{\underline{s}_n})(W_s-W_{s_n})_n\dot{W}_s^{(l)}\,ds\bigg|^2\bigg]$$
$$\leq w_1\sum_{k=1}^m\int_r^T E\Big[\Big|\sum_{l=1}^d(\partial_x B_{k,l}\Sigma)(\underline{s}_n,{}_nY^{\underline{s}_n})\Delta W_{s_n}^{(l)}\Big|^2\Big]\,ds\leq c_{1,6}|\mathbb{T}_n|$$

with $c_{1,6}:=c^4w_1(T-r)$. As before, we have used that $\Delta W_{s_n}^{(1)},\dots,\Delta W_{s_n}^{(d)}$ are pairwise independent and independent of $\mathscr{F}_{\underline{s}_n}$ for all $s\in[r,T]$. Hence, by setting $c_1:=7(c_{1,0}+\cdots+c_{1,6})$, the assertion follows. □

5.4 Proofs of Theorems 8 and 1

Proof of Theorem 8. (i) Applying Girsanov's theorem shows that existence and uniqueness follow from Proposition 5 when $B=\underline{B}+B_H\dot{h}$. Further, Propositions 31 and 12 imply the second claim.

(ii) This assertion is an immediate application of Proposition 5 when $B=\underline{B}+R+B_H\dot{h}$ and Σ is replaced by $\overline{B}+\Sigma$.

(iii) By Propositions 33 and 23 and Lemma 21, to establish (2.18), it suffices to show that there is $c_1>0$ such that

$$E\bigg[\max_{j\in\{0,\dots,k_n\}}\bigg|\int_r^{t_{j,n}}\Big(\overline{B}(s,{}_nY^s)-\overline{B}(\underline{s}_n,{}_nY^{\underline{s}_n})\Big)_n\dot{W}_s-R(\underline{s}_n,{}_nY^{\underline{s}_n})\,ds\bigg|^2\bigg]\leq c_1|\mathbb{T}_n|$$

for all $n\in\mathbb{N}$. To this end, we utilize the decomposition (5.6). First of all, Proposition 34 allows us to apply Lemma 19, which yields $c_{1,1}>0$ so that

$$E\bigg[\bigg(\int_r^T|\overline{B}(s,{}_nY^s)-\overline{B}(\underline{s}_n,{}_nY^{\underline{s}_n})-\partial_x\overline{B}(\underline{s}_n,{}_nY^{\underline{s}_n})({}_nY_s-{}_nY_{\underline{s}_n})||_n\dot{W}_s|\,ds\bigg)^2\bigg]$$
$$\leq c_{1,1}|\mathbb{T}_n|$$

for all $n\in\mathbb{N}$. Secondly, since $\partial_x\overline{B}$ is bounded, a combination of Lemma 35 with Lemma 19 gives $c_{1,2}>0$ such that

$$E\bigg[\bigg(\int_r^T|\partial_x\overline{B}(\underline{s}_n,{}_nY^{\underline{s}_n})({}_nY_s-{}_nY_{\underline{s}_n}-\Phi_{h,n}(s,{}_nY^s,W^s))||_n\dot{W}_s|\,ds\bigg)^2\bigg]\leq c_{1,1}|\mathbb{T}_n|$$

for all $n\in\mathbb{N}$. Thirdly, using again the boundedness of $\partial\overline{B}$, it follows from Proposition 37 and Lemma 19 that there is $c_{1,3}>0$ such that

$$E\bigg[\max_{j\in\{0,\dots,k_n\}}\bigg|\int_r^{t_{j,n}}\partial_x\overline{B}(\underline{s}_n,{}_nY^{\underline{s}_n})\Phi_{h,n}(s,{}_nY^s,W^s)_n\dot{W}_s-R(\underline{s}_n,{}_nY^{\underline{s}_n})\,ds\bigg|^2\bigg]$$

$$\leq c_{1,3}|\mathbb{T}_n|$$

for each $n \in \mathbb{N}$. Hence, by setting $c_1 := 3(c_{1,1} + c_{1,2} + c_{1,3})$, we obtain the desired estimate. For this reason, (2.18) holds.

To justify the second assertion, we set ${}_nU := {}_nY - Z \in \mathscr{C}^{1/2-}_{r,\infty}([0,T],\mathbb{R}^m)$ for all $n \in \mathbb{N}$, then $(\|{}_nU^r\|)_{n\in\mathbb{N}}$ and $(\max_{j\in\{0,\dots,k_n\}} |{}_nU_{t_{j,n}}|/|\mathbb{T}_n|^\alpha)_{n\in\mathbb{N}}$ converge in probability to zero, by (2.18) and Chebyshev's inequality. Let $p > 1$ be such that $2\alpha < 1 - 1/p$, then Proposition 31 and corollary 32 give $c_p > 0$ satisfying

$$E\Big[\|{}_nU^s - {}_nU^t\|^{2p}\Big] \leq c_p|s-t|^p$$

for all $n \in \mathbb{N}$ and $s,t \in [r,T]$. As Remark 15 entails that condition (3.7) is satisfied for $q = p-1$, Proposition 14 implies that $(\|{}_nU\|_{\alpha,r})_{n\in\mathbb{N}}$ converges in probability to zero, which establishes (2.17). □

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

Lemma 38. *Let $(\tilde{\Omega}, \tilde{\mathscr{F}}, \tilde{P})$ be a probability space, $(\tilde{E},\tilde{\rho})$ be a metric space, $D \subset \tilde{E}$ and $Y : \tilde{\Omega} \to \tilde{E}$ be measurable such that $\tilde{P} \circ Y^{-1}$ is inner regular.*

(i) Let $(Y_n)_{n\in\mathbb{N}}$ be a sequence of $\tilde{E}$-valued maps on Ω that converges to Y in probability. If $Y_n \in D$ a.s. for all $n \in \mathbb{N}$, then $\mathrm{supp}(\tilde{P}\circ Y^{-1}) \subset \overline{D}$.

(ii) Suppose that for each $y \in D$ there is a sequence $(P_{y,n})_{n\in\mathbb{N}}$ of probability measures on $(\Omega,\mathscr{F})$ such that $\tilde{P}_{y,n} \ll \tilde{P}$ for all $n \in \mathbb{N}$ and

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

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

Proof. (i) Let $y \in \mathrm{supp}(\tilde{P}\circ Y^{-1})$ and $k \in \mathbb{N}$. Since $(Y_n)_{n\in\mathbb{N}}$ converges to Y in probability, there exists $n_k \in \mathbb{N}$ such that $\tilde{P}(\tilde{\rho}(Y_n,Y) > 1/(2k)) < \tilde{P}(\tilde{\rho}(Y,y) < 1/(2k))$ for all $n \in \mathbb{N}$ with $n \geq n_k$. Hence,

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

for any such $n \in \mathbb{N}$. So, there is $\omega_k \in \tilde{\Omega}$ such that $y_k := Y_{n_k}(\omega_k) \in D$ and $\tilde{\rho}(y_k,y) < 1/k$. 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, suppose that there are $y \in \overline{D}$ and $\varepsilon > 0$ such that $\tilde{P}(\tilde{\rho}(Y,y) \geq \varepsilon) = 1$. 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 $\tilde{\rho}(y_{n_\varepsilon},y) < \varepsilon/2$. Then from

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

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

Lemma 39. *Let $h \in H_r^1([0,T],\mathbb{R}^d)$ and $n \in \mathbb{N}$, then the a.s. continuous local martingale ${}_{h,n}Z$ given by* (2.11) *is a martingale.*

Proof. We recall that $\int_{t_{i,n}}^{t_{i+1,n}} \dot{h}(s)\, dW_s$ is independent of $\mathscr{F}_{t_{i,n}}$ and normally distributed with zero mean and variance given by $\int_{t_{i,n}}^{t_{i+1,n}} |\dot{h}(s)|^2\, ds$, which yields that

$$\begin{aligned} &E\bigg[\exp\bigg(\int_{t_{i,n}}^{t_{i+1,n}} (\dot{h}(s) - {}_n\dot{W}_s)\, dW_s - \frac{1}{2}\int_{t_{i,n}}^{t_{i+1,n}} |\dot{h}(s) - {}_n\dot{W}_s|^2\, ds\bigg)\bigg|\mathscr{F}_{t_{i,n}}\bigg] \\ &= E\bigg[\exp\bigg(\int_{t_{i,n}}^{t_{i+1,n}} (\dot{h}(s) - x)\, dW_s - \frac{1}{2}\int_{t_{i,n}}^{t_{i+1,n}} |\dot{h}(s) - x|^2\, ds\bigg)\bigg]\bigg|_{x = \frac{\Delta W_{t_{i,n}}}{\Delta t_{i+1,n}}} = 1 \end{aligned}$$

a.s. for each $i \in \{1, \ldots, k_n - 1\}$, since $\Delta W_{t_{i+1,n}}$ is also independent of $\mathscr{F}_{t_{i,n}}$ and we have $\int_{t_{i,n}}^{t_{i+1,n}} {}_n\dot{W}_s\, dW_s = (\Delta W'_{t_{i,n}}/\Delta t_{i+1,n})\Delta W_{t_{i+1,n}}$. Hence, $E[{}_{h,n}Z_T] = 1$ follows by induction, from which we infer the claim. □

Proof of Theorem 1. (i) By Lemma 30, pathwise uniqueness holds for (1.1) and Proposition 5 provides a unique strong solution X such that $X_s = \hat{x}(s)$ for all $s \in [0,r]$ a.s. and $X \in \mathscr{C}_{r,\infty}^{1/2-}([0,T],\mathbb{R}^m)$. For this reason, (i) holds.

(ii) Let $h \in H_r^1([0,T],\mathbb{R}^d)$ and set $F_h := b - (1/2)\rho + \sigma\dot{h}$. We first check that F_h satisfies conditions (O.i) and (O.ii). Since σ and $\partial_x\sigma$ are bounded, there is $c_1 \geq 0$ such that $|\sigma| \vee |\rho| \leq c_1$. Then

$$|F_h(t,x)| \leq c(1 + \|x\|^\kappa) + c_1(1 + |\dot{h}(t)|) \leq c_2(1 + |\dot{h}(t)|)(1 + \|x\|)$$

for all $(t,x) \in [r,T) \times S$ with $c_2 := 3\max\{c, c_1\}$. Moreover, since σ and $\partial_x\sigma$ are d_∞-Lipschitz continuous, so is the map ρ. Thus, let $\lambda_1 \geq 0$ be a Lipschitz constant for ρ, then

$$|F_h(t,x) - F_h(t,y)| \leq (\lambda_1 + \lambda(1 + |\dot{h}(t)|))\|x - y\| \leq \lambda_2(1 + |\dot{h}(t)|)\|x - y\|$$

for all $t \in [r,T)$ and $x, y \in S$ with $\lambda_2 := 2\max\{\lambda, \lambda_1\}$. Hence, an application of Proposition 2 to the map F_h yields the first assertion and we may set $x_h := y_{F_h}$.

Regarding the second, let us also choose $l \in H_r^1([0,T],\mathbb{R}^d)$. We define $c_3 := 2^2(T - r + 1)$, then the above estimation shows that

$$\|x_h^t - x_l^t\|_{H,r}^2 \leq c_3 \int_r^t |F_h(s, x_h) - F_l(s, x_l^s)|^2\, ds$$

$$\leq c_3 \int_r^t 2\lambda_2^2(1 + |\dot{h}(s)|^2)\|x_h^s - x_l^s\|^2 + c_1^2|\dot{h}(s) - \dot{l}(s)|^2\, ds$$

for given $t \in [r,T]$. For this reason, $\|x_h - x_l\|_{H,r}^2 \leq c_4 e^{c_4\|h\|_{H,r}^2}\|h - l\|_{H,r}^2$ with $c_4 := (c_1^2 + 2\lambda_2^2)c_3 \exp(2\lambda_2^2 c_3(T-r))$, by Gronwall's inequality. As c_4 merely depends on $T - r$, c_1 and λ_2, the second claim follows.

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

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

is inner regular, where $\mathscr{B}(C_r^\alpha([0,T],\mathbb{R}^m))$ is the Borel σ-field with respect to the complete norm $\|\cdot\|_{\alpha,r}$. Note that the support of (5.14) consists of all $x \in C_r^\alpha([0,T],\mathbb{R}^m)$ satisfying $P(\{\|X - x\|_{\alpha,r} \geq \varepsilon\} \cap N_\alpha^c) > 0$ for all $\varepsilon > 0$, that is, it is the support of $P \circ X^{-1}$ in $C_r^\alpha([0,T],\mathbb{R}^m)$.

For $n \in \mathbb{N}$ we define $Y_n : N_\alpha^c \to C_r^\alpha([0,T],\mathbb{R}^m)$ by $Y_n(\omega) := x_{{}_nW(\omega)}$, then $Y_n \in \{x_h \,|\, h \in H_r^1([0,T],\mathbb{R}^d)\}$. So, Lemma 38 entails that the support of (5.14) is included in the closure of $\{x_h \,|\, h \in H_r^1([0,T],\mathbb{R}^d)\}$ with respect to $\|\cdot\|_{\alpha,r}$ once we have shown that

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

This, however, already follows from Theorem 8 by the choice $\underline{B} = b - (1/2)\rho$, $B_H = 0$, $\overline{B} = \sigma$ and $\Sigma = 0$. To obtain the converse inclusion in (1.7), let $h \in H_r^1([0,T],\mathbb{R}^d)$ and for each $n \in \mathbb{N}$ define $P_{h,n} : \mathscr{F} \to [0,1]$ by

$$P_{h,n}(A) := E[{}_{h,n}Z_T \mathbb{1}_A],$$

where the a.s. continuous local martingale ${}_{h,n}Z : [0,T] \times \Omega \to (0,\infty)$ given by ${}_{h,n}Z^r = 1$ a.s. and (2.11) is shown in Lemma 39 to be a martingale. So, $P_{h,n}$ is a probability measure satisfying $P_{h,n} \sim P$ and by Lemma 38, if

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

then the closure of $\{x_l \,|\, l \in H_r^1([0,T],\mathbb{R}^d)\}$ 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 ${}_{h,n}W : [0,T] \times \Omega \to \mathbb{R}^d$ defined by

$${}_{h,n}W_t := W_t - \int_r^{r\vee t} \dot{h}(s) - {}_n\dot{W}_s\, ds$$

is a d-dimensional $(\mathscr{F}_t)_{t\in[0,T]}$-Brownian motion under $P_{h,n}$ and X is a strong solution to (2.13) under $P_{h,n}$. Hence, an application of Theorem 8 in the case that $\underline{B} = b$, $B_H = \sigma$, $\overline{B} = -\sigma$ and $\Sigma = \sigma$ gives (2.12). As this readily implies (5.15), the proof is complete. $\square$

References

[1] S. Aida, *Support theorem for diffusion processes on hilbert spaces*, Publ. Res. Inst. Math. Sci., 26 (1990), pp. 947–965.

[2] S. Aida, S. Kusuoka, and 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, Harlow, Essex: Longman Scientific & Technical; New York: Wiley, 1993, pp. 3–34.

[3] A. Ananova and R. Cont, *Pathwise integration with respect to paths of finite quadratic variation*, Journal de Mathématiques Pures et Appliquées, 107 (2017), pp. 737–757.

[4] V. Bally, A. Millet, and M. Sanz-Sole, *Approximation and support theorem in holder norm for parabolic stochastic partial differential equations*, Ann. Probab., 23 (1995), pp. 178–222.

[5] G. Ben Arous, M. Gradinaru, and M. Ledoux, *Hölder norms and the support theorem for diffusions*, Annales de l'IHP - Probabilités et Statistiques, 30 (1994), pp. 415–436.

[6] R. Cont, *Functional Ito Calculus and functional Kolmogorov equations*, in Stochastic Integration by Parts and Functional Ito Calculus (Lecture Notes of the Barcelona Summer School in Stochastic Analysis, July 2012), F. Utzet and J. Vives, eds., Advanced Courses in Mathematics, Birkhauser Basel, 2016, pp. 115–208.

[7] R. Cont and D.-A. Fournié, *A functional extension of the Ito formula*, C. R. Math. Acad. Sci. Paris, 348 (2009), pp. 57–61.

[8] R. Cont and D.-A. Fournié, *Change of variable formulas for non-anticipative functionals on path space*, J. Funct. Anal., 259 (2010), pp. 1043–1072.

[9] R. Cont and D.-A. Fournié, *Functional Itô calculus and stochastic integral representation of martingales*, The Annals of Probability, 41 (2013), pp. 109–133.

[10] R. Cont and Y. Lu, *Weak approximation of martingale representations*, Stochastic Process. Appl., 126 (2016), pp. 857–882.

[11] R. Cont and D. Purba, *On pathwise quadratic variation*, working paper, 2017.

[12] B. Dupire, *Functional Itô calculus*, working paper, 2009.

[13] P. K. Friz and N. B. Victoir, *Multidimensional stochastic processes as rough paths*, vol. 120 of Cambridge Studies in Advanced Mathematics, Cambridge University Press, Cambridge, 2010. Theory and applications.

[14] I. Gyöngy and T. Pröhle, *On the approximation of stochastic differential equation and on Stroock-Varadhan's support theorem*, Computers & Mathematics with Applications, 19 (1990), pp. 65 – 70.

[15] M. Ledoux, Z. Qian, and T. Zhang, *Large deviations and support theorem for diffusion processes via rough paths*, Stochastic Processes and their Applications, 102 (2002), pp. 265 – 283.

[16] A. Millet and M. Sanz-Solé, *A simple proof of the support theorem for diffusion processes*, in Séminaire de Probabilités, XXVIII, vol. 1583 of Lecture Notes in Math., Springer, Berlin, 1994, pp. 36–48.

[17] S. E. A. Mohammed, *Stochastic Functional Differential Equations*, Research Notes in Mathematics, Pitman, 1984.

[18] M. S. Pakkanen, *Stochastic integrals and conditional full support*, Journal of Applied Probability, 47 (2010), p. 650–667.

[19] P. E. Protter, *Stochastic integration and differential equations*, Springer-Verlag, Berlin, 2005. Second edition.

[20] D. W. Stroock and 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, Berkeley, Calif., 1972, University of California Press, pp. 333–359.

[21] D. W. Stroock and S. R. S. Varadhan, *Multidimensional diffusion processes*, vol. 233 of Grundlehren der Mathematischen Wissenschaften, Springer-Verlag, Berlin, 1979.