Support characterization for regular path-dependent stochastic Volterra integral equations

Alexander Kalinin*

August 29, 2019

Abstract

We consider a stochastic Volterra integral equation with regular path-dependent coefficients and a Brownian motion as integrator in a multidimensional setting. Under an imposed absolute continuity condition, the unique solution is a semimartingale that admits almost surely Hölder continuous paths. Based on functional Itô calculus, we prove that the support of its law in the Hölder norm can be described by a flow of mild solutions to ordinary integro-differential equations that are constructed by means of the vertical derivative of the diffusion coefficient.

MSC2010 classification: 60H20, 28C20, 60G17, 45D05, 45J05.
Keywords: support of a measure, path-dependent Volterra process, functional Volterra integral equation, functional Itô calculus, vertical derivative, Hölder space.

1 Support representations via flows

The support of the law of a continuous stochastic process consists of all continuous paths around any neighborhood the process may remain with positive probability. Determining this class of paths for a diffusion process, viewed as solution to a stochastic differential equation (SDE), establishes a relation between the coefficients of the equation and the law of its solution.

In the pioneering work of Stroock and Varadhan [16], the support of the law of a diffusion process is characterized by an associated flow of classical solutions to ordinary differential equations. While Aida [1] generalizes the time-homogeneous case to a Hilbert space, allowing for an infinite dimension, Gyöngy and Pröhle [10] deal with coefficients that are of affine growth and not necessarily bounded. Moreover, Pakkanen [14] provides sufficient conditions for a stochastic integral to have the full support property.

An extension of the Stroock-Varadhan support theorem to any α-Hölder norm, where $\alpha \in (0, 1/2)$, is given in Ben Arous et al. [4]. The case of time-homogeneous coefficients was independently proven by Millet and Sanz-Solé [13] and later extended to a parabolic

*Department of Mathematics, Imperial College London, United Kingdom. alex.kalinin@mail.de.
The author gratefully acknowledges support from Imperial College through a Chapman fellowship.

stochastic partial differential equation (SPDE) in Bally et al. [3]. By using the vertical derivative as functional space derivative and generalizing the approach in [13] with the relevant Girsanov changes of measures, a path-dependent version of the Stroock-Varadhan support theorem in Hölder norms was recently derived in [7]. The contribution of this article is to extend this support characterization to stochastic Volterra integral equations with regular path-dependent coefficients by providing a flow of mild solutions to ordinary integro-differential equations.

Let $r, T \geq 0$ with $r < T$ and $d, m \in \mathbb{N}$. We work with the separable Banach space $C([0,T], \mathbb{R}^m)$ of all $\mathbb{R}^m$-valued continuous paths on $[0,T]$, endowed with the supremum norm given by $\|x\|_\infty = \sup_{t\in[0,T]} |x(t)|$, where $|\cdot|$ is used as absolute value function, Euclidean norm or Hilbert-Schmidt norm. Throughout, $\hat{x} \in C([0,T], \mathbb{R}^m)$ and

$$b : [r,T]^2 \times C([0,T], \mathbb{R}^m) \to \mathbb{R}^m \quad \text{and} \quad \sigma : [r,T]^2 \times C([0,T], \mathbb{R}^m) \to \mathbb{R}^{m\times d}$$

are two product measurable maps that are *non-anticipative* in the sense that they satisfy $b(t,s,x) = b(t,s,x^s)$ and $\sigma(t,s,x) = \sigma(t,s,x^s)$ for all $s,t \in [r,T]$ with $s \leq t$ and each $x \in C([0,T], \mathbb{R}^m)$, where x^s denotes the path x stopped at time s.

On a filtered probability space $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\in[0,T]}, P)$ that satisfies the usual conditions and which possesses a standard d-dimensional $(\mathscr{F}_t)_{t\in[0,T]}$-Brownian motion W, we consider the following path-dependent stochastic Volterra integral equation:

$$X_t = X_r + \int_r^t b(t,s,X)\,ds + \int_r^t \sigma(t,s,X)\,dW_s \quad \text{a.s.} \tag{1.1}$$

for $t \in [r,T]$ with initial condition $X_q = \hat{x}(q)$ for $q \in [0,r]$ a.s. An absolute continuity and affine growth condition on the coefficients b and σ ensure that any solution to (1.1) is a semimartingale with delayed Hölder continuous trajectories.

In fact, for each $\alpha \in (0,1]$ let $C_r^\alpha([0,T], \mathbb{R}^m)$ represent the non-separable Banach space of all $x \in C([0,T], \mathbb{R}^m)$ that are α-Hölder continuous on $[r,T]$, endowed with the *delayed α-Hölder norm* given by

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

By convenience, we set $C_r^0([0,T], \mathbb{R}^m) := C([0,T], \mathbb{R}^m)$ and $\|\cdot\|_{0,r} := \|\cdot\|_\infty$. Then, under the conditions stated below, there is a unique strong solution to (1.1) whose sample paths belong a.s. to the *delayed Hölder space* $C_r^\alpha([0,T], \mathbb{R}^m)$ for any $\alpha \in (0,1/2)$.

For $p \geq 1$ consider the separable Banach space $W_r^{1,p}([0,T], \mathbb{R}^m)$ of all $x \in C([0,T], \mathbb{R}^m)$ that are absolutely continuous on $[r,T]$ with a p-fold Lebesgue-integrable weak derivative $\dot{x}$, equipped with the *delayed Sobolev L^p-norm* defined by

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

Then it holds that $W_r^{1,p}([0,T], \mathbb{R}^m) \subsetneq C_r^{1/q}([0,T], \mathbb{R}^m)$ and $\|x\|_{1/q,r} \leq \|x\|_{1,p,r}$ for all $x \in W_r^{1,p}([0,T], \mathbb{R}^m)$ whenever $p > 1$ and q is its dual exponent. By allowing infinite values, we extend the definitions of $\|\cdot\|_\infty$ and $\|\cdot\|_{\alpha,r}$ at (1.2) to each path $x : [0,T] \to \mathbb{R}^m$ and the definition of $\|\cdot\|_{1,p,r}$ at (1.3) to every $x \in W_r^{1,1}([0,T], \mathbb{R}^m)$.

Based on the non-separable Banach space $D([0,T],\mathbb{R}^m)$ of all $\mathbb{R}^m$-valued càdlàg paths on $[0,T]$, endowed with the supremum norm $\|\cdot\|_\infty$, we use the following pseudometric on $[r,T]\times D([0,T],\mathbb{R}^m)$ given by

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

Then a functional on this Cartesian product that is d_∞-continuous is also non-anticipative and Lipschitz continuity relative to d_∞ merely requires 1/2-Hölder continuity in the time variable.

Let us now state the conditions under which the support theorem holds. By refering to *horizontal* and *vertical differentiability* of non-anticipative functionals from [5, 9], we in particular require that certain time and path space components of σ are of class $\mathbb{C}^{1,2}$, a property to be recalled in Section 2.1. In this context, let ∂_s be the horizontal, ∂_x the vertical and ∂_{xx} the second-order vertical differential operator.

To have a simple notation if these first- and second-order space derivatives appear, we set $\|y\| := (\sum_{k=1}^m \sum_{l=1}^d |y_{k,l}|^2)^{1/2}$ if $y \in (\mathbb{R}^{1\times m})^{m\times d}$ or $y \in (\mathbb{R}^{m\times m})^{m\times d}$. Further, let $\mathbb{I}_d$ be the identity matrix in $\mathbb{R}^{d\times d}$ and A' denote the transpose of a matrix $A \in \mathbb{R}^{d\times m}$.

(C.1) The map $[r,t)\times C([0,T],\mathbb{R}^m) \to \mathbb{R}^{m\times d}$, $(s,x)\mapsto \sigma(t,s,x)$ is of class $\mathbb{C}^{1,2}$ for each $t \in (r,T]$, the maps $b(\cdot,s,x)$ and $\sigma(\cdot,s,x)$ are absolutely continuous on $[s,T]$ and $\partial_x\sigma(\cdot,s,x)$ is absolutely continuous on $(s,T]$ for any $s \in [r,T)$ and each $x \in C([0,T],\mathbb{R}^m)$.

(C.2) The maps σ, $\partial_x\sigma$ and its weak time derivatives $\partial_t\sigma$, $\partial_t\partial_x\sigma$ are bounded. Further, there are $c,\lambda,\eta \geq 0$ and $\kappa \in [0,1)$ such that

$$|b(s,s,x)| + |\partial_t b(t,s,x)| \leq c\big(1+\|x\|_\infty^\kappa\big) \quad \text{and}$$
$$|\partial_s\sigma(t,s,x)| + \|\partial_{xx}\sigma(t,s,x)\| \leq c\big(1+\|x\|_\infty^\eta\big)$$

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

(C.3) There is $\lambda \geq 0$ satisfying $|b(s,s,x)-b(s,s,y)|+|\partial_t b(t,s,x)-\partial_t b(t,s,y)| \leq \lambda\|x-y\|_\infty$ and

$$\begin{aligned}|\sigma(u,t,x)-\sigma(u,s,y)| &+ |\partial_u\sigma(u,t,x)-\partial_u\sigma(u,s,y)| \\ &+\|\partial_x\sigma(u,t,x)-\partial_x\sigma(u,s,y)\| \leq \lambda d_\infty((t,x),(s,y))\end{aligned}$$

for any $s,t,u \in [r,T)$ with $s<t<u$ and every $x,y \in C([0,T],\mathbb{R}^m)$.

Under the assumption that $\sigma(t,\cdot,\cdot)$ is of class $\mathbb{C}^{1,2}$ on $[r,t)\times C([0,T],\mathbb{R}^m)$ for each $t \in (r,T]$, we may introduce the map $\rho : [r,T]^2 \times C([0,T],\mathbb{R}^m) \to \mathbb{R}^m$, which serves as *correction term*, coordinatewise by

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

if $s<t$ and, $\rho_k(t,s,x) := 0$, otherwise. Here, $\{e_1,\ldots,e_d\}$ stands for the standard basis of $\mathbb{R}^d$ and $[r,t)\times C([0,T],\mathbb{R}^m) \to \mathbb{R}^{1\times m}$, $(s,x)\mapsto \partial_x\sigma_{k,l}(t,s,x)$ is the vertical derivative

of the (k,l)-entry of the map $[r,t) \times C([0,T],\mathbb{R}^m) \to \mathbb{R}^{m\times d}$, $(s,x) \mapsto \sigma(t,s,x)$ for each $t \in (r,T]$, every $k \in \{1,\dots,m\}$ and any $l \in \{1,\dots,d\}$.

Finally, to describe the support of the unique strong solution to (1.1) by a flow, we study the following Volterra integral equation associated to any $h \in W_r^{1,p}([0,T],\mathbb{R}^d)$ with $p \geq 1$. Namely,

$$x_h(t) = x_h(r) + \int_r^t (b - (1/2)\rho)(t,s,x_h)\,ds + \int_r^t \sigma(t,s,x_h)\,dh(s) \tag{1.5}$$

for $t \in [r,T]$. By adding $\hat{x}$ as initial condition, the solution x_h lies in the *delayed Sobolev space* $W_r^{1,p}([0,T],\mathbb{R}^m)$, since it can also be viewed as a *mild solution* to an associated ordinary integro-differential equation, as concisely justified in Section 2.2.

Lemma 1.1. *Let* (C.1)-(C.3) *be valid.*

(i) Pathwise uniqueness holds for (1.1) *and there is a unique strong solution X such that $X^r = \hat{x}^r$ a.s. Further, X is a semimartingale and $E[\|X\|_{\alpha,r}^p] < \infty$ for any $\alpha \in [0,1/2)$ and all $p \geq 1$.*

(ii) For any $p \geq 1$ and each $h \in W_r^{1,p}([0,T],\mathbb{R}^d)$, there is a unique solution x_h to (1.5) *satisfying $x_h^r = \hat{x}^r$ and we have $x_h \in W_r^{1,p}([0,T],\mathbb{R}^m)$. Moreover, the flow map*

$$W_r^{1,p}([0,T],\mathbb{R}^d) \to W_r^{1,p}([0,T],\mathbb{R}^m), \quad h \mapsto x_h \tag{1.6}$$

is Lipschitz continuous on bounded sets.

Having clarified matters of uniqueness, existence and regularity, let us now consider the main result of this paper. Namely, a *support characterization* of solutions to (1.1) in delayed Hölder norms.

Theorem 1.2. *Let* (C.1)-(C.3) *hold, $\alpha \in [0,1/2)$ and $p \geq 2$. Then the support of the image measure of the unique strong solution X to* (1.1) *in $C_r^\alpha([0,T],\mathbb{R}^m)$ is the closure of the set of all solutions x_h to* (1.5)*, where $h \in W_r^{1,p}([0,T],\mathbb{R}^d)$. That is,*

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

Example 1.3. Suppose that there are four product measurable maps $K_b, K_\sigma : [r,T]^2 \to \mathbb{R}$, $\overline{b} : [r,T] \times C([0,T],\mathbb{R}^m) \to \mathbb{R}^m$ and $\overline{\sigma} : [r,T] \times C([0,T],\mathbb{R}^m) \to \mathbb{R}^{m\times d}$ such that

$$b(t,s,x) = K_b(t,s)\overline{b}(s,x) \quad \text{and} \quad \sigma(t,s,x) = K_\sigma(t,s)\overline{\sigma}(s,x)$$

for all $s,t \in [r,T]$ and any $x \in C([0,T],\mathbb{R}^m)$ and let the following three conditions hold:

(1) The functions $K_b(\cdot,s)$ and $K_\sigma(\cdot,s)$ are differentiable for each $s \in [r,T)$. Further, K_b, K_σ, $\partial_t K_b$ and $\partial_t K_\sigma$ are bounded.

(2) The map $\overline{\sigma}$ is of class $\mathbb{C}^{1,2}$ on $[r,T) \times C([0,T],\mathbb{R}^m)$ and together with its vertical derivative $\partial_x \overline{\sigma}$ it is bounded and d_∞-Lipschitz continuous.

(3) There are $c, \eta, \lambda \geq 0$ and $\kappa \in [0,1)$ such that

$$|\overline{b}(s,x)| \leq c(1 + \|x\|_\infty^\kappa), \quad |\overline{b}(s,x) - \overline{b}(s,y)| \leq \lambda \|x-y\|_\infty,$$
$$|K_\sigma(u,t) - K_\sigma(u,s)| + |\partial_u K_\sigma(u,t) - \partial_u K_\sigma(u,s)| \leq \lambda |s-t|^{1/2} \quad \text{and}$$
$$|\partial_s \overline{\sigma}(s,x)| + |\partial_{xx}\overline{\sigma}(s,x)| \leq c(1 + \|x\|_\infty^\eta)$$

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

Then Theorem 1.2 applies and in the specific case that $K_b = K_\sigma = 1$ it reduces to the support theorem in [7] with the same regularity conditions.

The structure of this paper is determined by the proof of the support theorem and can be comprised as follows. Section 2 provides supplementary material and a Hölder convergence result that yields Theorem 1.2 as a corollary. In detail, Section 2.1 gives a concise overview of horizontal and vertical differentiability of non-anticipative functionals. Section 2.2 relates the Volterra integral equation (1.5) to an ordinary integro-differential equation and shows that solutions to (1.1) are semimartingales by using a stochastic Fubini theorem. In Section 2.3 we consider the approach to prove the support theorem by introducing a more general setting and stating Theorem 2.3, the before mentioned convergence result.

Section 3 derives relevant estimates to infer convergence in Hölder norm in moment. To be precise, Section 3.1 gives a sufficient condition for a sequence of processes to converge in this sense by exploiting an explicit Kolmogorov-Chentsov estimate. In Section 3.2 we introduce the relevant notations in the context of sequence of partitions and recall a couple of auxiliary moment estimates from [7, 12]. The purpose of Section 3.3 is to deduce moment estimates for deterministic and stochastic Volterra integrals, generalizing the bounds from [7][Lemmas 20, 21 and Proposition 22].

Section 4 is devoted to a variety of specific moment estimates and decompositions, preparing the proof of Theorem 2.3. At first, Section 4.1 derives bounds for solutions to stochastic Volterra integral equations and gives two main decompositions, Proposition 4.3 and (4.7). Section 4.2 handles the first two remainders appearing in (4.7). While the second can be directly estimated, the first relies on the functional Itô formula in [6]. Section 4.3 intends to bound the third remainder in second moment, requiring another extensive decomposition. In Section 5 we prove the convergence result and the support representation, including assertions on uniqueness, existence and regularity.

2 Preparations and a convergence result in second moment

2.1 Differential calculus for non-anticipative functionals

We recall and discuss horizontal and vertical differentiability, as introduced in [5, 9]. To this end, let $t \in (r, T]$ and G be a non-anticipative functional on $[r,t) \times D([0,T], \mathbb{R}^m)$ that is considered at a point (s,x) of its domain:

(i) G is *horizontally differentiable* at (s,x) if the function $[0, T-s) \to \mathbb{R}$, $h \mapsto G(s+h, x^s)$ is differentiable at 0. If this is the case, then $\partial_s G(s,x)$ denotes its derivative there.

(ii) G is *vertically differentiable* at (s,x) if the function $\mathbb{R}^m \to \mathbb{R}$, $h \mapsto G(s, x + h\mathbb{1}_{[s,T]})$ is differentiable at 0. In this case, its derivative there is denoted by $\partial_x G(s,x)$.

(iii) G is *partially vertically differentiable* at (s,x) if for any $k \in \{1,\dots,m\}$ the function $\mathbb{R} \to \mathbb{R}$, $h \mapsto G(s, x + h\overline{e}_k\mathbb{1}_{[s,T]})$ is differentiable at 0, where $\{\overline{e}_1,\dots,\overline{e}_m\}$ is the standard basis of $\mathbb{R}^m$. In this event, $\partial_{x_k} G(s,x)$ represents its derivative there.

So, G is horizontally, vertically or partially vertically differentiable if it satisfies the respective property at any point of its domain. We observe that vertical differentiability entails partial vertical differentiability and $\partial_x G = (\partial_{x_1} G, \dots, \partial_{x_m} G)$.

We say that G is twice vertically differentiable if it is vertically differentiable and the same is true for $\partial_x G$. We then set $\partial_{xx} G := \partial_x(\partial_x G)$ and $\partial_{x_k x_l} G := \partial_{x_k}(\partial_{x_l} G)$ for any $k,l \in \{1,\dots,m\}$. If in addition $\partial_{xx} G$ is d_∞-continuous, then

$$\partial_{x_k x_l} G = \partial_{x_l x_k} G \quad \text{for each } k,l \in \{1,\dots,m\},$$

by Schwarz's Lemma, showing that $\partial_{xx} G$ is symmetric. Moreover, we call G *of class* $\mathbb{C}^{1,2}$ if it is once horizontally and twice vertically differentiable such that G, $\partial_s G$, $\partial_x G$ and $\partial_{xx} G$ are bounded on bounded sets and d_∞-continuous.

Clearly, horizontal differentiability applies to functionals on $[r,t) \times C([0,T],\mathbb{R}^m)$ as well by considering continuous paths only. Vertical differentiability, however, requires the evaluation along càdlàg paths. So, a functional F on $[r,t) \times C([0,T],\mathbb{R}^m)$ is of *class* $\mathbb{C}^{1,2}$ if it possesses an non-anticipative extension $G : [r,t) \times D([0,T],\mathbb{R}^m) \to \mathbb{R}$ that satisfies this property. Then the restricted derivatives

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

are well-defined, by Theorems 5.4.1 and 5.4.2 in [2]. That is, they do not dependent on the choice of the extension G. By combining these considerations with an absolute continuity condition, which ensures that only semimartingales appear, we can use the functional Itô formula from [6] to prove Proposition 4.4, a key ingredient when deriving (1.7).

Examples 2.1. (i) We suppose that $\alpha \in (0,1]$, $k \in \mathbb{N}$ and $\varphi : [r,t) \times (\mathbb{R}^m)^k \to \mathbb{R}^d$, $(s,x) \mapsto \varphi(s,\overline{x}_1,\dots,\overline{x}_m)$ is α-Hölder continuous. Let $t_0,\dots,t_k \in [r,t)$ satisfy $t_0 < \cdots < t_k$, then the $\mathbb{R}^d$-valued non-anticipative map G on $[r,t) \times D([0,T],\mathbb{R}^m)$ given by

$$G(s,x) := \varphi(s, x(t_0 \wedge s), \dots, x(t_k \wedge s))$$

is bounded on bounded sets and α-Hölder continuous with respect to d_∞. Furthermore, if φ is of class $C^{1,2}$ in the usual sense, then G is of class $\mathbb{C}^{1,2}$, because it satisfies $\partial_s G(s,x)$ $= (\partial_+\varphi/\partial s)(s, x(t_0 \wedge s), \dots, x(t_k \wedge s))$ and

$$\partial_x G(s,x) = \sum_{j=0,\, s \le t_j}^{k} D_{\overline{x}_j}\varphi(s, x(t_0 \wedge s), \dots, x(t_k \wedge s))$$

for any $s \in [r,t)$ and every $x \in D([0,T],\mathbb{R}^m)$, where $\partial_+\varphi/\partial s$ denotes the right-hand time derivative of φ and $D_{\overline{x}_j}\varphi$ the partial derivative of φ with respect to the j-th space variable $\overline{x}_j \in \mathbb{R}^m$ for each $j \in \{1,\dots,k\}$.

(ii) Let $\alpha \in (0,1]$, $K : [0,t) \to \mathbb{R}$ be continuously differentiable and φ be an $\mathbb{R}^{m\times d}$-valued Borel measurable bounded map on $[0,t) \times D([0,T],\mathbb{R}^m)$ that is α-Hölder continuous in $x \in D([0,T],\mathbb{R}^m)$, uniformly in $s \in [0,t)$. Then the *non-anticipative kernel integral map* $G : [r,t) \times D([0,T],\mathbb{R}^d) \to \mathbb{R}^{m\times d}$ defined by

$$G(s,x) := \int_0^s K(s-u)\varphi(u,x^u)\,du$$

is bounded and α-Hölder continuous relative to d_∞. In addition, if φ is d_∞-continuous, then G is of class $\mathbb{C}^{1,2}$, since $\partial_s G(s,x) = K(0)\varphi(s,x) + \int_0^s \dot{K}(s-u)\varphi(u,x)\,du$ for each $s \in [r,t)$ and any $x \in D([0,T],\mathbb{R}^m)$ and $\partial_x G = 0$.

2.2 Ordinary integro-differential equations and semimartingales

By utilizing an absolute continuity condition, we directly connect the Volterra integral equation (1.5) to an ordinary integro-differential equation and check that any solution to (1.1) solves a stochastic differential equation, ensuring that it is a semimartingale.

Let us first briefly analyze (1.5) for $h \in W_r^{1,1}([0,T],\mathbb{R}^d)$, under the hypothesis that $\sigma(t,\cdot,\cdot)$ is of class $\mathbb{C}^{1,2}$ on $[r,t) \times C([0,T],\mathbb{R}^m)$ for each $t \in (r,T]$. A *solution* to (1.5) is a path $x \in C([0,T],\mathbb{R}^m)$ such that

$$\int_r^t |(b-(1/2)\rho)(t,s,x)| + |\sigma(t,s,x)||\dot{h}(s)|\,ds \quad \text{and}$$
$$x(t) = x(r) + \int_r^t (b-(1/2)\rho)(t,s,x)\,ds + \int_r^t \sigma(t,s,x)\,dh(s)$$

for any $t \in [r,T]$, since the variation of h on $[r,s]$ is given by $\int_r^s |\dot{h}(u)|\,du$ for all $s \in [r,t]$. If we now assume that (C.1)-(C.3) are valid, then the d_∞-Lipschitz continuity of the map $[r,t) \times C([0,T],\mathbb{R}^m) \to \mathbb{R}^{1\times m}$, $(s,x) \mapsto \partial_x \sigma_{k,l}(t,s,x)$ entails that it admits a unique continuous extension to $[r,t] \times C([0,T],\mathbb{R}^m)$ for any $t \in (r,T]$, each $k \in \{1,\ldots,m\}$ and every $l \in \{1,\ldots,d\}$.

In this case, we may define $\overline{\rho} : [r,T]^2 \times C([0,T],\mathbb{R}^m) \to \mathbb{R}^m$ coordinatewise by letting $\overline{\rho}_k(t,s,x)$ agree with the right-hand side in (1.4), if $s \leq t$, and setting $\overline{\rho}(t,s,x) := 0$, otherwise. Then Fubini's theorem entails for each $x \in C([0,T],\mathbb{R}^m)$ that

$$\begin{aligned}&\int_r^t (b-(1/2)\rho)(t,s,x)\,ds + \int_r^t \sigma(t,s,x)\,dh(s)\\ &= \int_r^t (b-(1/2)\overline{\rho}+\sigma\dot{h})(s,s,x) + \int_r^s \partial_s(b-(1/2)\overline{\rho}+\sigma\dot{h})(s,u,x)\,du\,ds \end{aligned} \tag{2.1}$$

for every $t \in [r,T]$. Consequently, the path x solves (1.5) if and only if it is a *mild solution* to the path-dependent ordinary integro-differential equation

$$\dot{x}(t) = (b-(1/2)\overline{\rho}+\sigma\dot{h})(t,t,x) + \int_r^t \partial_t(b-(1/2)\overline{\rho}+\sigma\dot{h})(t,s,x)\,ds$$

for $t \in [r,T]$. Since all appearing maps are integrable, this means that the increment $x(t)-x(r)$ agrees with (2.1) for any $t \in [r,T]$. Let us now turn to the stochastic Volterra integral equation (1.1), without imposing any conditions for the moment.

Thus, we let $\mathscr{C}([0,T],\mathbb{R}^m)$ denote the completely metrizable topological space of all $(\mathscr{F}_t)_{t\in[0,T]}$-adapted continuous processes $X:[0,T]\times\Omega\to\mathbb{R}^m$ and recall that a *solution* to (1.1) is a process $X\in\mathscr{C}([0,T],\mathbb{R}^m)$ such that

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

For a process $\xi\in\mathscr{C}([0,T],\mathbb{R}^m)$ that is independent of W we let $(\mathscr{E}_t^0)_{t\in[0,T]}$ be the natural filtration of the adapted continuous process $[0,T]\times\Omega\to\mathbb{R}^{2m}$, $(t,\omega)\mapsto(\xi_t^r, W_{r\vee t}-W_r)(\omega)$. That is, $\mathscr{E}_t^0=\sigma(\xi_q : q\in[0,t])$ for $t\in[0,r]$ and

$$\mathscr{E}_t^0 := \mathscr{E}_r^0 \vee \sigma(W_s - W_r : s\in[r,t]) \quad \text{for } t\in(r,T].$$

In particular, $\mathscr{E}_t^0=\sigma(\xi_0)\vee\sigma(W_s : s\in[0,t])$ for all $t\in[0,T]$ if there is no delay. Let $(\mathscr{E}_t)_{t\in[0,T]}$ be the right-continuous filtration of the augmented filtration of $(\mathscr{E}_t^0)_{t\in[0,T]}$. Then a solution X to (1.1) satisfying $X^r=\xi^r$ a.s. is called *strong* if it is adapted to this complete filtration.

Finally, suppose that (C.1) and (C.2) hold. Then it follows from Fubini's theorem for stochastic integrals, stated in [17][Theorem 2.2] for instance, that any $X\in\mathscr{C}([0,T],\mathbb{R}^m)$ satisfies

$$\int_r^t b(t,s,X)\,ds + \int_r^t \sigma(t,s,X)\,dW_s = \int_r^t B_s(X)\,ds + \int_r^t \sigma(s,s,X)\,dW_s$$

a.s. for any $t\in[r,T]$, where the map $B:[r,T]\times\Omega\times\mathscr{C}([0,T],\mathbb{R}^m)\to\mathbb{R}^m$, which is product measurable and depends on whole processes rather than trajectories, is given by

$$B_s(Y) = b(s,s,Y) + \int_r^s \partial_s b(s,u,Y)\,du + \int_r^s \partial_s\sigma(s,u,Y)\,dW_u$$

for every $s\in[r,T]$ a.s. This shows that X solves (1.1) if and only if it is a solution to the path-dependent stochastic differential equation

$$X_t = B_t(X)\,dt + \sigma(t,t,X)\,dW_t \quad \text{for } t\in[r,T].$$

Moreover, it is automatically a semimartingale in this case.

2.3 Approach to the main result in a general setting

After these preliminary considerations, we proceed as follows to establish the support theorem. For any $n\in\mathbb{N}$ let $\mathbb{T}_n$ be a partition of $[r,T]$ of the form $\mathbb{T}_n=\{t_{0,n},\dots,t_{k_n,n}\}$ with $k_n\in\mathbb{N}$ and $t_{0,n},\dots,t_{k_n,n}\in[r,T]$ such that $r=t_{0,n}<\cdots<t_{k_n,n}=T$ and whose mesh $\max_{i\in\{0,\dots,k_n-1\}}(t_{i+1,n}-t_{i,n})$ is denoted by $|\mathbb{T}_n|$. We assume that the sequence $(\mathbb{T}_n)_{n\in\mathbb{N}}$ of partitions is *balanced* as defined in [8], which means that there is $c_{\mathbb{T}}\geq 1$ such that

$$|\mathbb{T}_n| \leq c_{\mathbb{T}} \min_{i\in\{0,\dots,k_n-1\}}(t_{i,n}-t_{i-1,n}) \quad \text{for all } n\in\mathbb{N}. \tag{2.2}$$

For the estimation of one term in Proposition 4.4, when the functional Itô formula is applied, we also require the following additional condition:

(C.4) There is $\overline{c}_{\mathbb{T}} > 0$ such that $k_n|\mathbb{T}_n| \leq \overline{c}_{\mathbb{T}}$ for each $n \in \mathbb{N}$.

However, unless explicitly stated, we shall not impose this condition. Moreover, we readily notice that any equidistant sequence of partitions satisfies both conditions.

Next, for any $k, n \in \mathbb{N}$ we are interested in the delayed linear interpolation of a map $x : [0,T] \to \mathbb{R}^k$ along $\mathbb{T}_n$. Namely, we define $L_n(x) : [0,T] \to \mathbb{R}^k$ by $L_n(x)(t) := x(r \wedge t)$, if $t \leq t_{1,n}$, and

$$L_n(x)(t) := x(t_{i-1,n}) + \frac{t - t_{i,n}}{t_{i+1,n} - t_{i,n}}(x(t_{i,n}) - x(t_{i-1,n})), \tag{2.3}$$

if $t \in (t_{i,n}, t_{i+1,n}]$ for some $i \in \{1, \ldots, k_n - 1\}$. Since $L_n(x)$ is piecewise continuously differentiable, it belongs to $W_r^{1,p}([0,T], \mathbb{R}^k)$ for every $p \geq 1$, and by construction, the process ${}_nW : [0,T] \times \Omega \to \mathbb{R}^d$ defined via ${}_nW_t := L_n(W)(t)$ is adapted.

Let us now assume that (C.1)-(C.3) and Lemma 1.1 hold. Then the support of $P \circ X^{-1}$ is included in the closure of $\{x_h \,|\, h \in W_r^{1,p}([0,T], \mathbb{R}^d)\}$ in $C_r^\alpha([0,T], \mathbb{R}^m)$ for $\alpha \in [0, 1/2)$ and $p \geq 2$ if we can prove that

$$\lim_{n \uparrow \infty} P(\|x_{{}_nW} - X\|_{\alpha,r} \geq \varepsilon) = 0 \quad \text{for any } \varepsilon > 0. \tag{2.4}$$

Moreover, if for each $h \in W_r^{1,p}([0,T], \mathbb{R}^d)$ there exists a sequence $(P_{h,n})_{n\in\mathbb{N}}$ of probability measures on $(\Omega, \mathscr{F})$ that are absolutely continuous to P such that

$$\lim_{n \uparrow \infty} P_{h,n}(\|X - x_h\|_{\alpha,r} \geq \varepsilon) = 0 \quad \text{for every } \varepsilon > 0, \tag{2.5}$$

then the converse inclusion holds. The sufficiency of (2.4) and (2.5) follows from a basic result on the support of probabiilty measures, see [7][Lemma 36] for example. To verify the validity of both limits, we consider a more general setting.

Let $\underline{B}$ be an $\mathbb{R}^m$-valued and $B_H, \overline{B}$ and Σ be $\mathbb{R}^{m\times d}$-valued non-anticipative product measurable maps on $[r,T]^2 \times C([0,T], \mathbb{R}^m)$. For any $n \in \mathbb{N}$ we study the path-dependent stochastic Volterra integral equation:

$$\begin{aligned} {}_nY_t = {}_nY_r + \int_r^t \underline{B}(t,s,{}_nY) + B_H(t,s,{}_nY)\dot{h}(s) + \overline{B}(t,s,{}_nY){}_n\dot{W}_s\,ds \\ + \int_r^t \Sigma(t,s,{}_nY)\,dW_s \quad \text{a.s. for } t \in [r,T]. \end{aligned} \tag{2.6}$$

Provided that the map $[r,t) \times C([0,T], \mathbb{R}^m) \to \mathbb{R}^{m\times d}$, $(s,x) \mapsto \overline{B}(t,s,x)$ is of class $\mathbb{C}^{1,2}$ for all $t \in (r,T]$, we introduce another path-dependent stochastic Volterra integral equation:

$$\begin{aligned} Y_t = Y_r + \int_r^t (\underline{B} + R)(t,s,Y) + B_H(t,s,Y)\dot{h}(s)\,ds \\ + \int_r^t (\overline{B} + \Sigma)(t,s,Y)\,dW_s \quad \text{a.s. for } t \in [r,T] \end{aligned} \tag{2.7}$$

with the $\mathbb{R}^m$-valued non-anticipative product measurable map R on $[r,T]^2 \times C([0,T], \mathbb{R}^m)$ given coordinatewise by

$$R_k(t,s,x) = \sum_{l=1}^d \partial_x \overline{B}_{k,l}(t,s,x)((1/2)\overline{B} + \Sigma)(s,s,x)e_l, \tag{2.8}$$

if $s < t$, and $R_k(t,s,x) := 0$, otherwise. In particular, (2.6) reduces to (2.7) in the case that $\overline{B} = 0$. We seek to show that if ${}_nY$ and Y are two continuous solutions to (2.6) and (2.7), respectively, satisfying ${}_nY^r = Y^r = \hat{x}^r$ a.s. for all $n \in \mathbb{N}$, then

$$\lim_{n\uparrow\infty} E[\|{}_nY - Y\|_{\alpha,r}^2] = 0. \tag{2.9}$$

Thus, by choosing $\underline{B} = b - (1/2)\rho$, $B_H = 0$, $\overline{B} = \sigma$ and $\Sigma = 0$, we obtain (2.4). If instead $\underline{B} = b$, $B_H = \sigma$, $\overline{B} = -\sigma$ and $\Sigma = \sigma$, then (2.5) is implied, as we will see. To derive the general convergence result (2.9), we introduce the following regularity conditions:

(C.5) The map $[r,t) \times C([0,T],\mathbb{R}^m) \to \mathbb{R}^{m\times d}$, $(s,x) \mapsto \overline{B}(t,s,x)$ is of class $\mathbb{C}^{1,2}$ for all $t \in (r,T]$, for any $F \in \{\underline{B}, B_H, \overline{B}, \Sigma\}$ the map $F(\cdot,s,x)$ is absolutely continuous on $[s,T]$ and $\partial_x\overline{B}$ is absolutely continuous on $(s,T]$ for each $s \in [r,T)$ and any $x \in C([0,T],\mathbb{R}^m)$.

(C.6) There are $c \geq 0$ and $\kappa \in [0,1)$ such that any two maps $F \in \{\underline{B}, B_H\}$ and $G \in \{\overline{B}, \Sigma\}$ satisfy $|F(s,s,x)| + |\partial_t F(t,s,x)| \leq c(1 + \|x\|_\infty^\kappa)$ and

$$|G(s,s,x)| + |\partial_t G(t,s,x)| \leq c$$

for each $s,t \in [r,T)$ with $s < t$ and every $x \in C([0,T],\mathbb{R}^m)$.

(C.7) There exists $\lambda \geq 0$ such that $|\underline{B}(s,s,x) - \underline{B}(s,s,y)| + |\partial_t\underline{B}(t,s,x) - \partial_t\underline{B}(t,s,y)| \leq \lambda\|x-y\|_\infty$ and for any $F \in \{B_H, \overline{B}, \Sigma\}$ it holds that

$$|F(u,t,x) - F(u,s,y)| + |\partial_u F(u,t,x) - \partial_u F(u,s,y)| \leq \lambda d_\infty((t,x),(s,y))$$

for each $s,t,u \in [r,T)$ with $s < t < u$ and every $x,y \in C([0,T],\mathbb{R}^m)$.

(C.8) There are $\overline{c}, \eta, \overline{\lambda} \geq 0$ such that $\|\partial_x\overline{B}(s,s,x)\| + \|\partial_t\partial_x\overline{B}(t,s,x)\| \leq \overline{c}$, $|\partial_s\overline{B}(t,s,x)| + \|\partial_{xx}\overline{B}(t,s,x)\| \leq \overline{c}(1 + \|x\|_\infty^\eta)$ and

$$\|\partial_x\overline{B}(u,t,x) - \partial_x\overline{B}(u,s,y)\| \leq \overline{\lambda} d_\infty((t,x),(s,y))$$

for any $s,t,u \in [r,T)$ with $s < t < u$ and each $x \in C([0,T],\mathbb{R}^m)$.

(C.9) There exist $\overline{b}_0 \in \mathbb{R}$ and a measurable function $\overline{b} : [r,T] \to \mathbb{R}$ such that $\int_r^T \overline{b}_1(s)^2\,ds < \infty$ and $\overline{b}_0\overline{B}(t,s,x) = \overline{b}(s)\Sigma(t,s,x)$ for every $s,t \in [r,T)$ with $s < t$ and each $x \in C([0,T],\mathbb{R}^m)$.

First, we question uniqueness, existence and regularity of solutions to (2.6) and (2.7). In this regard, let $\xi \in \mathscr{C}([0,T],\mathbb{R}^m)$ and $({}_n\xi)_{n\in\mathbb{N}}$ be a sequence in $\mathscr{C}([0,T],\mathbb{R}^m)$.

Lemma 2.2. *Assume that* (C.5)-(C.7) *are satisfied,* $h \in W_r^{1,2}([0,T],\mathbb{R}^d)$ *and for each* $n \in \mathbb{N}$ *there is is* $p > 2$ *such that* $E[\|\xi^r\|_\infty^p + \|{}_n\xi^r\|_\infty^p] < \infty$.

(i) Under (C.9)*, pathwise uniqueness holds for* (2.6) *and there exists a unique strong solution* ${}_nY$ *with* ${}_nY^r = {}_n\xi^r$ *a.s. for any* $n \in \mathbb{N}$*. Further, for each* $p > 2$ *and every* $\alpha \in [0, 1/2 - 1/p)$*, there is* $c_{\alpha,p} > 0$ *such that*

$$E[\|{}_nY\|_{\alpha,r}^p] \leq c_{\alpha,p}(1 + E[\|{}_n\xi^r\|_\infty^p]) \quad \textit{for all } n \in \mathbb{N}.$$

(ii) If (C.8) *holds, then we have pathwise uniqueness for* (2.7) *and a unique strong solution Y with $Y^r = \xi^r$ a.s. In this case, for each $p > 2$ and all $\alpha \in [0, 1/2 - 1/p)$ there is $\overline{c}_{\alpha,p} > 0$ with $E[\|Y\|_{\alpha,r}^p] \leq \overline{c}_{\alpha,p}(1 + E[\|\xi^r\|_\infty^p])$.*

Finally, we consider a convergence result in Hölder norm in second moment.

Theorem 2.3. *Let* (C.4)-(C.9) *hold, $h \in W_r^{1,2}([0,T],\mathbb{R}^d)$ and $\alpha \in [0, 1/2)$. Suppose that $\lim_{n\uparrow\infty} E[\|{}_n\xi^r - \xi^r\|_\infty^2]/|\mathbb{T}_n|^{2\alpha} = 0$ and there is $p > 2$ such that*

$$\alpha < 1/2 - 1/p \quad \text{and} \quad E[\|\xi^r\|_\infty^p] + \sup_{n\in\mathbb{N}} E[\|{}_n\xi^r\|_\infty^{(2\vee\eta)p}] < \infty.$$

Let ${}_nY$ and Y be the unique strong solutions to (2.6) *and* (2.7)*, respectively, such that ${}_nY^r = {}_n\xi^r$ and $Y^r = \xi^r$ a.s. for all $n \in \mathbb{N}$, then*

$$\lim_{n\uparrow\infty} E\big[\max_{j\in\{0,\dots,k_n\}} |{}_nY_{t_{j,n}} - Y_{t_{j,n}}|^2\big]/|\mathbb{T}_n|^{2\alpha} = 0. \tag{2.10}$$

In particular, (2.9) *is satisfied. That is, $({}_nY)_{n\in\mathbb{N}}$ converges in the norm $\|\cdot\|_{\alpha,r}$ in second moment to Y.*

3 Estimates for convergence in Hölder norm in moment

3.1 Convergence in moment along a sequence of partitions

We consider a sufficient condition for a sequence of processes to convergence in the norm $\|\cdot\|_{\alpha,r}$ in p-th moment, where $\alpha \in [0,1]$ and $p \geq 1$. Its derivation relies on an explicit Kolmogorov-Chentsov estimate [7][Proposition 12].

Namely, let X be an $\mathbb{R}^m$-valued right-continuous processes for which there are $c_0 \geq 0$, $p \geq 1$ and $q > 0$ such that $E[|X_s - X_t|^p] \leq c_0|s-t|^{1+q}$ for all $s,t \in [r,T]$. Then it follows that

$$E\bigg[\sup_{s,t\in[r,T]:\, s\neq t} \frac{|X_s - X_t|^p}{|s-t|^{\alpha p}}\bigg] \leq k_{\alpha,p,q}c_0(T-r)^{1+q-\alpha p} \tag{3.1}$$

for any $\alpha \in [0, q/p)$ with $k_{\alpha,p,q} := 2^{p+q}(2^{q/p-\alpha} - 1)^{-p}$. In particular, if $q \leq p$, then X itself, and not necessarily a modification, admits a.s. α-Hölder continuous paths on $[r,T]$.

Lemma 3.1. *Let $({}_nX)_{n\in\mathbb{N}}$ be a sequence of $\mathbb{R}^m$-valued right-continuous processes for which there are $c_0 \geq 0$, $p \geq 1$ and $q > 0$ with $q \leq p$ such that*

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

for all $n \in \mathbb{N}$, each $j \in \{0,\dots,k_n - 1\}$ and any $s,t \in [t_{j,n}, t_{j+1,n}]$. If $(\|{}_nX^r\|_\infty)_{n\in\mathbb{N}}$ and $(\max_{j\in\{1,\dots,k_n\}} |{}_nX_{t_{j,n}}|/|\mathbb{T}_n|^\alpha)_{n\in\mathbb{N}}$ converge in p-th moment to zero, then so does the sequence $(\|{}_nX\|_{\alpha,r})_{n\in\mathbb{N}}$ for every $\alpha \in [0, q/p)$.

Proof. For given $n \in \mathbb{N}$ a case distinction yields that

$$\sup_{s,t\in[r,T]:\, s\neq t} \frac{|{}_nX_s - {}_nX_t|}{|s-t|^\alpha} \leq 2 \max_{j\in\{0,\dots,k_n-1\}} \sup_{s,t\in[t_{j,n},t_{j+1,n}]:\, s\neq t} \frac{|{}_nX_s - {}_nX_t|}{|s-t|^\alpha}$$

$$+ \max_{i,j\in\{1,\dots,k_n\}:\, i\neq j} \frac{|{}_nX_{t_{i,n}} - {}_nX_{t_{j,n}}|}{|t_{i,n} - t_{j,n}|^\alpha}.$$

By virtue of the Kolmogorov-Chentsov estimate (3.1), it holds that

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

since $q > \alpha p$ and $\sum_{j=0}^{k_n-1}(t_{j+1,n} - t_{j,n}) = T - r$. Moreover, from condition (2.2) we infer that $|t_{i,n} - t_{j,n}| \geq |\mathbb{T}_n|/c_\mathbb{T}$ for all $i, j \in \{0,\dots,k_n\}$ with $i \neq j$. Hence,

$$E\Big[\max_{i,j\in\{1,\dots,k_n\}:\, i\neq j} \frac{|{}_nX_{t_{i,n}} - {}_nX_{t_{j,n}}|^p}{|t_{i,n} - t_{j,n}|^{\alpha p}}\Big] \leq 2^{p-1} c_\mathbb{T}^{\alpha p} E\big[\max_{j\in\{1,\dots,k_n\}} |{}_nX_{t_{j,n}}|^p\big]/|\mathbb{T}_n|^{\alpha p}$$

and the claim follows from the definition of the norm $\|\cdot\|_{\alpha,r}$. □

3.2 Sequential notation and auxiliary moment estimates

Let us introduce relevant notations related to the sequence of partitions $(\mathbb{T}_n)_{n\in\mathbb{N}}$. For fixed $n \in \mathbb{N}$ and $t \in [r,T)$, we choose $i \in \{0,\dots,k_n - 1\}$ such that $t \in [t_{i,n}, t_{i+1,n})$ and set

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

Verbalized, $\underline{t}_n$ is the predecessor of t_n relative to $\mathbb{T}_n$, provided $i \neq 0$, and $\overline{t}_n$ is the successor of t_n. For the sake of completeness, let $\underline{T}_n := t_{k_{n-1},n}$, $T_n := T$ and $\overline{T}_n := T$. Further, for $i \in \{0,\dots,k_n\}$ we set

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

For $p \geq 1$ we recall an interpolation error estimate in supremum for stochastic processes in p-th moment and an explicit integral moment estimate for the sequence $({}_nW)_{n\in\mathbb{N}}$ of adapted linear interpolations of W from [7][Lemmas 19 and 17].

(i) Let $({}_nX)_{n\in\mathbb{N}}$ be a sequence of $\mathbb{R}^m$-valued right-continuous processes for which there are $c_0 \geq 0$ and $q > 0$ such that $E[|{}_nX_s - {}_nX_t|^p] \leq c_0|s-t|^{1+q}$ for all $n \in \mathbb{N}$, each $j \in \{0,\dots,k_n - 1\}$ and every $s,t \in [t_{j,n}, t_{j+1,n}]$. Then there is $c_{p,q} > 0$ such that

$$E[\|L_n({}_nX) - {}_nX\|_\infty^p] \leq c_{p,q} c_0 |\mathbb{T}_n|^q \tag{3.2}$$

for all $n \in \mathbb{N}$. To be precise, $c_{p,q} = 2^{p-1}(1 + k_{0,p,q})(T-r)$.

(ii) Let Z be an $\mathbb{R}^d$-valued random vector satisfying $Z \sim \mathcal{N}(0, \mathbb{I}_d)$. Then the constant $\hat{w}_{p,q} := E[|Z|^{pq}]c_\mathbb{T}^{pq}$ satisfies

$$E\Big[\Big(\int_s^t |{}_n\dot{W}_u|^q\, du\Big)^p\Big] \leq \hat{w}_{p,q}|\mathbb{T}_n|^{-pq/2}(t-s)^p \tag{3.3}$$

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

Next, we let $p \geq 2$ and state a Burkholder-Davis-Ghundy inequality for stochastic integrals with respect to W from [12][Theorem 7.2]. Based on this bound, one can deduce an estimate for integrals relative to ${}_nW$ that is independent of $n \in \mathbb{N}$ and which is given in [7][Proposition 16].

(iii) For each $\mathbb{R}^{m\times d}$-valued progressively measurable process X for which $\int_r^T E[|X_u|^p]\,du$ is finite,

$$E\bigg[\sup_{v\in[s,t]}\bigg|\int_s^v X_u\,dW_u\bigg|^p\bigg] \leq w_p(t-s)^{p/2-1}\int_s^t E[|X_u|^p]\,du \tag{3.4}$$

for all $s,t \in [r,T]$ with $s \leq t$, where $w_p := ((p^3/2)/(p-1))^{p/2}$.

(iv) Any $\mathbb{R}^{m\times d}$-valued progressively measurable process X satisfies

$$E\bigg[\max_{v\in[s,t]}\int_s^v X_{\underline{u}_n}\,d_nW_u\bigg|^p\bigg] \leq \hat{w}_p(t-s)^{p/2}\max_{j\in\{0,\dots,k_n\}:\,t_{j,n}\in[\underline{s}_n,\underline{t}_n]} E[|X_{t_{j,n}}|^p] \tag{3.5}$$

for each $s,t \in [r,T]$ with $s \leq t$ with $\hat{w}_p := 3^p w_p c_{\mathbb{T}}^{p/2}$.

3.3 Moment estimates for Volterra integrals

The first integral bound that we consider follows from the auxiliary estimate (3.3).

Lemma 3.2. *Let $p > 1$ and assume for each $n \in \mathbb{N}$ that ${}_nX : [0,T]\times[0,T]\times\Omega \to \mathbb{R}_+$, $(t,s,\omega) \mapsto X_{t,s}(\omega)$ is a product measurable function. If there are $\overline{p} > p$, $c_{\overline{p}} > 0$ and $q \geq \overline{p}/2$ such that*

$$E\bigg[\max_{j\in\{0,\dots,k_n\}}\int_r^{t_{j,n}} {}_nX_{t_{j,n},s}^{\overline{p}}\,ds\bigg] \leq c_{\overline{p}}|\mathbb{T}_n|^q \quad \textit{for all } n \in \mathbb{N}. \tag{3.6}$$

Then there is $c_p > 0$ such that

$$E\bigg[\max_{j\in\{0,\dots,k_n\}}\bigg(\int_r^{t_{j,n}} X_{t_{j,n},s}|{}_n\dot{W}_s|\,ds\bigg)^p\bigg] \leq c_p|\mathbb{T}_n|^{p(q/\overline{p}-1/2)} \quad \textit{for any } n \in \mathbb{N}.$$

Proof. Let q_1 and q_2 denote the dual exponents of p and $\overline{p}/p$, respectively. Then two applications of Hölder's inequality yield that

$$E\bigg[\max_{j\in\{0,\dots,k_n\}}\bigg(\int_r^{t_{j,n}} X_{t_{j,n},s}|{}_n\dot{W}_s|\,ds\bigg)^p\bigg]$$
$$\leq \Bigg(E\bigg[\max_{j\in\{0,\dots,k_n\}}\bigg(\int_r^{t_{j,n}} {}_nX_{t_{j,n},s}^p\,ds\bigg)^{\overline{p}/p}\bigg]\Bigg)^{p/\overline{p}} c_{p,1}|\mathbb{T}_n|^{-p/2}$$

with $c_{p,1} := \hat{w}_{pq_2/q_1,q_1}^{1/q_2}(T-r)^{p/q_1}$, where $\hat{w}_{pq_2/q_1,q_1}$ is the constant introduced at (3.3). For this reason, the constant $c_p := (T-r)^{1-p/\overline{p}}c_{\overline{p}}^{p/\overline{p}}c_{p,1}$ satisfies the desired estimate. ☐

Remark 3.3. For any $n \in \mathbb{N}$ let ${}_nX$ be independent of the first time variable, that is, there is an $\mathbb{R}_+$-valued measurable process ${}_nY$ with ${}_nX_{t,s} = {}_nY_s$ for all $s,t \in [0,T]$. Then for condition (3.6) to hold, it suffices that there is $\overline{c}_{\overline{p}} > 0$ so that $E[{}_nY_s^{\overline{p}}] \leq \overline{c}_{\overline{p}}|\mathbb{T}_n|^q$ for every $s \in [r,T)$ and each $n \in \mathbb{N}$.

For the second and various other estimates in the following section, let us use for each $n \in \mathbb{N}$ the function $\gamma_n : [r,T] \to [0, c_{\mathbb{T}}]$ defined by

$$\gamma_n(s) := \frac{\Delta s_n}{\Delta \overline{s}_n}. \tag{3.7}$$

Put differently, $\gamma_n = \Delta t_{i,n}/\Delta t_{i+1,n}$ on $[t_{i,n}, t_{i+1,n})$ for all $i \in \{0, \ldots, k_n - 1\}$ and $\gamma_n(T) = 1$.

Lemma 3.4. *Assume that $F : [r,T]^2 \times C([0,T], \mathbb{R}^m) \to \mathbb{R}^m$ is a non-anticipative product measurable map for which there are $\lambda_0, c_0 \geq 0$ such that*

$$|F(u,t,x) - F(u,s,x)| \leq \lambda_0 d_\infty((t,x),(s,x)) \quad \text{and} \quad |F(t,s,x)| \leq c_0(1 + \|x\|_\infty)$$

for all $s,t,u \in [r,T]$ with $s < t < u$ and each $x \in C([0,T], \mathbb{R}^m)$. Further, let $({}_nY)_{n \in \mathbb{N}}$ be a sequence in $\mathscr{C}([0,T], \mathbb{R}^m)$ which there are $p \geq 1$ and $c_{p,0} \geq 0$ such that

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

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

$$E\bigg[\max_{j \in \{0,\ldots,k_n\}} \bigg| \int_r^{t_{j,n}} F(t_{j,n}, \underline{s}_n, {}_nY)(\gamma_n(s) - 1)\, ds \bigg|^p\bigg] \leq c_p |\mathbb{T}_n|^{p/2}(1 + E[\|{}_nY^r\|_\infty^p])$$

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

Proof. Let $E[\|{}_nY^r\|_\infty^p] < \infty$, as otherwise the claimed estimate is infinite. Clearly, a decomposition of the integral shows that

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

for all $j \in \{1, \ldots, k_n\}$. Hence, a first estimation gives

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

for $c_{p,1} := 2^{p-1}(T-r)^p\lambda_0^p(1 + c_{p,0})$ and a second yields that

$$E\bigg[\max_{j \in \{1,\ldots,k_n\}} \bigg| \int_{t_{j-1,n}}^{t_{j,n}} F(t_{j,n}, \underline{s}_n, {}_nY)\, ds \bigg|^p\bigg] \leq c_{p,2}|\mathbb{T}_n|^p(1 + E[\|{}_nY^r\|_\infty^p])$$

with $c_{p,2} := 2^{p-1}c_0^p(1 + c_{p,0})$. Thus, the constant $c_p := 2^{p-1}(c_{p,1} + (T-r)^{p/2}c_{p,2})$ satisfies the asserted estimate. □

The third estimate deals with Volterra integrals driven by ${}_nW$ and W, where $n \in \mathbb{N}$.

Proposition 3.5. *Let $F : [r,T]^2 \times C([0,T], \mathbb{R}^m) \to \mathbb{R}^{m \times d}$ be non-anticipative, product measurable and such that $F(\cdot, s, x)$ is absolutely continuous on $[s,T]$ for all $s \in [r,T]$ and each $x \in C([0,T], \mathbb{R}^m)$. Suppose that there are $\lambda_0, c_0 \geq 0$ such that*

$$\begin{aligned} |F(u,t,x) - F(u,s,x)| + |\partial_u F(u,t,x) - \partial_u F(u,s,x)| &\leq \lambda_0 d_\infty((t,x),(s,x)) \\ |F(t,s,x)| + |\partial_t F(t,s,x)| &\leq c_0(1 + \|x\|_\infty) \end{aligned}$$

for any $s,t,u \in [r,T)$ with $s < t < u$ and every $x \in C([0,T],\mathbb{R}^m)$. Moreover, let $({}_nY)_{n\in\mathbb{N}}$ be a sequence in $\mathscr{C}([0,T],\mathbb{R}^m)$ for which there are $p \geq 2$ and $c_{p,0} \geq 0$ such that

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

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

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

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

Proof. We suppose that $E[\|{}_nY^r\|_\infty^p]$ is finite and decompose the integral to get that

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

for each $j \in \{1,\dots,k_n\}$. Hence, we may apply Fubini's theorem for stochastic integrals from [17] to obtain that

$$\begin{aligned}\int_r^{t_{j-1,n}} F(t_{j,n},s_n,{}_nY) - F(t_{j,n},\underline{s}_n,{}_nY)\,dW_s &= \int_r^{t_{j-1,n}} F(s,s_n,{}_nY) - F(s,\underline{s}_n,{}_nY)\,dW_s\\ &+ \int_r^{t_{j,n}}\int_r^{t\wedge t_{j-1,n}} \partial_t F(t,s_n,{}_nY) - \partial_t F(t,\underline{s}_n,{}_nY)\,dW_s\,dt \quad \text{a.s.}\end{aligned}$$

for all $j \in \{1,\dots,k_n\}$. Regarding the first expression, we estimate that

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

for $c_{p,1} := 2^{p-1}w_p(T-r)^{p/2}\lambda_0^p(1+c_{p,0})$, where w_p is the constant satisfying (3.4). For the second expression we first calculate that

$$\begin{aligned}E\bigg[\max_{j\in\{1,\dots,k_n\}}\bigg|\int_r^{t_{j-1,n}}\int_r^t \partial_t F(t,s_n,{}_nY) - \partial_t F(t,\underline{s}_n,{}_nY)\,dW_s\,dt\bigg|^p\bigg]&\\ \leq c_{p,2}|\mathbb{T}_n|^{p/2}(1 + E[\|{}_nY^r\|_\infty^p])&\end{aligned}$$

with $c_{p,2} := 2^p(p+2)^{-1}w_p(T-r)^{3p/2}\lambda_0^p(1+c_{p,0})$. And secondly,

$$\begin{aligned}&E\bigg[\max_{j\in\{1,\dots,k_n\}}\bigg|\int_{t_{j-1,n}}^{t_{j,n}}\int_r^{t_{j-1,n}} \partial_t F(t,s_n,{}_nY) - \partial_t F(t,\underline{s}_n,{}_nY)\,dW_s\,dt\bigg|^p\bigg]\\ &\leq \sum_{j=1}^{k_n}(t_{j,n} - t_{j-1,n})^{p-1}\int_{t_{j-1,n}}^{t_{j,n}} E\bigg[\bigg|\int_r^{t_{j-1,n}} \partial_t F(t,s_n,{}_nY) - \partial_t F(t,\underline{s}_n,{}_nY)\,dW_s\bigg|^p\bigg]\,dt\\ &\leq c_{p,3}|\mathbb{T}_n|^{p-1}(1 + E[\|{}_nY^r\|_\infty^p]),\end{aligned}$$

where $c_{p,3} := 2^{p-1}w_p(T-r)^{p/2+1}\lambda_0^p(1+c_{p,0})$. Next, for the remaining term Fubini's theorem for stochastic integrals yields that

$$\int_{t_{j-1,n}}^{t_{j,n}} F(t_{j,n},\underline{s}_n,{}_nY)\,dW_s = \int_{t_{j-1,n}}^{t_{j,n}} F(s,\underline{s}_n,{}_nY)\,dW_s + \int_{t_{j-1,n}}^{t_{j,n}}\int_{t_{j-1,n}}^{t} \partial_t F(t,\underline{s}_n,{}_nY)\,dW_s\,dt \quad \text{a.s.} \tag{3.8}$$

for any $j \in \{1,\dots,k_n\}$. For the first term we have

$$\begin{aligned} E\bigg[\max_{j\in\{1,\dots,k_n\}}\bigg|\int_{t_{j-1,n}}^{t_{j,n}} F(s,\underline{s}_n,{}_nY)\,dW_s\bigg|^p\bigg] &\leq \sum_{j=1}^{k_n} E\bigg[\bigg|\int_{t_{j-1,n}}^{t_{j,n}} F(s,\underline{s}_n,{}_nY)\,dW_s\bigg|^p\bigg] \\ &\leq c_{p,4}|\mathbb{T}_n|^{p/2-1}(1+E[\|{}_nY^r\|_\infty^p]) \end{aligned}$$

with $c_{p,4} := 2^{p-1}w_p(T-r)c_0^p(1+c_{p,0})$. Finally, for the second stochastic integral in the decomposition (3.8) it holds that

$$\begin{aligned} &E\bigg[\max_{j\in\{1,\dots,k_n\}}\bigg|\int_{t_{j-1,n}}^{t_{j,n}}\int_{t_{j-1,n}}^{t} \partial_t F(t,\underline{s}_n,{}_nY)\,dW_s\,dt\bigg|^p\bigg] \\ &\quad\leq \sum_{j=1}^{k_n}(t_{j,n}-t_{j-1,n})^{p-1}\int_{t_{j-1,n}}^{t_{j,n}} E\bigg[\bigg|\int_{t_{j-1,n}}^{t} \partial_t F(t,\underline{s}_n,{}_nY)\,dW_s\bigg|^p\bigg]\,dt \\ &\quad\leq c_{p,5}|\mathbb{T}_n|^{3p/2-1}(1+E[\|{}_nY^r\|_\infty^p]) \end{aligned}$$

for $c_{p,5} := 2^p(p+2)^{-1}w_p(T-r)c_0^p(1+c_{p,0})$. Hence, the asserted estimate follows readily by setting $c_p := 5^{p-1}((T-r)(c_{p,1}+c_{p,2})+(T-r)^{p/2}c_{p,3}+c_{p,4}+(T-r)^pc_{p,5})$. □

4 Estimates and decompositions for the convergence result

4.1 Decomposition into remainder terms

We first give a moment estimate for solutions to (2.6) that does not depend on $n \in \mathbb{N}$.

Proposition 4.1. *Let* (C.5) *and* (C.6) *hold,* $h \in W_r^{1,2}([0,T],\mathbb{R}^d)$ *and* $\lambda \geq 0$ *be so that*

$$|\overline{B}(u,t,x)-\overline{B}(u,s,x)| + |\partial_u\overline{B}(u,t,x)-\partial_u\overline{B}(u,s,x)| \leq \lambda d_\infty((t,x),(s,x))$$

for any $s,t,u \in [r,T)$ *with* $s<t<u$ *and every* $x \in C([0,T],\mathbb{R}^m)$. *Then for each* $p \geq 2$ *there is* $c_p > 0$ *such that any* $n \in \mathbb{N}$ *and each solution* ${}_nY$ *to* (2.6) *satisfy*

$$E[\|{}_nY\|_\infty^p] + E[\|{}_nY^s - {}_nY^t\|_\infty^p]/|s-t|^{p/2} \leq c_p(1+E[\|{}_nY^r\|_\infty^p]) \tag{4.1}$$

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

Proof. We let $E[\|{}_nY^r\|_\infty^p] < \infty$ and may certainly assume in (C.6) that $\kappa > 0$. For given $l \in \mathbb{N}$ the stopping time $\tau_{l,n} := \inf\{t \in [0,T]\,|\,|{}_nY_t| \geq l\}\vee r$ satisfies $\|{}_nY^{\tau_{l,n}}\|_\infty \leq \|{}_nY^r\|_\infty \vee l$

and we readily estimate that

$$\begin{aligned}(E[\|{}_nY^{s\wedge\tau_{l,n}} - {}_nY^{t\wedge\tau_{l,n}}\|_\infty^p])^{1/p} &\leq \Big(\overline{c}_p(t-s)^{p/2-1}\int_s^t 1 + E[\|{}_nY^{u\wedge\tau_{l,n}}\|_\infty^{\kappa p}]\,du\Big)^{1/p}\\ &+ \bigg(E\bigg[\sup_{v\in[s,t]}\bigg|\int_s^{v\wedge\tau_{l,n}} \overline{B}(u,u,{}_nY)\,d_nW_u\bigg|^p\bigg]\bigg)^{1/p}\\ &+ \bigg(E\bigg[\bigg(\int_s^{t\wedge\tau_{l,n}}\bigg|\int_r^v \partial_v\overline{B}(v,u,{}_nY)\,d_nW_u\bigg|\,dv\bigg)^p\bigg]\bigg)^{1/p}\end{aligned}\tag{4.2}$$

for any fixed $s,t\in[r,T]$ with $s\leq t$ and $\overline{c}_p := 6^{p-1}(1+T-r)^p((T-r)^{p/2}+\|h\|_{1,2,r}^p+w_p)c^p$. We recall the constant $\hat{w}_{p/\kappa,1}$ such that (3.3) holds when p and q are replaced by p/κ and 1, respectively. Then

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

$$\bigg(E\bigg[\bigg(\int_{\underline{u}_n}^{u\wedge\tau_{l,n}}\int_r^v |\partial_v\overline{B}(v,u',{}_nY)_n\dot{W}_{u'}|\,du'\,dv\bigg)^{p/\kappa}\bigg]\bigg)^\kappa \leq (T-r)^p c_{p,1}(u-\underline{u}_n)^{p/2}$$

for any given $u\in[s,T]$ with the constant$c_{p,1} := 2^{p/2}\hat{w}_{p/\kappa,1}^\kappa c^p$. We let $\overline{c}_{p/\kappa}$ be defined just as $\overline{c}_p$ above with p replaced by p/κ to get that

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

for $c_{p,2} := 2^{p-1}(\overline{c}_{p/\kappa}^\kappa + (1+T-r)^p c_{p,1})$, due to the validity of (4.2). Hence, an application of Hölder's inequality yields that

$$\begin{aligned}&E\bigg[\bigg(\int_s^{t\wedge\tau_{l,n}} |(\overline{B}(u,u,{}_nY) - \overline{B}(\underline{u}_n,\underline{u}_n,{}_nY)\,{}_n\dot{W}_u|\,du\bigg)^p\bigg]\\ &\leq c_{p,3}(t-s)^{p/2-1}\int_s^t (1+E[\|{}_nY^{u\wedge\tau_{l,n}}\|_\infty^p])^\kappa\,du \quad \text{and}\\ &E\bigg[\bigg(\int_s^{t\wedge\tau_{l,n}}\int_r^v |(\partial_v\overline{B}(v,u,{}_nY) - \partial_v\overline{B}(v,\underline{u}_n,{}_nY))_n\dot{W}_u|\,du\,dv\bigg)^p\bigg]\\ &\leq (T-r)^p c_{p,3}(t-s)^{p/2-1}\int_s^t (1+E[\|{}_nY^{u\wedge\tau_{l,n}}\|_\infty^p])^\kappa\,du,\end{aligned}$$

where $c_{p,3} := 2^{p/2}3^p\hat{w}_{(p/2)/(1-\kappa)}^{1-\kappa}(\lambda^p(1+c_{p,2})+c^p(T-r)^{p/2})$. Moreover, the constant $\hat{w}_p$ appearing in (3.5) satisfies

$$E\bigg[\sup_{v\in[s,t]}\bigg|\int_s^{v\wedge\tau_{l,n}} \overline{B}(\underline{u}_n,\underline{u}_n,{}_nY)\,d_nW_u\bigg|^p\bigg] \leq \hat{w}_p c^p(t-s)^{p/2} \quad \text{and}$$

$$E\bigg[\bigg(\int_s^{t\wedge\tau_{l,n}}\bigg|\int_r^v \partial_v\overline{B}(v,\underline{u}_n,{}_nY)\,d_nW_u\bigg|\,dv\bigg)^p\bigg] \leq (T-r)^p\hat{w}_p c^p(t-s)^{p/2}.$$

Thus, with the constant $c_{p,4} := 3^{p-1}(2\overline{c}_p + (1+T-r)^p(c_{p,3}+\hat{w}_p c^p)$ we can now infer from (4.2) that

$$E[\|{}_nY^{s\wedge\tau_{l,n}} - {}_nY^{t\wedge\tau_{l,n}}\|_\infty^p] \leq c_{p,4}(t-s)^{p/2-1}\int_s^t 1 + E[\|{}_nY^{u\wedge\tau_{l,n}}\|_\infty^p]\,du. \tag{4.3}$$

Hence, Gronwall's inequality and Fatou's lemma imply that

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

where $c_{p,5} := 2^{p-1}\max\{1, T-r\}^{p/2}\max\{1, c_{p,4}\}e^{2^{p-1}(T-r)^{p/2}c_{p,4}}$. For this reason, we set $c_p := (1 + c_{p,4})(1 + c_{p,5})$ and apply Fatou's lemma to (4.3), which gives the result. □

Corollary 4.2. *Assume* (C.5), (C.6) *and* (C.8) *and let* $h \in W_r^{1,2}([0,T],\mathbb{R}^d)$. *Then for every* $p \geq 2$ *there is* $c_p > 0$ *such that each solution* Y *to* (2.7) *satisfies*

$$E[\|Y\|_\infty^p] + E[\|Y^s - Y^t\|_\infty^p]/|s-t|^{p/2} \leq c_p(1 + E[\|Y^r\|_\infty^p]) \tag{4.4}$$

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

Proof. As the map R given by (2.8) is bounded, the assertion is a direct consequence of Proposition 4.1 by replacing $\underline{B}$ by $\underline{B} + R$, $\overline{B}$ by 0 and Σ by $\overline{B} + \Sigma$. □

For $n \in \mathbb{N}$ let us recall the linear operator L_n and the function γ_n given at (2.3) and (3.7), respectively, and deduce the main decomposition to establish the limit (2.10).

Proposition 4.3. *Let* (C.5)-(C.8) *hold and* $h \in W_r^{1,2}([0,T],\mathbb{R}^d)$. *Then for each* $p \geq 2$ *there is* $c_p > 0$ *such that each* $n \in \mathbb{N}$ *and any two solutions* ${}_nY$ *and* Y *of* (2.6) *and* (2.7), *respectively, satisfy*

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

Proof. We suppose that $E[\|{}_nY^r\|_\infty^p]$ and $E[\|Y^r\|_\infty^p]$ are finite and aim to derive the estimate by applying Gronwall's inequality to the increasing function $\varphi_n : [r,T] \to \mathbb{R}_+$ given by

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

To this end, let us write the difference of ${}_nY$ and Y as follows:

$$\begin{aligned}
{}_nY_t - Y_t = {}_nY_r - Y_r &+ \int_r^t \underline{B}(s,s,{}_nY) - \underline{B}(s,s,Y)\,ds \\
&+ \int_r^t B_H(s,s,{}_nY) - B_H(s,s,Y)\,dh(s) \\
&+ {}_n\Delta_t + \int_r^t \int_r^v \partial_v\underline{B}(v,u,{}_nY) - \partial_v\underline{B}(v,u,Y)\,du\,dv
\end{aligned}$$

$$+\int_r^t\int_r^v \partial_v B_H(v,u,{}_nY) - \partial_v B_H(v,u,Y)\,dh(u)\,dv$$
$$+\int_r^t \Sigma(s,s,{}_nY) - \Sigma(s,s,Y)\,dW_s$$
$$+\int_r^t\int_r^v \partial_v\Sigma(v,u,{}_nY) - \partial_v\Sigma(v,u,Y)\,dW_u\,dv$$

for each $t\in[r,T]$ a.s. with a process ${}_n\Delta\in\mathscr{C}([0,T],\mathbb{R}^m)$ satisfying

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

for any $t\in[r,T]$ a.s. So, we let the terms ${}_nY_r - Y_r$ and ${}_n\Delta$ unchanged, then for the constant $c_{p,1} := 15^{p-1}(1+T-r)^p(T-r)^{p/2-1}((T-r)^{p/2}+\|h\|_{1,2,r}^p + w_p)\lambda^p$ we have

$$\varphi_n(t)^{1/p} \le \delta_{n,1}^{1/p} + \delta_n(t)^{1/p} + \left(c_{p,1}\int_r^{t_n} \delta_{n,1} + \delta_{n,2}(s) + \varepsilon_n(s) + \varphi_n(s)\,ds\right)^{1/p} \tag{4.5}$$

for all $t\in[r,T]$, where we have set $\delta_{n,1} := E[\|{}_nY^r - Y^r\|_\infty^p]$ and the measurable functions $\delta_n, \delta_{n,2}, \varepsilon_n : [r,T]\to\mathbb{R}_+$ are defined by

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

To obtain the estimate (4.5), we used the chain of inequalities: $E[\|L_n({}_nY)^{\underline{s}_n} - L_n(Y)^{\underline{s}_n}\|_\infty^p]$ $\le E[\|{}_nY^r - Y^r\|_\infty^p \vee \max_{j\in\{0,\dots,k_n\}:\, t_{j,n}\le s} |{}_nY_{t_{j,n}} - Y_{t_{j,n}}|^p] \le \delta_{n,1} + \varphi_n(s)$, valid for every $s\in[r,T]$.

For the estimation of δ_n let us define two processes ${}_{n,3}\Delta, {}_{n,5}\Delta \in \mathscr{C}([0,T],\mathbb{R}^m)$ by ${}_{n,3}\Delta_t := \int_r^t R(t,\underline{s}_n,{}_nY)(\gamma_n(s)-1)\,ds$ and

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

and choose ${}_{n,4}\Delta\in\mathscr{C}([0,T],\mathbb{R}^m)$ such that ${}_{n,4}\Delta_t = \int_r^t \overline{B}(t,\underline{s}_n,{}_nY)\,d({}_nW_s - W_s)$ for any $t\in[r,T]$ a.s. Then ${}_n\Delta$ admits the following representation:

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

for all $t\in[r,T]$ a.s. Due to the assumptions, we may assume without loss of generality that the Lipschitz constant λ is large enough such that

$$|R(u,t,x) - R(u,s,y)| \le \lambda d_\infty((t,x),(s,y))$$

for any $s,t,u \in [r,T)$ with $s < t < u$ and every $x,y \in C([0,T],\mathbb{R}^m)$. Thus, for the constant $c_{p,2} := 10^{p-1}(1+T-r)^p(T-r)^{p/2-1}((T-r)^{p/2}+w_p)\lambda^p$ we get that

$$\delta_n(t)^{1/p} \leq \delta_{n,3}(t)^{1/p} + \delta_{n,4}(t)^{1/p} + \delta_{n,5}(t)^{1/p} + \bigg(c_{p,2}\int_r^{t_n} \delta_{n,1} + (s-\underline{s}_n)^{p/2} + \delta_{n,2}(s) + \varepsilon_n(s) + \varphi_n(s)\,ds\bigg)^{1/p} \tag{4.6}$$

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

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

Thanks to Proposition 4.1 and Corollary 4.2, there are $\underline{c}_p, \overline{c}_p > 0$ such that (4.1) and (4.4) hold when c_p is replaced by $\underline{c}_p$ and $\overline{c}_p$, respectively. By combining (4.5) with (4.6), we see that

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

for fixed $t \in [r,T]$, where $c_{p,3} := 10^{p-1}(c_{p,1} + c_{p,2})$ and $c_{p,4} := 2^{p/2}(T-r)(1 + \underline{c}_p + \overline{c}_p)c_{p,3}$. For this reason, Gronwall's inequality gives

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

with $c_p := e^{c_{p,3}(T-r)}(5^{p-1} + c_{p,4})$, which implies the desired estimate. □

By the estimate (3.2), Lemma 3.4 and Proposition 3.5, to prove (2.10), only the last remainder in the estimation of Proposition 4.3 should be investigated in more detail. Thus, let $\Phi_{h,n} : [r,T] \times C([0,T],\mathbb{R}^m) \times C([0,T],\mathbb{R}^d) \to \mathbb{R}^m$ be defined via

$$\Phi_{h,n}(s,y,w) := B_H(\underline{s}_n,\underline{s}_n,y)(h(s) - h(\underline{s}_n)) + \overline{B}(\underline{s}_n,\underline{s}_n,y)(L_n(w)(s) - L_n(w)(\underline{s}_n)) + \Sigma(\underline{s}_n,\underline{s}_n,y)(w(s) - w(\underline{s}_n)) + \int_{\underline{s}_n}^{s}\int_r^{\underline{s}_n} \partial_v\overline{B}(v,u,y)\,dL_n(w)(u)\,dv$$

for each $h \in W_r^{1,2}([0,T],\mathbb{R}^d)$ and any $n \in \mathbb{N}$. Whenever ${}_nY$ is a solution to (2.6), then we will utilize the following decomposition to deal with the considered remainder:

$$\begin{aligned}
&(\overline{B}(t_{j,n},s,{}_nY) - \overline{B}(t_{j,n},\underline{s}_n,{}_nY))_n\dot{W}_s - R(t_{j,n},\underline{s}_n,{}_nY)\gamma_n(s) \\
&= (\overline{B}(t_{j,n},s,{}_nY) - \overline{B}(t_{j,n},\underline{s}_n,{}_nY) - \partial_x\overline{B}(t_{j,n},\underline{s}_n,{}_nY)({}_nY_s - {}_nY_{\underline{s}_n}))_n\dot{W}_s \\
&\quad + \partial_x\overline{B}(t_{j,n},\underline{s}_n,{}_nY)({}_nY_s - {}_nY_{\underline{s}_n} - \Phi_{h,n}(s,{}_nY,W))_n\dot{W}_s \\
&\quad + \partial_x\overline{B}(t_{j,n},\underline{s}_n,{}_nY)\Phi_{h,n}(s,{}_nY,W)_n\dot{W}_s - R(t_{j,n},\underline{s}_n,{}_nY)\gamma_n(s)
\end{aligned} \tag{4.7}$$

for all $j \in \{1,\dots,k_n\}$ and each $s \in [r,t_{j,n})$.

4.2 Moment estimates for the first two remainders

The first result in this section together with Lemma 3.2 provide an estimate of the first remainder appearing in (4.7).

Proposition 4.4. *Let* (C.4)-(C.6) *be valid,* $h \in W_r^{1,2}([0,T],\mathbb{R}^d)$ *and* F *be a product measurable functional on* $[r,T]\times[r,T)\times C([0,T],\mathbb{R}^m)$ *so that the following two conditions hold:*

(i) There exists $\lambda \geq 0$ *such that* $|\overline{B}(u,t,x) - \overline{B}(u,s,x)| + |\partial_u\overline{B}(u,t,x) - \partial_u\overline{B}(u,s,x)| \leq \lambda d_\infty((t,x),(s,x))$ *for any* $s,t,u \in [r,T)$ *with* $s < t < u$ *and all* $x \in C([0,T],\mathbb{R}^m)$.

(ii) The functional $[r,t)\times C([0,T],\mathbb{R}^m) \to \mathbb{R}$, $(s,x) \mapsto F(t,s,x)$ *is of class* $\mathbb{C}^{1,2}$ *for any* $t \in (r,T]$ *and there are* $c_0,\eta,\lambda_0 \geq 0$ *such that*

$$|\partial_s F(t,s,x)| + |\partial_{xx}F(t,s,x)| \leq c_0(1 + \|x\|_\infty^\eta),$$
$$|\partial_x F(u,t,x) - \partial_x F(u,s,x)| \leq \lambda_0 d_\infty((t,x),(s,x))$$

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

Then for any $p \geq 2$ *there is* $c_p > 0$ *such that for all* $n \in \mathbb{N}$ *and each solution* ${}_nY$ *to* (2.6),

$$E\Big[\max_{j\in\{1,\dots,k_n\}}\int_r^{t_{j,n}} |F(t_{j,n},s,{}_nY) - F(t_{j,n},\underline{s}_n,{}_nY) - \partial_x F(t_{j,n},\underline{s}_n,{}_nY)({}_nY_s - {}_nY_{\underline{s}_n})|^p\,ds\Big]$$
$$\leq c_p|\mathbb{T}_n|^{p-1}(1 + E[\|{}_nY^r\|_\infty^{(\eta\vee 2)p}]).$$

Proof. For any $j \in \{1,\dots,k_n\}$ let the product measurable map ${}_{n,j}\Delta : [r,t_{j,n})^2\times\Omega \to \mathbb{R}^{1\times m}$ be given by ${}_{n,j}\Delta_{s,u} := \partial_x F(t_{j,n},u,{}_nY) - \partial_x F(t_{j,n},\underline{s}_n,{}_nY)$, if $u \in [\underline{s}_n,s]$, and ${}_{n,j}\Delta_{s,u} := 0$, otherwise. Then from the functional Itô formula in [6] we infer that

$$\begin{aligned}
&F(t_{j,n},s,{}_nY) - F(t_{j,n},\underline{s}_n,{}_nY) - \partial_x F(t_{j,n},\underline{s}_n,{}_nY)({}_nY_s - {}_nY_{\underline{s}_n})\\
&= \int_{\underline{s}_n}^s \partial_u F(t_{j,n},u,{}_nY) + \frac{1}{2}\mathrm{tr}(\partial_{xx}F(t_{j,n},u,{}_nY)(\Sigma\Sigma')(u,u,{}_nY))\,du\\
&\quad + \int_{\underline{s}_n}^s {}_{n,j}\Delta_{s,u}(\underline{B}(u,u,{}_nY) + B_H(u,u,{}_nY)\dot{h}(u) + \overline{B}(u,u,{}_nY){}_n\dot{W}_u)\,du\\
&+ \int_{\underline{s}_n}^s {}_{n,j}\Delta_{s,v}\int_r^v \partial_v\underline{B}(v,u,{}_nY) + \partial_v B_H(v,u,{}_nY)\dot{h}(u) + \partial_v\overline{B}(v,u,{}_nY){}_n\dot{W}_u\,du\,dv\\
&+ \int_{\underline{s}_n}^s {}_{n,j}\Delta_{s,u}\Sigma(u,u,{}_nY)\,dW_u + \int_{\underline{s}_n}^s {}_{n,j}\Delta_{s,v}\int_r^v \partial_v\Sigma(v,u,{}_nY)\,dW_u\,dv
\end{aligned} \tag{4.8}$$

for each $s \in [r,t_{j,n})$ a.s. Now, for $\overline{\eta} := \eta\vee 2$ Proposition 4.1 gives a constant $\underline{c}_{\overline{\eta}p} > 0$ such that (4.1) holds when p and c_p are replaced by $\overline{\eta}p$ and $\underline{c}_{\overline{\eta}p}$, respectively. Then for the first two terms on the right-hand side in (4.8) we have

$$E\Big[\max_{j\in\{1,\dots,k_n\}}\sup_{s\in[r,t_{j,n})}\Big|\int_{\underline{s}_n}^s \partial_u F(t_{j,n},u,{}_nY) + \frac{1}{2}\mathrm{tr}(\partial_{xx}F(t_{j,n},u,{}_nY)(\Sigma\Sigma')(u,u,{}_nY))\,du\Big|^p\Big]$$
$$\leq 2^{p-1}c_0^p(s - \underline{s}_n)^p E[(1 + \|{}_nY\|_\infty^\eta)^p]$$

$$
\begin{aligned}
&+ 2^{-1}c_0^p(s - \underline{s}_n)^{p-1}\int_{\underline{s}_n}^s E\big[(1 + \|{}_nY^u\|_\infty^\eta)^p|(\Sigma\Sigma')(u,u,{}_nY)|^p\big]\,du\\
&\leq c_{p,1}|\mathbb{T}_n|^p\big(1 + E\big[\|{}_nY^r\|_\infty^{\overline{\eta}p}\big]\big)^{\eta/\overline{\eta}}
\end{aligned}
$$

with $c_{p,1} := 2^{2p-1}c_0^p(2^p + c^{2p})(1 + \underline{c}_{\overline{\eta}p})^{\eta/\overline{\eta}}$. We note that $|{}_{n,j}\Delta_{s,u}| \leq \lambda_0 d_\infty((s,{}_nY),(\underline{s}_n,{}_nY))$ for each $j \in \{1,\dots,k_n\}$ and all $s,u \in [r,t_{j,n})$ and by setting $\overline{c}_p := 2^{3p/2}\lambda_0^p(1 + \underline{c}_{\overline{\eta}p})^{1/\overline{\eta}}$, we obtain that

$$
\lambda_0^p\big(E\big[d_\infty((s,{}_nY),(\underline{s}_n,{}_nY))^{2p}\big]\big)^{1/2} \leq \overline{c}_p|\mathbb{T}_n|^{p/2}\big(1 + E\big[\|{}_nY^r\|_\infty^{\overline{\eta}p}\big]\big)^{1/\overline{\eta}}
$$

for each $s \in [r,T]$. Consequently, the Cauchy-Schwarz inequality gives us the following bound for the third and sixth expression in the decomposition (4.8):

$$
\begin{aligned}
&E\Big[\max_{j\in\{1,\dots,k_n\}}\int_r^{t_{j,n}}\Big|\int_{\underline{s}_n}^s {}_{n,j}\Delta_{s,v}\Big(\underline{B}(v,v,{}_nY) + \int_r^v \partial_v\underline{B}(v,u,{}_nY)\,du\Big)\,dv\Big|^p\,ds\Big]\\
&\leq 2^{p-1}c^p\int_r^T (s-\underline{s}_n)^p\lambda_0^pE\big[d_\infty((s,{}_nY),(\underline{s}_n,{}_nY))^p(1 + \|{}_nY^s\|_\infty^\kappa)^p\big]\,ds\\
&\quad + 2^{p-1}c^p\int_r^T \lambda_0^pE\Big[d_\infty((s,{}_nY),(\underline{s}_n,{}_nY))^p(s-\underline{s}_n)^{p-1}\int_{\underline{s}_n}^s\Big(\int_r^v 1 + \|{}_nY^u\|_\infty^\kappa\,du\Big)^p\,dv\Big]\,ds\\
&\leq c_{p,2}|\mathbb{T}_n|^p\big(1 + E[\|{}_nY^r\|_\infty^{\overline{\eta}p}\big)^{2/\overline{\eta}}\int_r^T (s-\underline{s}_n)^{p/2}\,ds
\end{aligned}
$$

for $c_{p,2} := 2^{5p/2-1}(1 + (T-r)^p)c^p(1 + \underline{c}_{\overline{\eta}p})^{1/\overline{\eta}}\overline{c}_p$. For the fourth expression we apply the Cauchy-Schwarz inequality twice, which entails that

$$
\begin{aligned}
&E\Big[\max_{j\in\{1,\dots,k_n\}}\int_r^{t_{j,n}}\Big|\int_{\underline{s}_n}^s {}_{n,j}\Delta_{s,u}B_H(u,u,{}_nY)\,dh(u)\Big|^p\,ds\Big]\\
&\leq \|h\|_{1,2,r}^p c^p\int_r^T (s-\underline{s}_n)^{p/2}\lambda_0^pE\big[d_\infty((s,{}_nY),(\underline{s}_n,{}_nY))^p(1 + \|{}_nY\|_\infty^\kappa)^p\big]\,ds\\
&\leq c_{p,3}|\mathbb{T}_n|^p\big(1 + E\big[\|{}_nY^r\|_\infty^{\overline{\eta}p}\big]\big)^{2/\overline{\eta}},
\end{aligned}
$$

where $c_{p,3} := 2^{3p/2}\|h\|_{1,2,r}^p c^p(1 + \underline{c}_{\overline{\eta}p})^{1/\overline{\eta}}\overline{c}_p$. Proceeding similarly, it follows for the seventh expression that

$$
\begin{aligned}
&E\Big[\max_{j\in\{1,\dots,k_n\}}\int_r^{t_{j,n}}\Big|\int_{\underline{s}_n}^s {}_{n,j}\Delta_{s,v}\int_r^v \partial_vB_H(v,u,{}_nY)\,dh(u)\,dv\Big|^p\,ds\Big]\\
&\leq \int_r^T (s-\underline{s}_n)^{p-1}\int_{\underline{s}_n}^s \lambda_0^pE\Big[d_\infty((s,{}_nY),(\underline{s}_n,{}_nY))^p\Big|\int_r^v \partial_vB_H(v,u,{}_nY)\,dh(u)\Big|^p\Big]\,dv\,ds\\
&\leq c_{p,4}|\mathbb{T}_n|^p\big(1 + E\big[\|{}_nY^r\|_\infty^{\overline{\eta}p}\big]\big)^{2/\overline{\eta}}\int_r^T (s-\underline{s}_n)^{p/2}\,ds
\end{aligned}
$$

with $c_{p,4} := (T-r)^{p/2}c_{p,3}$. We turn to the fifth and eight term in (4.8) and once again apply the Cauchy-Schwarz inequality, which leads us to

$$
E\Big[\max_{j\in\{1,\dots,k_n\}}\int_r^{t_{j,n}}\Big|\int_{\underline{s}_n}^s {}_{n,j}\Delta_{s,v}\Big(\overline{B}(v,v,{}_nY)_n\dot{W}_v + \int_r^v \partial_v\overline{B}(v,u,{}_nY)\,d_nW_u\Big)\,dv\Big|^p\,ds\Big]
$$

$$
\begin{aligned}
&\le 2^{p-1}c^p \int_r^T (s-\underline{s}_n)^{p/2}\lambda_0^p E\Big[d_\infty((s,{}_nY),(\underline{s}_n,{}_nY))^p\Big(\int_{\underline{s}_n}^s |{}_n\dot{W}_v|^2\,dv\Big)^{p/2}\Big]\,ds\\
&\quad + 2^{p-1}c^p\int_r^T (s-\underline{s}_n)^{p-1}\int_{\underline{s}_n}^s \lambda_0^p E\Big[d_\infty((s,{}_nY),(\underline{s}_n,{}_nY))^p\Big(\int_r^v |{}_n\dot{W}_u|\,du\Big)^p\Big]\,dv\,ds\\
&\le c_{p,5}|\mathbb{T}_n|^p\big(1+E\big[\|{}_nY^r\|_\infty^{\overline{\eta}p}\big]\big)^{1/\overline{\eta}}
\end{aligned}
$$

for $c_{p,5} := 2^{2p-1}\hat{w}_{p,2}^{1/2}(1+(T-r)^{p+1}/(p+1))c^p\overline{c}_p$. By using the constant $\overline{c}_{\mathbb{T}}$ appearing in condition (C.4), we derive the following estimate for the ninth term:

$$
\begin{aligned}
&E\Big[\max_{j\in\{1,\dots,k_n\}}\int_r^{t_{j,n}}\Big|\int_{\underline{s}_n}^s {}_{n,j}\Delta_{s,u}\Sigma(u,u,{}_nY)\,dW_u\Big|^p\,ds\Big]\\
&\le w_p c^p\sum_{j=1}^{k_n}\int_r^{t_{j,n}}(s-\underline{s}_n)^{p/2-1}\int_{\underline{s}_n}^s \lambda_0^p E\big[d_\infty((s,{}_nY),(\underline{s}_n,{}_nY))^p\big]\,du\,ds\\
&\le c_{p,6}|\mathbb{T}_n|^{p-1}\big(1+E\big[\|{}_nY^r\|_\infty^{\overline{\eta}p}\big]\big)^{1/\overline{\eta}},
\end{aligned}
$$

where $c_{p,6} := 2^{p/2}w_p c^p\underline{c}_p(T-r)\overline{c}_{\mathbb{T}}$. Finally, for the last expression we now readily estimate that

$$
\begin{aligned}
&E\Big[\max_{j\in\{1,\dots,k_n\}}\int_r^{t_{j,n}}\Big|\int_{\underline{s}_n}^s {}_{n,j}\Delta_{s,v}\int_r^v \partial_v\Sigma(v,u,{}_nY)\,dW_u\,dv\Big|^p\,ds\Big]\\
&\le \int_r^T (s-\underline{s}_n)^{p-1}\int_{\underline{s}_n}^s E\Big[\lambda_0^p d_\infty((s,{}_nY),(\underline{s}_n,{}_nY))^p\Big|\int_r^v \partial_v\Sigma(v,u,{}_nY)\,dW_u\Big|^p\Big]\,dv\,ds\\
&\le c_{p,7}|\mathbb{T}_n|^p\big(1+E\big[\|{}_nY^r\|_\infty^{\overline{\eta}p}\big]\big)^{1/\overline{\eta}}\int_r^T (s-\underline{s}_n)^{p/2}\,ds
\end{aligned}
$$

for $c_{p,7} := 2^{p/2}c^p\overline{c}_p w_{2p}^{1/2}(T-r)^{p/2}$. So, we let $c_{p,8} := (T-r)((T-r)c_{p,1}+c_{p,3}+c_{p,5})$ and $c_{p,9} := (T-r)^{p/2+2}(c_{p,2}+c_{p,4}+c_{p,7})$ and conclude by setting $c_p := 7^{p-1}(c_{p,6}+c_{p,8}+c_{p,9})$. □

Next, we give a bound for the second remainder in (4.7), which allows for another application of Lemma 3.2, according to Remark 3.3.

Lemma 4.5. *Let* (C.5)-(C.7) *be valid and* $h \in W_r^{1,2}([0,T],\mathbb{R}^d)$. *Then for each* $p \ge 2$ *there is* $c_p > 0$ *such that each* $n \in \mathbb{N}$ *and any solution* ${}_nY$ *to* (2.6) *satisfy*

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

for every $s \in [r,T)$.

Proof. From Fubini's theorem for deterministic and stochastic integrals and the definition of $\Phi_{h,n}$ we get that

$$
\begin{aligned}
&{}_nY_s - {}_nY_{\underline{s}_n} - \Phi_{h,n}(s, {}_nY, W) = \int_{\underline{s}_n}^s \underline{B}(u,u,{}_nY)\,du \\
&+ \int_{\underline{s}_n}^s B_H(u,u,{}_nY) - B_H(\underline{s}_n,\underline{s}_n,{}_nY)\,dh(u) \\
&+ \int_{\underline{s}_n}^s \overline{B}(u,u,{}_nY) - \overline{B}(\underline{s}_n,\underline{s}_n,{}_nY)\,d_nW_u \\
&+ \int_{\underline{s}_n}^s \int_r^v \partial_v \underline{B}(v,u,{}_nY) + \partial_v B_H(v,u,{}_nY)\,dh(u)\,dv \\
&+ \int_{\underline{s}_n}^s \int_{\underline{s}_n}^v \partial_v \overline{B}(v,u,{}_nY)\,d_nW_u\,dv \\
&+ \int_{\underline{s}_n}^s \Sigma(u,u,{}_nY) - \Sigma(\underline{s}_n,\underline{s}_n,{}_nY)\,dW_u + \int_{\underline{s}_n}^s \int_r^v \partial_v\Sigma(v,u,{}_nY)\,dW_u\,dv \quad \text{a.s.}
\end{aligned}
\tag{4.9}
$$

Proposition 4.1 provides a constant $\underline{c}_{2p} > 0$ such that (4.1) holds when p and c_p are replaced by $2p$ and $\underline{c}_{2p}$, respectively. We set $\overline{c}_{p,2} := \lambda^p + (T-r)^{p/2}c^p$ and $\overline{c}_{p,1} := (1+\underline{c}_{2p})^{1/2}$ and define eight constants as follows:

$$
\begin{aligned}
&c_{p,1} := 2^{2p}c^p\overline{c}_{p,1},\ c_{p,2} := 2^{3p}\|h\|_{1,2,r}^p\overline{c}_{p,1}\overline{c}_{p,2},\ c_{p,3} := 2^{3p/2}3^p\hat{w}_{p,2}^{1/2}\overline{c}_{p,1}\overline{c}_{p,2}, \\
&c_{p,4} := (T-r)^p c_{p,1},\ c_{p,5} := 2^{2p}(T-r)^{p/2}\|h\|_{1,2,r}^p c^p\overline{c}_{p,1},\ c_{p,6} := 2^{3p/2}\hat{w}_{p,1}(T-r)^{p/2}c^p, \\
&c_{p,7} := 2^p3^p w_p\overline{c}_{p,1}\overline{c}_{p,2} \text{ and } c_{p,8} := 2^p w_p(T-r)^{p/2}c^p.
\end{aligned}
$$

By using the inequalities of Jensen and Cauchy-Schwarz and (3.4), it follows readily that the p-th moment of the i-th expression in the decomposition (4.9) is bounded by $c_{p,i}|\mathbb{T}_n|^p(1+E[\|{}_nY^r\|_\infty^{2p}])^{1/2}$ for all $i \in \{1,\dots,8\}$. We set $c_p := 8^{p-1}(c_{p,1}+\cdots+c_{p,8})$ and the asserted estimate follows. □

4.3 A second moment estimate for the third remainder

We directly bound the third remainder in (4.7) by repeatedly using an estimate that follows for any $n \in \mathbb{N}$ with $k_n \geq 2$ from Doob's L^2-maximal inequality; see [7][Lemma 33] for details.

(v) For every $l \in \{1,\dots,d\}$ assume that $({}_lU_i)_{i\in\{1,\dots,k_n-1\}}$ and $({}_lV_i)_{i\in\{1,\dots,k_n-1\}}$ are two sequences of $\mathbb{R}^{1\times m}$-valued and $\mathbb{R}^m$-valued random vectors, respectively, such that ${}_lU_i$ is $\mathscr{F}_{t_{i-1,n}}$-measurable, ${}_lV_i$ is $\mathscr{F}_{t_{i,n}}$-measurable,

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

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

$$E\Big[\max_{j\in\{1,\dots,k_n\}}\Big|\sum_{i=1}^{j-1}\sum_{l=1}^d {}_lU_i\,{}_lV_i\Big|^2\Big] \leq 4\sum_{i=1}^{k_n-1}\sum_{l_1,l_2=1}^d E\big[{}_{l_1}U_i\,{}_{l_1}V_i\,{}_{l_2}V_i'\,{}_{l_2}U_i'\big]. \tag{4.10}$$

Proposition 4.6. *Let* (C.5)-(C.8) *be satisfied and* $h \in W_r^{1,2}([0,T],\mathbb{R}^d)$. *Then there is* $c_2 > 0$ *such that for each* $n \in \mathbb{N}$ *and any solution* ${}_nY$ *to* (2.6) *it holds that*

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

Proof. By the definition (2.8) of the mapping R, we can write the k-th coordinate of $\partial_x\overline{B}(t_{j,n},\underline{s}_n,{}_nY)\Phi_{h,n}(s,{}_nY,W)_n\dot{W}_s - R(t_{j,n},\underline{s}_n,{}_nY)\gamma_n(s)$ in the form

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

for each $j \in \{1,\ldots,k_n\}$, any $k \in \{1,\ldots,m\}$ and all $s \in [r,t_{j,n})$, where we write $X^{(l)}$ for the l-th coordinate of any $\mathbb{R}^d$-valued process X for each $l \in \{1,\ldots,d\}$. Based on this identity, we use the following decomposition:

$$\begin{aligned}
&\Phi_{h,n}(s,{}_nY,W)_n\dot{W}_s^{(l)} - ((1/2)\overline{B}+\Sigma)(\underline{s}_n,\underline{s}_n,{}_nY)\gamma_n(s)e_l\\
&= B_H(\underline{s}_n,\underline{s}_n,{}_nY)(h(s_n)-h(\underline{s}_n))_n\dot{W}_s^{(l)} + \overline{B}(\underline{s}_n,\underline{s}_n,{}_nY)({}_nW_{s_n} - {}_nW_{\underline{s}_n})_n\dot{W}_s^{(l)}\\
&\quad + \Sigma(\underline{s}_n,\underline{s}_n,{}_nY)(\Delta W_{s_n}{}_n\dot{W}_s^{(l)} - \gamma_n(s)e_l) + B_H(\underline{s}_n,\underline{s}_n,{}_nY)(h(s)-h(s_n))_n\dot{W}_s^{(l)}\\
&\quad + \overline{B}(\underline{s}_n,\underline{s}_n,{}_nY)\big(({}_nW_s - {}_nW_{s_n})_n\dot{W}_s^{(l)} - (1/2)\gamma_n(s)e_l\big)\\
&\quad + \Sigma(\underline{s}_n,\underline{s}_n,{}_nY)(W_s - W_{s_n})_n\dot{W}_s^{(l)} + \bigg(\int_{\underline{s}_n}^s\int_r^{\underline{s}_n}\partial_v\overline{B}(v,u,{}_nY)\,d_nW_u\,dv\bigg)_n\dot{W}_s^{(l)}
\end{aligned} \tag{4.11}$$

with $l \in \{1,\ldots,d\}$. To handle the first appearing term, we decompose the integral and apply Fubini's theorem for stochastic integrals to rewrite that

$$\begin{aligned}
&\int_r^{t_{j,n}} \partial_x\overline{B}_{k,l}(t_{j,n},\underline{s}_n,{}_nY)B_H(\underline{s}_n,\underline{s}_n,{}_nY)(h(s_n)-h(\underline{s}_n))\,d_nW_s^{(l)}\\
&= \int_r^{t_{j-1,n}} \partial_x\overline{B}_{k,l}(s,s_n,{}_nY)B_H(s_n,s_n,{}_nY)(h(\overline{s}_n)-h(s_n))\,dW_s^{(l)}\\
&\quad + \int_r^{t_{j,n}}\int_r^{t\wedge t_{j-1,n}} \partial_t\partial_x\overline{B}_{k,l}(t,s_n,{}_nY)B_H(s_n,s_n,{}_nY)(h(\overline{s}_n)-h(s_n))\,dW_s^{(l)}\,dt \quad \text{a.s.}
\end{aligned}$$

for any $j \in \{1,\ldots,k_n\}$, every $k \in \{1,\ldots,m\}$ and each $l \in \{1,\ldots,d\}$. By Proposition 4.1, there is $\underline{c}_2 > 0$ such that (4.1) holds for $p = 2$ with $\underline{c}_2$ instead of c_p. Therefore,

$$\begin{aligned}
&E\bigg[\max_{j\in\{0,\ldots,k_n\}}\sum_{k=1}^m\bigg|\sum_{l=1}^d\int_r^{t_{j,n}} \partial_x\overline{B}_{k,l}(t_{j,n},\underline{s}_n,{}_nY)B_H(\underline{s}_n,\underline{s}_n,{}_nY)(h(s_n)-h(\underline{s}_n))\,d_nW_s^{(l)}\bigg|^2\bigg]\\
&\leq 2w_2c^2\overline{c}^2\int_r^T E\big[(1+\|{}_nY^{s_n}\|_\infty^\kappa)^2\big]|h(\overline{s}_n)-h(s_n)|^2\,ds\\
&\quad + 2w_2(T-r)c^2\overline{c}^2\int_r^T\int_r^t E\big[(1+\|{}_nY^{s_n}\|_\infty^\kappa)^2\big]|h(\overline{s}_n)-h(s_n)|^2\,ds\,dt\\
&\leq c_{2,1}|\mathbb{T}_n|\big(1+E\big[\|{}_nY^r\|_\infty^2\big]\big)^\kappa
\end{aligned}$$

with $c_{2,1} := 2^3 w_2(1+(T-r)^2/2)(T-r)\|h\|_{1,2,r}^2 c^2\overline{c}^2(1+\underline{c}_2)^\kappa$. Proceeding similarly, we obtain for the second term in the decomposition (4.11) that

$$\begin{aligned}
&\int_r^{t_{j,n}} \partial_x \overline{B}_{k,l}(t_{j,n},\underline{s}_n,{}_nY)\overline{B}(\underline{s}_n,\underline{s}_n,{}_nY)({}_nW_{s_n}-{}_nW_{\underline{s}_n})\,d_nW_s^{(l)}\\
&= \int_r^{t_{j-1,n}} \partial_x \overline{B}_{k,l}(s,s_n,{}_nY)\overline{B}(s_n,s_n,{}_nY)\Delta W_{s_n}\,dW_s^{(l)}\\
&\quad + \int_r^{t_{j,n}}\int_r^{t\wedge t_{j-1,n}} \partial_t\partial_x \overline{B}_{k,l}(t,s_n,{}_nY)\overline{B}(s_n,s_n,{}_nY)\Delta W_{s_n}\,dW_s^{(l)}\,dt \quad \text{a.s.}
\end{aligned}$$

for every $j \in \{1,\dots,k_n\}$, each $k\in\{1,\dots,m\}$ and any $l\in\{1,\dots,d\}$. Hence, by setting $c_{2,2} := 2w_2(1+(T-r)^2/2)(T-r)dc^2\overline{c}^2$, it follows readily that

$$\begin{aligned}
&E\Big[\max_{j\in\{0,\dots,k_n\}}\sum_{k=1}^m\Big|\sum_{l=1}^d\int_r^{t_{j,n}} \partial_x \overline{B}_{k,l}(t_{j,n},\underline{s}_n,{}_nY)\overline{B}(\underline{s}_n,\underline{s}_n,{}_nY)({}_nW_{s_n}-{}_nW_{\underline{s}_n})\,d_nW_s^{(l)}\Big|^2\Big]\\
&\leq 2w_2c^2\overline{c}^2\int_r^T\Big(E[|\Delta W_{t_n}|^2]+(T-r)\int_r^t E[|\Delta W_{s_n}|^2]\,ds\Big)\,dt \leq c_{2,2}|\mathbb{T}_n|.
\end{aligned}$$

To deal with the third term in (4.11), we utilize the $\mathbb{R}^d$-valued $\mathscr{F}_{t_{i,n}}$-measurable random vector

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

which is independent of $\mathscr{F}_{t_{i-1,n}}$ and satisfies $E[{}_{l,n}V_i]=0$ for any $i\in\{1,\dots,k_n\}$ and each $l\in\{1,\dots,d\}$. We note that if $\mathbb{I}_{l_2,l_1}\in\mathbb{R}^{d\times d}$ denotes the matrix whose (l_2,l_1)-entry is 1 and whose all other entries are zero, then

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

whenever $i\in\{1,\dots,k_n\}$ and $l_1,l_2\in\{1,\dots,d\}$. Now, by decomposing the integral once again, we obtain that

$$\begin{aligned}
&\int_r^{t_{j,n}} \partial_x \overline{B}_{k,l}(t_{j,n},\underline{s}_n,{}_nY)\Sigma(\underline{s}_n,\underline{s}_n,{}_nY)(\Delta W_{s_n}{}_n\dot{W}_s^{(l)}-\gamma_n(s)e_l)\,ds\\
&= \sum_{i=1}^{j-1}\partial_x \overline{B}_{k,l}(t_{i,n},t_{i-1,n},{}_nY)\Sigma(t_{i-1,n},t_{i-1,n},{}_nY)_{l,n}V_i\\
&\quad + \sum_{i_2=1}^{j-1}\int_{t_{i_2,n}}^{t_{i_2+1,n}}\sum_{i_1=1}^{i_2}\partial_t\partial_x \overline{B}_{k,l}(t,t_{i_1-1,n},{}_nY)\Sigma(t_{i_1-1,n},t_{i_1-1,n},{}_nY)\,{}_{l,n}V_{i_1}\,dt
\end{aligned}$$

for all $j\in\{1,\dots,k_n\}$, each $k\in\{1,\dots,m\}$ and every $l\in\{1,\dots,d\}$. Consequently, the estimate (4.10) and Young's inequality give us that

$$\begin{aligned}
&E\Big[\max_{j\in\{0,\dots,k_n\}}\sum_{k=1}^m\Big|\int_r^{t_{j,n}}\sum_{l=1}^d \partial_x \overline{B}_{k,l}(t_{j,n},\underline{s}_n,{}_nY)\Sigma(\underline{s}_n,\underline{s}_n,{}_nY)(\Delta W_{s_n}{}_n\dot{W}_s^{(l)}-\gamma_n(s)e_l)\,ds\Big|^2\Big]\\
&\leq 2^4|\mathbb{T}_n|\sum_{i=1}^{k_n-1}\Delta t_{i,n}\sum_{k=1}^m\sum_{l=1}^d E\big[|\partial_x \overline{B}_{k,l}(t_{i,n},t_{i-1,n},{}_nY)\Sigma(t_{i-1,n},t_{i-1,n},{}_nY)|^2\big]
\end{aligned}$$

$$+\,2^4(T-r)\int_r^T \sum_{i=1}^{k_n-1}(\Delta t_{i,n})^2 \sum_{k=1}^m\sum_{l=1}^d E\big[\big|\partial_t\partial_x\overline{B}_{k,l}(t,t_{i-1,n},{}_nY)\Sigma(t_{i-1,n},t_{i-1,n},{}_nY)\big|^2\big]\,dt$$
$$\leq c_{2,3}|\mathbb{T}_n|,$$

where $c_{2,3} := 2^4(1+(T-r)^2)(T-r)c^2\overline{c}^2$. For the fourth expression in (4.11) we integrate by parts, after another decomposition of the integral, which yields that

$$\begin{aligned}&\int_r^{t_{j,n}} \partial_x\overline{B}(t_{j,n},\underline{s}_n,{}_nY)B_H(\underline{s}_n,\underline{s}_n,{}_nY)(h(s)-h(s_n))\,d_nW_s^{(l)}\\&=\int_r^{t_{j,n}} \partial_x\overline{B}_{k,l}(s,\underline{s}_n,{}_nY)B_H(\underline{s}_n,\underline{s}_n,{}_nY)\Delta W_{s_n}^{(l)}\frac{(\overline{s}_n-s)}{\Delta\overline{s}_n}\,dh(s)\\&\quad+\int_r^{t_{j,n}}\int_r^t \partial_t\partial_x\overline{B}_{k,l}(t,\underline{s}_n,{}_nY)B_H(\underline{s}_n,\underline{s}_n,{}_nY)\Delta W_{s_n}^{(l)}\frac{(\overline{s}_n-s)}{\Delta\overline{s}_n}\,dh(s)\,dt\end{aligned}$$

for each $j\in\{1,\dots,k_n\}$, any $k\in\{1,\dots,m\}$ and every $l\in\{1,\dots,d\}$. Hence, from the Cauchy-Schwarz inequality we get that

$$\begin{aligned}&E\Big[\max_{j\in\{0,\dots,k_n\}}\sum_{k=1}^m\Big|\sum_{l=1}^d\int_r^{t_{j,n}} \partial_x\overline{B}(t_{j,n},\underline{s}_n,{}_nY)B_H(\underline{s}_n,\underline{s}_n,{}_nY)(h(s)-h(s_n))\,d_nW_s^{(l)}\Big|^2\Big]\\&\leq 2\|h\|_{1,2,r}^2\int_r^T\sum_{k=1}^m\sum_{l=1}^d E\big[|\partial_x\overline{B}_{k,l}(s,\underline{s}_n,{}_nY)B_H(\underline{s}_n,\underline{s}_n,{}_nY)|^2\big]E\big[|\Delta W_{s_n}^{(l)}|^2\big]\,ds\\&\quad+2\|h\|_{1,2,r}^2(T-r)\int_r^T\int_r^t\sum_{k=1}^m E\Big[\Big|\sum_{l=1}^d \partial_x\overline{B}_{k,l}(t,\underline{s}_n,{}_nY)B_H(\underline{s}_n,\underline{s}_n,{}_nY)\Delta W_{s_n}^{(l)}\Big|^2\Big]\,ds\,dt\\&\leq c_{2,4}|\mathbb{T}_n|(1+E[\|{}_nY^r\|_\infty^2])^\kappa\end{aligned}$$

with $c_{2,4} := 2^3(1+(T-r)^2/2)(T-r)\|h\|_{1,2,r}^2c^2\overline{c}^2(1+\underline{c}_2)^\kappa$, because $\Delta W_n^{(1)},\dots,\Delta W_{s_n}^{(d)}$ are pairwise independent and independent of $\mathscr{F}_{\underline{s}_n}$ for all $s\in[r,T]$.

The fifth term in (4.11) can be treated in a similar way as the third. Namely, we set ${}_{l,n}U_s := ({}_nW_s-{}_nW_{s_n}){}_n\dot{W}_s^{(l)}-(1/2)\gamma_n(s)e_l$ for every $s\in[r,T]$ and rewrite that

$$\begin{aligned}&\int_r^{t_{j,n}} \partial_x\overline{B}_{k,l}(t_{j,n},\underline{s}_n,{}_nY)\overline{B}(\underline{s}_n,\underline{s}_n,{}_nY)_{l,n}U_s\,ds\\&=\frac{1}{2}\sum_{i=1}^{j-1}\partial_x\overline{B}_{k,l}(t_{i,n},t_{i-1,n},{}_nY)\overline{B}(t_{i-1,n},t_{i-1,n},{}_nY)_{l,n}V_i\\&\quad+\frac{1}{2}\sum_{i_2=1}^{j-1}\int_{t_{i_2,n}}^{t_{i_2+1,n}}\sum_{i_1=1}^{i_2}\partial_t\partial_x\overline{B}_{k,l}(t,t_{i_1-1,n},{}_nY)\overline{B}(t_{i_1-1,n},t_{i_1-1,n},{}_nY)_{l,n}V_{i_1}\,dt\end{aligned}$$

for all $j\in\{1,\dots,k_n\}$, each $k\in\{1,\dots,m\}$ and every $l\in\{1,\dots,d\}$. Thus, from the estimate (4.10) we can again infer that

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

$$
\begin{aligned}
&\leq 2^2|\mathbb{T}_n| \sum_{i=1}^{k_n-1} \Delta t_{i,n} \sum_{k=1}^{m}\sum_{l=1}^{d} E\big[\big|\partial_x \overline{B}_{k,l}(t_{i,n}, t_{i-1,n}, {}_nY)\overline{B}(t_{i-1,n}, t_{i-1,n}, {}_nY)\big|^2\big] \\
&\quad + 2^2(T-r)\int_r^T \sum_{i=1}^{k_n-1} (\Delta t_{i,n})^2 \sum_{k=1}^{m}\sum_{l=1}^{d} E\big[\big|\partial_t\partial_x \overline{B}_{k,l}(t, t_{i-1,n}, {}_nY)\overline{B}(t_{i-1,n}, t_{i-1,n}, {}_nY)\big|^2\big]\,dt \\
&\leq c_{2,5}|\mathbb{T}_n|
\end{aligned}
$$

for $c_{2,5} := 2^2(1+(T-r)^2)(T-r)c^2\overline{c}^2$. For the sixth expression in (4.11) we decompose the integral and apply Itô's formula to the effect that

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

for all $j \in \{1,\dots,k_n\}$, each $k \in \{1,\dots,m\}$ and every $l \in \{1,\dots,d\}$. Hence, by utilizing that $\Delta W_{s_n}^{(1)},\dots,\Delta W_{s_n}^{(d)}$ are pairwise independent for any $s \in [r,T]$, we estimate that

$$
\begin{aligned}
&E\bigg[\max_{j\in\{0,\dots,k_n\}} \sum_{k=1}^{m}\bigg|\sum_{l=1}^{d}\int_r^{t_{j,n}} \partial_x \overline{B}_{k,l}(t_{j,n}, \underline{s}_n, {}_nY)\Sigma(\underline{s}_n, \underline{s}_n, {}_nY)(W_s - W_{s_n})\,d_nW_s^{(l)}\bigg|^2\bigg] \\
&\leq 2w_2|\mathbb{T}_n|\int_r^T \sum_{k=1}^{m}\sum_{l=1}^{d} E\big[\big|\partial_x \overline{B}_{k,l}(s, \underline{s}_n, {}_nY)\Sigma(\underline{s}_n, \underline{s}_n, {}_nY)\big|^2\big]\,ds \\
&\quad + 2w_2(T-r)\int_r^T\int_r^t \sum_{k=1}^{m}\sum_{l=1}^{d} E\big[\big|\partial_t\partial_x \overline{B}_{k,l}(t, \underline{s}_n, {}_nY)\Sigma(\underline{s}_n, \underline{s}_n, {}_nY)\big|^2\big]\Delta s_n\,ds\,dt \\
&\leq c_{2,6}|\mathbb{T}_n|,
\end{aligned}
$$

where $c_{2,6} := 2w_2(1+(T-r)^2/2)(T-r)c^2\overline{c}^2$.

Finally, for the seventh expression in (4.11) we define an $\mathbb{R}^m$-valued $\mathscr{F}_{t_{i-1,n}}$-measurable random vector by

$$
{}_{l,n}X_i := \frac{1}{\Delta t_{i+1,n}}\int_{t_{i,n}}^{t_{i+1,n}}\int_{t_{i-1,n}}^{s}\int_r^{t_{i-1,n}} \partial_v \overline{B}(v, u, {}_nY)\,d_nW_u\,dv\,ds,
$$

which satisfies $E\big[|{}_{l,n}X_i|^2\big] \leq 2^2\hat{w}_{2,1}(T-r)^2c_{\mathbb{T}}c^2(t_{i+1,n} - t_{i,n})$, for any $i \in \{1,\dots,k_n-1\}$ and every $l \in \{1,\dots,d\}$. Then we have

$$
\begin{aligned}
&\int_r^{t_{j,n}} \partial_x \overline{B}_{k,l}(t_{j,n}, \underline{s}_n, {}_nY)\bigg(\int_{\underline{s}_n}^{s}\int_r^{\underline{s}_n} \partial_v \overline{B}(v, u, {}_nY)\,d_nW_u\,dv\bigg)\,d_nW_s^{(l)} \\
&= \sum_{i=1}^{j-1} \partial_x \overline{B}_{k,l}(t_{i,n}, t_{i-1,n}, {}_nY)\,{}_{l,n}X_i \Delta W_{t_{i,n}}^{(l)} \\
&\quad + \sum_{i_2=1}^{j-1}\int_{t_{i_2,n}}^{t_{i_2+1,n}} \sum_{i_1=1}^{i_2} \partial_t\partial_x \overline{B}_{k,l}(t, t_{i_1-1,n}, {}_nY)\,{}_{l,n}X_{i_1}\Delta W_{t_{i_1,n}}^{(l)}\,dt
\end{aligned}
$$

for all $j \in \{1,\dots,k_n\}$, each $k \in \{1,\dots,m\}$ and every $l \in \{1,\dots,d\}$. As $\Delta W_{t_{i,n}}^{(1)},\dots,W_{t_{i,n}}^{(d)}$ are pairwise independent and independent of $\mathscr{F}_{t_{i-1,n}}$ for every $i \in \{1,\dots,k_n\}$, it follows that

$$E\Big[\max_{j\in\{0,\dots,k_n\}}\sum_{k=1}^m\Big|\sum_{l=1}^d\int_r^{t_{j,n}}\partial_x\overline{B}_{k,l}(t_{j,n},\underline{s}_n,{}_nY)\Big(\int_{\underline{s}_n}^s\int_r^{\underline{s}_n}\partial_v\overline{B}(v,u,{}_nY)\,d_nW_u\,dv\Big)\,d_nW_s^{(l)}\Big|^2\Big]$$
$$\leq 2^3\sum_{i=1}^{k_{n-1}}\Delta t_{i,n}\sum_{k=1}^m\sum_{l=1}^d E[|\partial_x\overline{B}_{k,l}(t_{i,n},t_{i-1,n},{}_nY)_{l,n}X_i|^2]$$
$$+2(T-r)\int_r^T\sum_{k=1}^m E\Big[\max_{j\in\{1,\dots,k_n\}}\Big|\sum_{i=1}^{j-1}\sum_{l=1}^d\partial_t\partial_x\overline{B}_{k,l}(t,t_{i-1,n},{}_nY)_{l,n}X_i\Delta W_{t_{i,n}}^{(l)}\Big|^2\Big]\,dt$$
$$\leq c_{2,7}|\mathbb{T}_n|$$

with $c_{2,7} := 2^5(1+(T-r)^2)(T-r)^3\hat{w}_{2,1}c_{\mathbb{T}}c^2\overline{c}^2$, by virtue of the estimate (4.10). Hence, we complete the proof by setting $c_2 := 7(c_{2,1}+\cdots+c_{2,7})$. □

5 Proofs of the convergence result in second moment and the support representation

5.1 Proofs of Lemmas 2.2 and 1.1

Proof of Lemma 2.2. (i) If $\overline{b}_0 = 0$ holds in (C.9), then (2.6) reduces to a pathwise Volterra integral equation. In this case, pathwise uniqueness and strong existence are covered by the deterministic results in [11] or can essentially be inferred from [15]. Otherwise, we may assume that $\overline{b}_0 = 1$ and introduce a martingale ${}_n\overline{Z} \in \mathscr{C}([0,T],\mathbb{R})$ by ${}_n\overline{Z}^r = 1$ and

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

for all $t \in [r,T]$ a.s. Then ${}_n\overline{W} \in \mathscr{C}([0,T],\mathbb{R}^d)$ defined via ${}_n\overline{W}_t := W_t + \int_r^{r\vee t}\overline{b}(s)\,d_nW_s$ is a d-dimensional $(\mathscr{F}_t)_{t\in[0,T]}$-Brownian motion under the probability measure $\overline{P}_n$ on $(\Omega,\mathscr{F})$ given by $\overline{P}_n(A) := E[\mathbb{1}_A\,{}_n\overline{Z}_T]$, by Girsanov's theorem.

We observe that a process $Y \in \mathscr{C}([0,T],\mathbb{R}^m)$ is a solution to (2.6) under P if and only if it solves the path-dependent stochastic Volterra integral equation

$$Y_t = Y_r + \int_r^t\underline{B}(t,s,Y) + B_H(t,s,Y)\dot{h}(s)\,ds + \int_r^t\Sigma(t,s,Y)\,d_n\overline{W}_s \tag{5.1}$$

a.s. for all $t \in [r,T]$ under $\overline{P}_n$. Consequently, pathwise uniqueness and strong existence follow from the stochastic results in [11] or [15] when considering the drift $\underline{B} + B_H\dot{h}$ and the diffusion Σ.

Regarding the claimed estimate, we let $p > 2$ and $\alpha \in [0,1/2-1/p)$. Then from Proposition 4.1 we obtain $c_p > 0$ such that (4.1) holds and the Kolmogorov-Chentsov estimate (3.1) implies that

$$E[(\|{}_nY\|_{\alpha,r} - \|{}_n\xi^r\|_\infty)^p] \leq k_{\alpha,p,p/2-1}c_p(T-r)^{p(1/2-\alpha)}(1+E[\|{}_n\xi^r\|_\infty^p])$$

for every $n \in \mathbb{N}$. Hence, we set $c_{\alpha,p} := 2^p k_{\alpha,p,p/2-1}(1+c_p)\max\{1, T-r\}^{p(1/2-\alpha)}$, then the triangle inequality gives the desired result.

(ii) Pathwise uniqueness, strong existence and the asserted bound can be directly inferred from (i) by replacing $\underline{B}$ by $\underline{B}+R$, $\overline{B}$ by 0 and Σ by $\overline{B}+\Sigma$, since (C.9) holds in this case with $\overline{b}=0$. ☐

Proof of Lemma 1.1. (i) Pathwise uniqueness, the existence of a unique strong solution and the integrability condition follow from assertion (ii) of Lemma 2.2 by letting $\underline{B} = b$, $B_H = \overline{B} = 0$, $\Sigma = \sigma$ and $\xi = \hat{x}$.

(ii) For $h \in W_r^{1,p}([0,T], \mathbb{R}^d)$ we set $F_h := b - (1/2)\rho + \sigma\dot{h}$. First, since $\partial_x\sigma(\cdot, s, x)$ is absolutely continuous on $[s,T)$, so is ρ and hence, F_h for any $s \in [r,T)$ and each $x \in C([0,T], \mathbb{R}^m)$. Secondly, there are $c_0, \lambda_0 \geq 0$ such that $\max\{|\sigma|, |\partial_t\sigma|, |\rho|, |\partial_t\rho|\} \leq c_0$ and

$$|\rho(s,s,x) - \rho(s,s,y)| + |\partial_t\rho(t,s,x) - \partial_t\rho(t,s,y)| \leq \lambda_0\|x-y\|_\infty$$

for all $s,t \in [r,T)$ with $s<t$ and every $x,y \in C([0,T],\mathbb{R}^m)$. These conditions ensure that the map F_h satisfies $|F_h(s,s,x)| + |\partial_t F_h(t,s,x)| \leq c_1(1+|\dot{h}(s)|)(1+\|x\|_\infty)$ and

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

for any $s,t \in [r,T)$ with $s<t$ and each $x,y \in C([0,T],\mathbb{R}^m)$, where $c_1 := 3\max\{c_0, c\}$ and $\lambda_1 := 2\max\{\lambda_0, \lambda\}$. As these are all the necessary assumptions, we invoke [11] to get a unique mild solution x_h to (1.5), which satisfies $x_h \in W_r^{1,p}([0,T],\mathbb{R}^m)$.

To show the the second assertion, we also let $g \in W_r^{1,p}([0,T],\mathbb{R}^d)$. Then for the constant $c_{p,1} := 2^{2p-2}(1+T-r)^p\max\{1,T-r\}^{p-1}\max\{c_0^p, \lambda_1^p\}$ we have

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

for each $t \in [r,T]$, since $\|y\|_\infty \leq \max\{1,T-r\}^{1-1/p}\|y\|_{1,p,r}$ for any $y \in W_r^{1,p}([0,T],\mathbb{R}^m)$. Hence, Gronwall's inequality gives $\|x_g - x_h\|_{1,p,r}^p \leq c_p\exp(c_p\|h\|_{1,p,r}^p)\|g-h\|_{1,p,r}^p$, where we have defined $c_p := c_{p,1}\exp((T-r)c_{p,1})$. ☐

5.2 Proofs of Theorems 2.3 and 1.2

Proof of Theorem 2.3. By Lemma 3.1, which is applicable due to Proposition 4.1 and Corollary 4.2, we merely have to show the first assertion, as the second follows from the first.

In this regard, the decomposition of Proposition 4.3 in second moment, Lemma 3.4 and a combination of the estimate (3.2) and Proposition 3.5 with Hölder's inequality show that this limit holds once we can justify that there is $c_2 > 0$ such that

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

does not exceed $c_2|\mathbb{T}_n|$ for every $n \in \mathbb{N}$. Based on the decomposition (4.7) and the hypothesis that $\partial_x\overline{B}$ is bounded, this fact follows from Proposition 4.4 and Lemma 4.5, in conjunction with Lemma 3.2 and Remark 3.3, and Proposition 4.6. ☐

Proof of Theorem 1.2. We let N_α denote the P-null set of all $\omega \in \Omega$ such that $X(\omega)$ fails to be α-Hölder continuous on $[r,T]$ and recall that the support of $P\circ X^{-1}$ in $C_r^\alpha([0,T],\mathbb{R}^m)$ coincides with the support of the inner regular probability measure

$$\mathscr{B}(C_r^\alpha([0,T],\mathbb{R}^m)) \to [0,1], \quad B \mapsto P(\{X \in B\} \cap N_\alpha^c). \tag{5.2}$$

Then an application of Theorem 2.3 in the case that $\underline{B} = b - (1/2)\rho$, $B_H = 0$, $\overline{B} = \sigma$, $\Sigma = 0$ and $\xi = \hat{x}$ gives us (2.4), which in turn implies that the support of (5.2) is included in the closure of $\{x_h \,|\, h \in W_r^{1,p}([0,T],\mathbb{R}^d)\}$ relative to $\|\cdot\|_{\alpha,r}$.

Now we let $h \in W_r^{1,p}([0,T],\mathbb{R}^d)$ be fixed and recall that for any $n \in \mathbb{N}$ and each $x \in C([0,T],\mathbb{R}^d)$ there is a unique mild solution $y_{h,n} \in C([0,T],\mathbb{R}^d)$ to the ordinary integral equation with running value condition

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

As the map $C([0,T],\mathbb{R}^d) \to C([0,T],\mathbb{R}^d)$, $x \mapsto y_{h,n,x}$ is Lipschitz continuous on bounded sets, we may let ${}_{h,n}W \in \mathscr{C}([0,T],\mathbb{R}^d)$ be given by ${}_{h,n}W_t := y_{h,n,W}(t)$ and introduce a martingale ${}_{h,n}Z \in \mathscr{C}([0,T],\mathbb{R})$ by requiring that ${}_{h,n}Z^r = 1$ and

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

for any $t \in [r,T]$ a.s. By Girsanov's theorem, ${}_{h,n}W$ is a d-dimensional $(\mathscr{F}_t)_{t\in[0,T]}$-Brownian motion under the probability measure $P_{h,n}$ on $(\Omega,\mathscr{F})$ given by $P_{h,n}(A) := E[\mathbb{1}_A\,{}_{h,n}Z_T]$ and X is a strong solution to the stochastic Volterra integral equation

$$X_t = X_r + \int_r^t b(t,s,X) + \sigma(t,s,X)\big(\dot{h}(s) - \dot{L}_n({}_{h,n}W)\big)(s)\,ds + \int_r^t \sigma(t,s,X)\,d_{h,n}W_s$$

for all $t \in [0,T]$ a.s. under $P_{h,n}$. Hence, let ${}_nY$ be the unique strong solution to (2.6) when $\underline{B} = b$, $B_H = \sigma$, $\overline{B} = -\sigma$ and $\Sigma = \sigma$ with ${}_nY^r = \hat{x}^r$ a.s., then uniqueness in law implies that $P(\|{}_nY - x_h\| \geq \varepsilon) = P_{h,n}(\|X - x_h\|_{\alpha,r} \geq \varepsilon)$ for any $\varepsilon > 0$. This shows that Theorem 2.3 also yields (2.5) and the claimed representation follows. ☐

References

[1] Shigeki Aida. Support theorem for diffusion processes on Hilbert spaces. *Publ. Res. Inst. Math. Sci.*, 26(6):947–965, 1990.

[2] V. Bally, L. Caramellino, and R. Cont. *Stochastic integration by parts and functional Itô calculus.* Advanced Courses in Mathematics. CRM Barcelona. Lecture notes of the Barcelona Summer School on Stochastic Analysis held in Barcelona, July 23–27, 2012.

[3] Vlad Bally, Annie Millet, and Marta Sanz-Solé. Approximation and support theorem in Hölder norm for parabolic stochastic partial differential equations. *Ann. Probab.*, 23(1):178–222, 1995.

[4] Gérard Ben Arous, Mihai Grădinaru, and Michel Ledoux. Hölder norms and the support theorem for diffusions. *Ann. Inst. H. Poincaré Probab. Statist.*, 30(3):415–436, 1994.

[5] R. Cont and D. Fournie. A functional extension of the Ito formula. *C. R. Math. Acad. Sci. Paris*, 348(1-2):57–61, 2010.

[6] R. Cont and D.-A. Fournié. Functional Itô calculus and stochastic integral representation of martingales. *Ann. Probab.*, 41(1):109–133, 2013.

[7] R. Cont and A. Kalinin. On the support of solutions to stochastic differential equations with path-dependent coefficients. *Stochastic Process. Appl.*, 2019.

[8] Rama Cont and Purba Das. Quadratic variation and quadratic roughness. *arXiv preprint arXiv:1907.03115*, 2019.

[9] B. Dupire. Functional Itô calculus. *Quant. Finance*, 19(5):721–729, 2019.

[10] I. Gyöngy and T. Pröhle. On the approximation of stochastic differential equation and on Stroock-Varadhan's support theorem. *Comput. Math. Appl.*, 19(1):65–70, 1990.

[11] A. Kalinin. Deterministic and stochastic volterra integral equation with path-dependency. *Preprint*, 2019.

[12] X. Mao. *Stochastic differential equations and applications*. Horwood Publishing Limited, Chichester, second edition, 2008.

[13] Annie Millet and Marta Sanz-Solé. A simple proof of the support theorem for diffusion processes. In *Séminaire de Probabilités, XXVIII*, volume 1583 of *Lecture Notes in Math.*, pages 36–48. Springer, Berlin, 1994.

[14] M. S. Pakkanen. Stochastic integrals and conditional full support. *J. Appl. Probab.*, 47(3):650–667, 2010.

[15] P. Protter. Volterra equations driven by semimartingales. *Ann. Probab.*, 13(2):519–530, 1985.

[16] Daniel 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 (Univ. California, Berkeley, Calif., 1970/1971), Vol. III: Probability theory*, pages 333–359, 1972.

[17] Mark Veraar. The stochastic Fubini theorem revisited. *Stochastics*, 84(4):543–551, 2012.