Annales de l'Institut Henri Poincaré - Probabilités et Statistiques
2021, Vol. 57, No. 1, 132–168
https://doi.org/10.1214/20-AIHP1074

ANNALES
DE L'INSTITUT
HENRI
POINCARÉ
PROBABILITÉS
ET STATISTIQUES

Skorohod and rough integration for stochastic differential equations driven by Volterra processes

Thomas Cass[a] and Nengli Lim[b]

[a]*Department of Mathematics, Imperial College London, Huxley Building, London, SW7 2AZ, United Kingdom. E-mail: thomas.cass@imperial.ac.uk*
[b]*Singapore University of Technology and Design, 8 Somapah Rd., Singapore 487372. E-mail: nengli_lim@sutd.edu.sg*

Received 17 June 2019; revised 13 May 2020; accepted 29 May 2020

Abstract. Given a solution Y to a rough differential equation (RDE), a recent result (*Ann. Probab.* **47** (2019) 1–60) extends the classical Itô-Stratonovich formula and provides a closed-form expression for $\int Y \circ \mathrm{d}\mathbf{X} - \int Y \, \mathrm{d}X$, i.e. the difference between the rough and Skorohod integrals of Y with respect to X, where X is a Gaussian process with finite p-variation less than 3. In this paper, we extend this result to Gaussian processes with finite p-variation such that $3 \leq p < 4$. The constraint this time is that we restrict ourselves to Volterra Gaussian processes with kernels satisfying a natural condition, which however still allows the result to encompass many standard examples, including fractional Brownian motion with Hurst parameter $H > \frac{1}{4}$. As an application we recover Itô formulas in the case where the vector fields of the RDE governing Y are commutative.

Résumé. Étant donnée Y une solution d'une équation différentielle rugueuse (RDE), un résultat récent (*Ann. Probab.* **47** (2019) 1–60) étend la formule d'Itô-Stratonovich et propose une expression explicite pour $\int Y \circ \mathrm{d}\mathbf{X} - \int Y \, \mathrm{d}X$, c'est-à-dire pour la différence entre l'intégrale rugueuse et l'intégrale de Skorohod de Y par rapport à X, où X est un processus Gaussien avec p-variation plus petite que 3. Dans cet article, nous étendons ce résultat au cas de processus Gaussiens avec p-variation telle que $3 \leq p < 4$. La contrainte ici est que nous nous restreignons au cas de processus Gaussiens de type Volterra avec des noyaux satisfaisant une condition naturelle, ce qui permet néanmoins de traiter beaucoup d'exemples classiques incluant le cas du mouvement Brownien fractionnaire avec paramètre de Hurst $H > \frac{1}{4}$. Comme application, nous retrouvons la formule d'Itô dans le cas où les champs de vecteurs de la RDE gouvernant Y sont commutatifs.

MSC2020 subject classifications: 60H07; 60L20

Keywords: Rough path theory; Volterra processes; Malliavin calculus

1. Introduction

Lyons' rough path theory is a framework for giving a path-wise interpretation to stochastic differential equations of the form

$$\mathrm{d}Y_t = V(Y_t) \circ \mathrm{d}X_t, \qquad Y_0 = y_0, \tag{1}$$

in particular for a broad class of continuous, vector-valued Gaussian processes X and sufficiently smooth vector fields V. A fundamental contribution of Lyons [24,25] was to realize that this needs X to be enriched to a *rough path* $\mathbf{X}$ whose components comprise not only X, but also the higher-order iterated integrals up to some finite degree. The model (1) can then be interpreted as a rough differential equation:

$$\mathrm{d}Y_t = V(Y_t) \circ \mathrm{d}\mathbf{X}_t, \qquad Y_0 = y_0. \tag{2}$$

In this paper we will assume that $X = (X^1, \ldots, X^d)$ has i.i.d components, each centered with covariance function R, and that X is defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. For simplicity we assume that $\mathcal{F}$ is generated by X. The process X then gives rise to an isonormal Gaussian process w.r.t. the Hilbert space $\mathcal{H}_1^d = \bigoplus_{i=1}^d \mathcal{H}_1^{(i)}$ where, for all $i = 1, \ldots, d$,

$\mathcal{H}_1^{(i)} = \mathcal{H}_1$ and $\mathcal{H}_1$ is the completion of the real vector space

$$\text{span}\{\mathbb{1}_{[0,t)}(\cdot) : |t \in [0, T]\},$$

endowed with the inner-product $\langle \mathbb{1}_{[0,t)}(\cdot), \mathbb{1}_{[0,s)}(\cdot)\rangle_{\mathcal{H}_1} = R(t, s)$. The solution Y to (2) can also be viewed as a Wiener functional on $(\Omega, \mathcal{F}, \mathbb{P})$, and its properties can be studied using the Malliavin calculus. A number of recent works have opened up the interplay between Lyons' and Malliavin's calculi, see e.g. [4,5,19] and [6]. In particular, in a recent paper [7], the authors have proven a conversion formula for the difference between the rough path integral of Y w.r.t $\mathbf{X}$ and the Skorohod integral δ^X of Y (i.e. the $L^2(\Omega)$ adjoint of the Malliavin derivative operator). In more detail this result shows, for the case where Y and X are both $\mathbb{R}^d$-valued, the following almost sure identity

$$\int_0^T \langle Y_t \circ \mathrm{d}\mathbf{X}_t\rangle - \delta^X(Y) = \frac{1}{2}\int_0^T \mathrm{tr}\big[V(Y_t)\big]\,\mathrm{d}R(t) + \int_{[0,T]^2} \mathbb{1}_{[0,t)}(s)\,\mathrm{tr}\big[J_t^{\mathbf{X}}\big(J_s^{\mathbf{X}}\big)^{-1}V(Y_s) - V(Y_t)\big]\,\mathrm{d}R(s,t). \tag{3}$$

Here, $J_t^{\mathbf{X}}$ denotes the Jacobian of the flow map $y_0 \to Y_t$, and the second part of the correction term is a proper 2D Young–Stieltjes integral (see [15,16]) with respect to the covariance function of X. When X is standard Brownian motion, this last term vanishes since the integrand is zero on the diagonal and $\mathrm{d}R(s,t) = \delta_{\{s=t\}}\,\mathrm{d}s\,\mathrm{d}t$. This, together with the fact that $R(t) := R(t,t) = t$, allows us to recover the classical Itô-Stratonovich conversion formula.

In [7], conditions need to be imposed in the proof of the formula (3) which limit the range of applications. An important assumption, for instance, is that the covariance function of X has finite (two-parameter) ρ-variation for $\rho \in [1, \frac{3}{2})$. This implies that the sample paths of X will have finite p-variation, for some $p \in [2, 3)$, and this excludes interesting examples such as fractional Brownian motion with $H \in (\frac{1}{4}, \frac{1}{3}]$.

The purpose of the present paper is to extend the correction formula (3) to these less regular cases. To do so we will assume that the Gaussian process X is a Volterra process; i.e. the covariance function R of each component can be written as

$$R(s,t) = \int_0^{t\wedge s} K(t,r)K(s,r)\,\mathrm{d}r,$$

for some kernel K, a square-integrable function $K : [0,T]^2 \to \mathbb{R}$ with $K(t,s) = 0, \forall s \geq t$. We will present conditions on K that allow us to generalize (3). In doing so, we need to overcome a number of serious obstacles. We highlight here the three most salient of these, outline the contribution of the present work and, at the same time, provide a road-map for the paper:

(i) We need to prove that the solution Y belongs to the domain of the Skorohod integral δ^X. In fact, we prove the stronger statement that Y belongs to the Malliavin Sobolev space $\mathbb{D}^{1,2}(\mathcal{H}_1^d) \subset \mathrm{Dom}(\delta^X)$. To show that Y, a path-valued random variable, can be understood as a random variable in the Hilbert space $\mathcal{H}_1^d$, we need to identify a class of functions with a subset of $\mathcal{H}_1^d$. This was proved in [7], by taking advantage of the assumption that $\rho \in [1, \frac{3}{2})$, but the less regular cases need a new argument that exploits the structure of the Volterra kernel. To handle the Malliavin derivative $\mathcal{D}Y$, we need a similar result that identifies a class of two-parameter functions as a subset of $\mathcal{H}_1^d \otimes \mathcal{H}_1^d$.

(ii) For the examples considered in this paper, the Gaussian rough path $\mathbf{X}$ will consist of iterated integrals up to degree three; i.e. $\mathbf{X} = (1, X, \mathbf{X}^2, \mathbf{X}^3)$. This contrasts with the result in [7], where only the case $\mathbf{X} = (1, X, \mathbf{X}^2)$ needs to be considered. This increases the complexity of the arguments significantly; indeed, the rough integral in the left side of (3) is now well approximated locally by terms up to third-order

$$\int_s^t \langle Y_r \circ \mathrm{d}\mathbf{X}_r\rangle \simeq \langle Y_s, X_{s,t}\rangle + V(Y_s)\mathbf{X}_{s,t}^2 + V^2(Y_s)\big(\mathbf{X}_{s,t}^3\big).$$

A key step in [7] is the proof that the second-order terms in this approximation satisfy

$$\lim_{\|\pi(n)\|\to 0}\left\| \sum_{i:\pi(n)=\{t_i^n\}} V(Y_{t_i^n})\left(\mathbf{X}^2_{t_i^n,t_{i+1}^n} - \frac{1}{2}\sigma^2\big(t_i^n, t_{i+1}^n\big)\mathcal{I}_d\right)\right\|_{L^2(\Omega)} = 0.$$

For the present work we need to address the same problem for the third order terms, namely the existence of an $L^2(\Omega)$-limit for sums of terms of the form

$$V^2(Y_{t_i^n})\big(\mathbf{X}^3_{t_i^n,t_{i+1}^n}\big)$$

over a sequence of partitions with mesh tending to zero. An important discovery of this paper is the somewhat surprising conclusion that these terms have vanishing $L^2(\Omega)$-limit, without the need to subtract any re-balancing terms. This is the concluding result of Section 4.

(iii) The proof of point (ii) relies on a rather intricate interplay between estimates from Malliavin's calculus and rough path analysis. From the latter theory, we need estimates on the directional Malliavin derivatives of RDE solutions. It is well known that an RDE solution of the form (2) can be differentiated in a direction $h \in C^{q-\mathrm{var}}([0,T],\mathbb{R}^d)$ by considering the perturbed RDE solution driven by the translated rough path $T_{\epsilon h}\mathbf{X}$ and then evaluating the derivative in ϵ at zero. For $T_{\epsilon h}\mathbf{X}$ to make sense, $\mathbf{X}$ and h must have Young-complementary regularity, i.e. $\frac{1}{p}+\frac{1}{q}>1$, in which case Duhamel's formula gives

$$\mathcal{D}_h Y_t = \int_0^t J_t^{\mathbf{X}}\big(J_s^{\mathbf{X}}\big)^{-1} V(Y_s)\,\mathrm{d}h(s), \tag{4}$$

a well-defined Young integral. In Malliavin calculus, h will typically be an element of the Cameron–Martin space (written as $\mathcal{H}^d$ in this paper), and this has spurred interest in results that prove that $\mathcal{H}^d$ can be continuously embedded into q-variation spaces, see e.g. [5,13]. By combining these results with Young's inequality, one can then say e.g. that

$$|\mathcal{D}_h Y_t| \lesssim |h|_{q\text{-var}} \lesssim |h|_{\mathcal{H}^d}, \tag{5}$$

and these arguments can be generalized to higher order directional derivatives, allowing one control over the Hilbert–Schmidt norm of the Malliavin derivative; see [21]. Note however, that quality is lost in (5) by use of the embedding. For the proof in (ii) we need subtler estimates on the higher order derivatives of the form

$$\big|\mathcal{D}^n_{h_i,\dots,h_n} Y_t\big| \le C_n(\mathbf{X}) \prod_{j=1}^n |h_j|_{q\text{-var}}. \tag{6}$$

The derivatives of order 2 and higher complicate matters because they are no longer representable as Young integrals as in (4); instead genuine rough integrals appear. Much of the work underpinning point (ii) goes into deriving closed-form expressions for these higher order derivatives and then estimating them so as to arrive at (6). We must also pay careful attention to the random variable $C_n(\mathbf{X})$ in (6) which, for our application, must have finite positive moments of all orders. The first half of Section 4 is devoted to this material.

The culmination of these arguments is presented in Section 5, where we give a set of conditions under which a conversion formula holds for $\int_0^T \langle Y_t \circ \mathrm{d}\mathbf{X}_t\rangle - \delta^X(Y)$. This formula is reminiscent of the one obtained for the case of second-order rough paths, but there are interesting differences too. Most notably the second term in (3),

$$\int_{[0,T]^2} \mathbb{1}_{[0,t)}(s)\,\mathrm{tr}\big[J_t^{\mathbf{X}}\big(J_s^{\mathbf{X}}\big)^{-1} V(Y_s) - V(Y_t)\big]\,\mathrm{d}R(s,t), \tag{7}$$

which exists for $2 \le p < 3$ as a well-defined 2D Young–Stieltjes integral, can only be identified as an L^2-limit of a sequence of approximating sums. The difference between the two cases stems from the lack of complementary Young regularity of the integrand and R. Interestingly the integrand, while being continuous on $[0,T]^2$, is not Hölder bi-continuous and so we cannot even appeal to the relaxed criteria discussed in point (i) above. It is unknown at present whether the limit is interpretable as a 2D Young–Stieltjes integral. We discuss in detail two important corollaries of our result. The first is where X is a fractional Brownian motion with H in $(\frac{1}{4},\frac{1}{3}]$, and the second is the case where the vector fields defining (2) commute. In this latter case, we show that the second term (7) in the correction formula disappears and, as a special case, we can recover Itô-type formulas for Gaussian processes, thus connecting our work to a substantial recent corpus e.g. [2,3,20,28] and [29].

2. Preliminaries

2.1. *Rough path concepts and notation*

We briefly review the basic notation used in this article; the standard references [14,24,26] and [16] can be consulted for more detail. We let $T^n(\mathbb{R}^d)$ denote the degree n truncated tensor algebra $T^n(\mathbb{R}^d) := \mathbb{R} \oplus \mathbb{R}^d \oplus \cdots \oplus (\mathbb{R}^d)^{\otimes n}$ equipped with addition, scalar multiplication and the (truncated) tensor product defined in the usual way. The unit element is $e =$

$(1, 0 \ldots, 0)$, and $T_e^n(\mathbb{R}^d) \subset T^n(\mathbb{R}^d)$ denotes the Lie group of tensors whose zeroth order term equals unity; its Lie algebra is $A_T^n(\mathbb{R}^d)$. The exponential and logarithm maps are written as $\exp : A_T^n(\mathbb{R}^d) \to T_e^n(\mathbb{R}^d)$ and $\log : T_e^n(\mathbb{R}^d) \to A_T^n(\mathbb{R}^d)$. We let $G^n(\mathbb{R}^d)$ will be the step-n nilpotent group with d generators equipped with some (any) symmetric, sub-additive homogeneous norm $\|\cdot\|$ which induces a left-invariant metric d, allowing one to define a p-variation distance on the space of $G^n(\mathbb{R}^d)$-valued paths in the customary way, cf. [15].

For $p \geq 1$, the set of weakly geometric (resp. geometric) p-rough paths is denoted by $\mathcal{C}^{p\text{-var}}([0,T]; G^{\lfloor p \rfloor}(\mathbb{R}^d))$ (resp. $\mathcal{C}^{0,p\text{-var}}([0,T]; G^{\lfloor p \rfloor}(\mathbb{R}^d))$). Given a Banach space $(E, \|\cdot\|_E)$, $V^{p\text{-var}}([0,T]; E)$ is the set of E-valued paths of finite p-variation w.r.t. the norm on E; the subspace of continuous (resp. piecewise continuous) paths is $\mathcal{C}^{p\text{-var}}([0,T]; E)$ (resp. $\mathcal{C}_{pw}^{p\text{-var}}([0,T]; E)$). For such paths f, we define the p-variation norm:

$$\|f\|_{\mathcal{V}^p;[0,T]} := \|f\|_{p\text{-var};[0,T]} + \sup_{t \in [0,T]} \|f_t\|_E.$$

The extension of p-variation to two-parameter functions will be heavily used. The definition hinges on the notion of the rectangular increment which for a function $f : [0,T]^2 \to E$ is given on the rectangle $[s,t] \times [u,v]$ by

$$f\begin{pmatrix} s, t \\ u, v \end{pmatrix} := f(s,u) + f(t,v) - f(s,v) - f(t,s); \tag{8}$$

see [15,30] for a complete description. On occasion, we will use the notation

$$f(\Delta_i, v) := f(u_{i+1}, v) - f(u_i, v)$$

and similarly for $f(u, \Delta_j)$. Two functions f and g defined on a rectangle will be said to have complementary regularity if they respectively have finite p and q variation such that $p^{-1} + q^{-1} > 1$. In this case, the 2D Young–Stieltjes integral of f against g (and *vice versa*) exists, and the following 2-parameter version of Young's inequality holds:

$$\left| \int_{[s,t]\times[u,v]} f \, \mathrm{d}g \right| \leq C_{p,q} \left\|\!\left| f \right|\!\right\| \|g\|_{q\text{-var},[s,t]\times[u,v]}, \tag{9}$$

where

$$\left\|\!\left| f \right|\!\right\| = \left| f(s,u) \right| + \left\| f(s,\cdot) \right\|_{p\text{-var};[u,v]} + \left\| f(\cdot,u) \right\|_{p\text{-var};[s,t]} + \|f\|_{p\text{-var},[s,t]\times[u,v]};$$

see Theorem 2.12 of [7,30] and [15].

2.2. *Gaussian rough paths*

We will work with a continuous Gaussian process $X_t = (X_t^{(1)}, \ldots, X_t^{(d)}) \in \mathbb{R}^d$, $t \in [0,T]$, which is assumed to have zero-mean and to have i.i.d. components, defined on the canonical completed probability space $(\Omega, \mathcal{F}, \mathbb{P})$, where $\Omega = \mathcal{C}([0,T]; \mathbb{R}^d)$. The common covariance function of the components is $R : [0,T]^2 \to \mathbb{R}$, the variance $R(t,t)$ will be abbreviated by using $R(t)$, and for the rectangular increment of R over $[s,t]^2$ (recall (8)) we use the notation $\sigma^2(s,t)$ in place of $\mathbb{E}[(X_{s,t}^{(1)})^2]$. The Hilbert space $\mathcal{H}^d = \bigoplus_{i=1}^d \mathcal{H}$ is the Cameron–Martin space (or reproducing kernel Hilbert space), which is densely and continuously embedded in Ω, and is given abstractly as the completion of the linear span of the functions

$$\left\{ R(t,\cdot)^{(u)} := R(t,\cdot) e_u | t \in [0,T], u = 1, \ldots, d \right\}$$

under the inner-product

$$\left\langle R(t,\cdot)^{(u)}, R(s,\cdot)^{(v)} \right\rangle_{\mathcal{H}^d} = \delta_{uv} R(t,s), \quad u, v = 1, \ldots, d,$$

wherein $\{e_u\}_{u=1}^d$ is the standard basis of $\mathbb{R}^d$ and δ_{uv} is the Kronecker delta. The reproducing property is captured by $\langle f., R(t,\cdot)^{(u)} \rangle_{\mathcal{H}^d} = f_t^{(u)}$, $t \in [0,T]$ for any $f = (f^{(1)}, \ldots, f^{(d)}) \in \mathcal{H}^d$.

If R has finite 2D ρ-variation for ρ in $[1,2)$, then [15] proved that X lifts to a geometric p-rough path for $p > 2\rho$ by taking limits of smooth approximations (see also [9] for the case of fractional Brownian motion). Moreover, Proposition 17 in [15] shows that for all $h \in \mathcal{H}^d$ the following embedding holds

$$\|h\|_{\rho\text{-var};[0,T]} \leq \|h\|_{\mathcal{H}^d} \sqrt{\|R\|_{\rho\text{-var};[0,T]^2}},$$

Thus if $\rho \in [1, \frac{3}{2})$, the sample paths of X will be complementary regular (a.s.) w.r.t. any Cameron–Martin path. In the case $\rho \in [\frac{3}{2}, 2)$, Theorem 1 in [13] shows that if R satisfies the stronger condition of mixed $(1, \rho)$-variation [30], then there exists $C < \infty$ such that $\|h\|_{q\text{-var};[0,T]} \le C\|h\|_{\mathcal{H}^d}$ for all h in $\mathcal{H}^d$, where $q = 2\rho(\rho+1)^{-1}$. One can easily verify that this gives complementary regularity as long as $p < 4$.

The following condition collects the assumptions we impose on X, or equivalently R.

Condition 1. *Let X be a continuous, centered Gaussian process in $\mathbb{R}^d$ with i.i.d. components. We assume that*

(a) $\|R\|_{\rho\text{-var};[0,T]^2} < \infty$ *for some* $\rho \in [1, 2)$,

(b) *the geometric rough path lift of X is of finite p-variation,* $p \in [1, 4)$, *and that there exists* $q \ge 1$ *satisfying* $p^{-1} + q^{-1} > 1$ *such that* $\|h\|_{q\text{-var};[0,T]} \le C\|h\|_{\mathcal{H}^d}$ *for all* $h \in \mathcal{H}^d$.

For some theorems, we will need to impose further conditions on the covariance function to control the $L^2(\Omega)$ norm of the iterated integrals. In these cases, we will assume there exists $C < \infty$ such that

$$\big\|R(t,\cdot) - R(s,\cdot)\big\|_{q\text{-var};[0,T]} \le C|t-s|^{\frac{1}{\rho}}, \quad \forall s,t \in [0,T]. \tag{10}$$

2.3. *Volterra processes and fractional Brownian motion*

A *Volterra kernel* K is a square-integrable function $K : [0,T]^2 \to \mathbb{R}$ such that $K(t,s) = 0 \forall s \ge t$. Associated with any Volterra kernel is a lower triangular, Hilbert–Schmidt operator $\mathbb{K} : L^2([0,T]) \to L^2([0,T])$ given by

$$\mathbb{K}(f)(\cdot) = \int_0^T K(\cdot, s) f(s)\,\mathrm{d}s \quad \text{for all } f \in L^2\big([0,T]\big).$$

Given a standard Brownian motion B and a Volterra kernel K, we define a Volterra process $X = (X_t)_{t\in[0,T]}$ as the Itô integral

$$X_t = \int_0^t K(t,s)\,\mathrm{d}B_s;$$

this is a centered Gaussian process with covariance function

$$R(s,t) = \int_0^{t\wedge s} K(t,r)K(s,r)\,\mathrm{d}r.$$

We will consider Volterra processes for which the following conditions prevail.

Condition 2. *There exist constants $C < \infty$ and $\alpha \in [0, 1/4)$ such that*

(i) $|K(t,s)| \le Cs^{-\alpha}(t-s)^{-\alpha}$ *for all* $0 < s < t \le T$;

(ii) $\frac{\partial K(t,s)}{\partial t}$ *exists for all* $0 < s < t \le T$ *and satisfies* $|\frac{\partial K(t,s)}{\partial t}| \le C(t-s)^{-(\alpha+1)}$.

3. Convergence in $\mathbb{D}^{1,2}(\mathcal{H}_1^d)$

In this section, we will discuss the various isomorphisms and subspaces of the Cameron–Martin space and its tensor product. The motivation is as follows: let Y be a solution to RDE (2) and given a partition $\pi = \{r_i\}$ of $[0,T]$, denote

$$Y^\pi(t) := \sum_i Y_{r_i} \mathbb{1}_{[r_i, r_{i+1})}(t).$$

Now recall the following inequality from Proposition 1.3.1 in [27]

$$\mathbb{E}\big[\delta^X\big(Y^\pi - Y\big)^2\big] \le \mathbb{E}\big[\big\|Y^\pi - Y\big\|^2_{\mathcal{H}_1^d}\big] + \mathbb{E}\big[\big\|\mathcal{D}Y^\pi - \mathcal{D}Y\big\|^2_{\mathcal{H}_1^d\otimes\mathcal{H}_1^d}\big], \tag{11}$$

which in particular implies that $\mathrm{Dom}(\delta^X) \supseteq \mathbb{D}^{1,2}(\mathcal{H}_1^d)$, where we use $\mathcal{H}_1^d$ to denote the completion of the linear span of

$$\big\{\mathbb{1}^{(u)}_{[0,t)}(\cdot) := \mathbb{1}_{[0,t)}(\cdot)e_u \,|\, t \in [0,T], u = 1, \dots, d\big\}$$

(cf. [2,27]) with respect to the inner-product given by

$$\langle \mathbb{1}^{(u)}_{[0,t)}(\cdot), \mathbb{1}^{(v)}_{[0,s)}(\cdot)\rangle_{\mathcal{H}^d_1} = \delta_{uv} R(t,s).$$

Thus if we can show that almost surely, Y and DY can be identified as elements of $\mathcal{H}^d_1$ and $\mathcal{H}^d_1 \otimes \mathcal{H}^d_1$ respectively, and furthermore $\|Y^\pi - Y\|_{\mathcal{H}^d_1}$ and $\|\mathcal{D}Y^\pi - \mathcal{D}Y\|_{\mathcal{H}^d_1\otimes\mathcal{H}^d_1}$ both vanish as $\|\pi\| \to 0$, then with further integrability assumptions one can use (11) and dominated convergence to show that $\delta^X(Y^\pi)$ converges to $\delta^X(Y)$ in $L^2(\Omega)$.

Given a Banach space E and a Volterra kernel K satisfying Condition 2 for some $\alpha \in [0, 1/4)$, we introduce the linear operator $\mathcal{K}^*$ (see [2,10])

$$\big(\mathcal{K}^*\phi\big)(s) := \phi(s)K(T,s) + \int_s^T \big[\phi(r) - \phi(s)\big] K(\mathrm{d}r, s), \tag{12}$$

where the signed measure $K(\mathrm{d}r,s) := \frac{\partial K(r,s)}{\partial r}\,\mathrm{d}r$. The domain $D(\mathcal{K}^*)$ of $\mathcal{K}^*$ consists of measurable functions $\phi : [0,T] \to E$ for which the integral on the right-hand side exists for all s in $[0,T]$.

Remark 3.1. Note in particular that if ϕ is a λ-Hölder continuous function in the norm of E for some $\lambda > \alpha$, then $\phi \in D(\mathcal{K}^*)$ and $\mathcal{K}^*\phi$ is in $L^2([0,T]; E)$. Also for any a in $[0,T]$, $\phi\mathbb{1}_{[0,a)}$ is in $D(\mathcal{K}^*)$ whenever ϕ is, and we have the identity

$$\mathcal{K}^*(\phi\mathbb{1}_{[0,a)})(s) = \mathbb{1}_{[0,a)}(s)\left(\phi(s)K(a,s) + \int_s^a \big[\phi(r) - \phi(s)\big] K(\mathrm{d}r, s)\right). \tag{13}$$

3.1. *Convergence in $\mathcal{H}^d_1$*

The main aim of this subsection is to investigate the (almost sure) regularity required of Y to identify it as an element of $\mathcal{H}^d_1$, and to have $\|Y^\pi - Y\|_{\mathcal{H}^d_1} \to 0$. For Volterra processes, the first issue is to find criteria ensuring that the step-function approximations to a given Hölder continuous function converge in $\mathcal{H}^d_1$. We recall the following result from [23] (see also Proposition 8 of [1]).

Proposition 3.2. *Let $(E, \|\cdot\|_E)$ be a Banach space and $K : [0,T]^2 \to \mathbb{R}$ be a kernel satisfying Condition 2 for some $\alpha \in [0, \frac{1}{4})$. Let $\phi : [0,T] \to E$ be λ-Hölder continuous, i.e. there exists $C < \infty$ such that*

$$\big\|\phi(t_1) - \phi(t_2)\big\|_E \le C|t_1 - t_2|^\lambda, \quad \forall t_1, t_2 \in [0,T],$$

and for any partition $\pi = \{s_i\}$ of $[0,T]$, let $\phi^\pi : [0,T] \to E$ denote

$$\phi^\pi(t) = \sum_i \phi(s_i)\mathbb{1}_{[s_i,s_{i+1})}(t).$$

Then if $\lambda > \alpha$ we have

$$\lim_{\|\pi\|\to 0} \int_0^T \big\|\mathcal{K}^*\big(\phi^\pi - \phi\big)(t)\big\|_E^2 \,\mathrm{d}t = 0,$$

where $\mathcal{K}^*$ is defined as in (12).

Rather than dealing with the Hilbert space $\mathcal{H}^d_1$ as an abstract completion, it will be useful to realize it as a closed subspace of an L^2 space. To this end, we define $\mathcal{H}^d_2$ to be the closure in $L^2([0,T]; \mathbb{R}^d)$ of the linear subspace generated by

$$\big\{K(t,\cdot)^{(u)} := K(t,\cdot)e_u \,|\, t \in [0,T], u = 1,\ldots,d\big\}.$$

The inner-product is the usual one in $L^2([0,T]; \mathbb{R}^d)$, namely $\langle f, g\rangle_{\mathcal{H}^d_2} = \int_0^T \langle f_s, g_s\rangle\,\mathrm{d}s$ where $\langle\cdot,\cdot\rangle$ denotes the Euclidean inner-product in $\mathbb{R}^d$. The following proposition is more or less immediate (see Proposition 2.2.4 of [22]).

Proposition 3.3. *$\mathcal{H}^d_1$ and $\mathcal{H}^d_2$ are isomorphic as Hilbert spaces under the map $\mathcal{K}^*$.*

Remark 3.4. In the case of standard Brownian motion the isomorphism $\mathcal{K}^*$ is the identity operator and $\mathcal{H}_1^d = \mathcal{H}_2^d = L^2([0,T];\mathbb{R}^d)$.

Since the RDE solutions we work with are path-valued, it will be convenient to find subspaces of $\mathcal{H}_1^d$ whose elements are actual paths. We let

$$\Lambda_\alpha^d := \bigcup_{\lambda>\alpha} \mathcal{C}_{pw}^{\lambda\text{-Höl}}\big([0,T];\mathbb{R}^d\big),$$

where $\mathcal{C}_{pw}^{\lambda\text{-Höl}}([0,T];\mathbb{R}^d)$ denotes the space of piecewise λ-Hölder continuous functions. By equipping Λ_α^d with the inner-product

$$\langle f,g\rangle_{\Lambda_\alpha^d} := \big\langle \mathcal{K}^*(f), \mathcal{K}^*(g)\big\rangle_{L^2([0,T];\mathbb{R}^d)},$$

whilst suppressing its dependence on K in the notation, the following proposition shows that we can regard Λ_α^d as a dense subspace of $\mathcal{H}_1^d$.

Proposition 3.5. *Suppose K is a kernel satisfying Condition 2 for some $\alpha \in [0,\frac{1}{4})$. Then Λ_α^d is a dense subspace of $\mathcal{H}_1^d$, and the inclusion map $i : (\Lambda_\alpha^d, \langle\cdot,\cdot\rangle_{\Lambda_\alpha^d}) \to (\mathcal{H}_1^d, \langle\cdot,\cdot\rangle_{\mathcal{H}_1^d})$ is an isometry.*

Proof. Let $f \in \Lambda_\alpha^d$ and let $\pi(n) = \{r_i^{(n)}\}$ be a sequence of partitions whose mesh vanishes as $n \to \infty$. We define

$$f^{\pi(n)}(t) := \sum_i f\big(r_i^{(n)}\big)\mathbb{1}_{[r_i^{(n)}, r_{i+1}^{(n)})}(t).$$

Note that for each n, $f^{\pi(n)}$ is in $\Lambda_\alpha^d \cap \mathcal{H}_1^d$. Moreover, Proposition 3.2 tells us that $\|\mathcal{K}^*(f^{\pi(n)} - f)\|_{L^2([0,T];\mathbb{R}^d)} \to 0$. Hence, using the fact from Proposition 3.3 that $\|f\|_{\mathcal{H}_1^d} = \|\mathcal{K}^* f\|_{\mathcal{H}_2^d}$ for all $f \in \mathcal{H}_1^d$, we see that $f^{\pi(n)}$ is Cauchy in $\mathcal{H}_1^d$. We again identify f with the limit of the sequence, and under this identification we have

$$\|f\|_{\mathcal{H}_1^d} = \big\|\mathcal{K}^*(f)\big\|_{L^2([0,T];\mathbb{R}^d)}. \tag{14}$$

Since Λ_α^d contains all the generating functions $\{\mathbb{1}_{[0,t)}^{(u)}(\cdot)\}$ of $\mathcal{H}_1^d$, its closure is $\mathcal{H}_1^d$. □

We recall from [7] a similar result in terms of p-variation. In that paper, $\mathcal{H}_1^d$ was derived from a Gaussian covariance function R which was assumed to be of finite 2D ρ-variation, $\rho \in [1,2)$. It was shown that

$$\mathcal{W}_\rho^d := \bigcup_{q<\frac{\rho}{\rho-1}} \mathcal{C}_{pw}^{q\text{-var}}\big([0,T];\mathbb{R}^d\big), \tag{15}$$

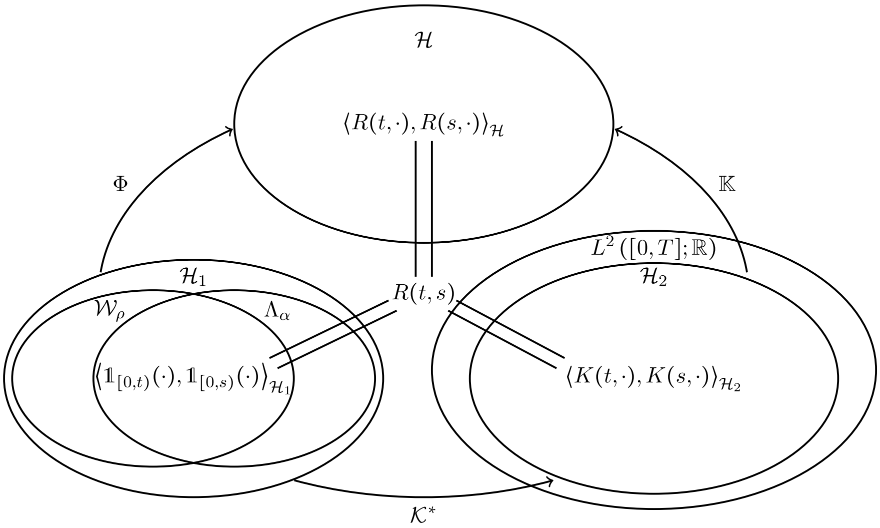
when equipped with the inner-product

$$\langle f,g\rangle_{\mathcal{W}_\rho^d} := \int_{[0,T]^2} \langle f_s, g_t\rangle_{\mathbb{R}^d}\, \mathrm{d}R(s,t),$$

is a dense subspace of $\mathcal{H}_1^d$ with the inclusion map again being an isometry. In the case when $\lambda > \alpha \wedge (1-\frac{1}{\rho})$, any f and g belonging to $\mathcal{C}_{pw}^{\lambda\text{-Höl}}([0,T];\mathbb{R}^d)$ also belong to $\mathcal{W}_\rho^d \cap \Lambda_\alpha^d$, and we have

$$\langle f,g\rangle_{\mathcal{W}_\rho^d} = \int_{[0,T]^2} \langle f_s, g_t\rangle_{\mathbb{R}^d}\, \mathrm{d}R(s,t) = \int_0^T \big\langle \mathcal{K}^* f(r), \mathcal{K}^* g(r)\big\rangle_{\mathbb{R}^d}\, \mathrm{d}r = \langle f,g\rangle_{\Lambda_\alpha^d}.$$

The following figure depicts schematically the relationship between the various subspaces in the case of a Volterra process satisfying Condition 2 (assuming $d = 1$ for convenience).



Note that $\mathbb{K}$ gives an isomorphism from $\mathcal{H}_2^d$ onto $\mathcal{H}^d$ because $R(t,\cdot)=\int_0^T K(\cdot,r)K(t,r)\,\mathrm{d}r$, and we let $\Phi:\mathcal{H}_1^d\to\mathcal{H}^d$ denote the Hilbert space isomorphism obtained from extending the map $\mathbb{1}_{[0,t)}^{(u)}(\cdot)\mapsto R(t,\cdot)^{(u)}$, $t\in[0,T]$, $u=1,\dots,d$.

3.2. *Convergence in $\mathcal{H}_1^d\otimes\mathcal{H}_1^d$*

The results of the previous subsection allow us to interpret RDE solutions (paths) as $\mathcal{H}_1^d$-valued random variables. The Malliavin derivatives of these random variables, when they exist, will take values in $\mathcal{H}_1^d\otimes\mathcal{H}_1^d$, and we therefore need similar results which identify suitable function spaces which are subspaces of this tensor product space.

Throughout, E will denote a general Banach space with norm $\|\cdot\|_E$. The following operator was defined in [23].

Definition 3.6. Let $\mathcal{K}^*\otimes\mathcal{K}^*$ denote the operator

$$\begin{aligned}\big(\mathcal{K}^*\otimes\mathcal{K}^*\big)\psi(u,v) &:= \psi(u,v)K(T,v)K(T,u)+K(T,v)A^K\big(\psi(\cdot,v)\big)(u)\\ &\quad + K(T,u)A^K\big(\psi(u,\cdot)\big)(v)+B^K(\psi)(u,v),\end{aligned}$$

where

$$\begin{aligned}A^K(\phi)(s) &:= \int_s^T\big[\phi(r)-\phi(s)\big]K(\mathrm{d}r,s),\\ B^K(\psi)(u,v) &:= \int_v^T\int_u^T\psi\begin{pmatrix}u & r_1\\ v & r_2\end{pmatrix}K(\mathrm{d}r_1,u)K(\mathrm{d}r_2,v),\end{aligned}$$

which is defined for any measurable function $\psi:[0,T]^2\to E$ for which the integrals on the right side exist.

Using Proposition 3.3 and the fact that

$$\big(\mathcal{K}^*\otimes\mathcal{K}^*\big)\psi(s,t)=\big(\mathcal{K}^*\psi_1\big)(s)\otimes\big(\mathcal{K}^*\psi_2\big)(t) \tag{16}$$

when $\psi(s,t)=\psi_1(s)\psi_2(t)$, it is also clear that $\mathcal{K}^*\otimes\mathcal{K}^*$ maps $\mathcal{H}_1^d\otimes\mathcal{H}_1^d$ isometrically onto $\mathcal{H}_2^d\otimes\mathcal{H}_2^d$, which is a closed subspace of $L^2([0,T];\mathbb{R}^d)\otimes L^2([0,T];\mathbb{R}^d)\cong L^2([0,T]^2;\mathbb{R}^d\otimes\mathbb{R}^d)$.

To go beyond product functions in the domain of $\mathcal{K}^*\otimes\mathcal{K}^*$, we also recall the class of strongly Hölder bi-continuous functions from [23].

Definition 3.7. Let $0<\lambda\le 1$. We say that a function $\phi:[0,T]^2\to E$ is strongly λ-Hölder bi-continuous in the norm of E (or simply strongly λ-Hölder bi-continuous in the case where E is finite-dimensional), if for all $u_1,u_2,v_1,v_2\in[0,T]$ we have

$$\sup_{v\in[0,T]}\big\|\phi(u_2,v)-\phi(u_1,v)\big\|_E\le C|u_2-u_1|^\lambda,\qquad \sup_{u\in[0,T]}\big\|\phi(u,v_2)-\phi(u,v_1)\big\|_E\le C|v_2-v_1|^\lambda,$$

and

$$\left\| \phi \begin{pmatrix} u_1 & u_2 \\ v_1 & v_2 \end{pmatrix} \right\|_E \leq C|u_2 - u_1|^\lambda |v_2 - v_1|^\lambda. \tag{17}$$

The following result and its proof can be found as Theorem 4.1 in [23].

Theorem 3.8. *Let $\psi : [0,T]^2 \to E$ be a function which is strongly λ-Hölder bi-continuous in the norm of E. For any partition $\pi = \{(u_i, v_j)\}$ of $[0,T]^2$, let $\psi^\pi : [0,T]^2 \to E$ denote*

$$\psi^\pi(u,v) := \sum_{i,j} \psi(u_i, v_j) \mathbb{1}_{[u_i,u_{i+1})}(u) \mathbb{1}_{[v_j,v_{j+1})}(v).$$

In addition, let $\mathcal{K}^* \otimes \mathcal{K}^*$ denote the operator in Definition 3.6, where the Volterra kernel K satisfies Condition 2 for some $\alpha \in [0, \frac{1}{4})$. Then if $\lambda > \alpha$, we have

$$\lim_{\|\pi\| \to 0} \int_{[0,T]^2} \left\| \left(\mathcal{K}^* \otimes \mathcal{K}^* \left(\psi^\pi - \psi\right)\right)(u,v) \right\|_E^2 \mathrm{d}u \, \mathrm{d}v = 0,$$

and

$$\lim_{\|\pi\| \to 0} \int_0^T \left\| \left(\mathcal{K}^* \otimes \mathcal{K}^* \left(\psi^\pi - \psi\right)\right)(r,r) \right\|_E \mathrm{d}r = 0.$$

For this paper, the result above, coupled with the fact that $\mathcal{H}_1^d \otimes \mathcal{H}_1^d$ is isomorphic to $\mathcal{H}_2^d \otimes \mathcal{H}_2^d$, shows that the strongly λ-Hölder bi-continuous functions are contained in $\mathcal{H}_1^d \otimes \mathcal{H}_1^d$ for the class of Volterra kernels we are considering. For orientation here, contrast this to Proposition 3.2, which showed a similar inclusion in $\mathcal{H}_1^d$ for the class of λ-Hölder continuous functions.

3.3. *The Malliavin derivative and convergence in the tensor norm*

Here, we will apply the results of the last subsection to the Malliavin derivatives of RDE solutions. When $\mathbf{X} \in \mathcal{C}^{0,p\text{-var}}([0,T]; G^{\lfloor p \rfloor}(\mathbb{R}^d))$ satisfies Condition 1, for all $h \in \mathcal{H}_1^d$, $\Phi(h)$ can be embedded in $\mathcal{C}^{q\text{-var}}([0,T]; \mathbb{R}^d)$ where $\frac{1}{p} + \frac{1}{q} > 1$. Furthermore, the Malliavin derivative of Y satisfying (2) is given by [16]

$$\mathcal{D}_h Y_t = \int_0^t J_t^{\mathbf{X}} \left(J_s^{\mathbf{X}}\right)^{-1} V(Y_s) \, \mathrm{d}\Phi(h)(s) = \int_0^T \mathbb{1}_{[0,t)}(s) J_t^{\mathbf{X}} \left(J_s^{\mathbf{X}}\right)^{-1} V(Y_s) \, \mathrm{d}\Phi(h)(s).$$

Denoting

$$\mathcal{D}_s Y_t = \mathbb{1}_{[0,t)}(s) J_t^{\mathbf{X}} \left(J_s^{\mathbf{X}}\right)^{-1} V(Y_s) \tag{18}$$

with respect to any partition $\pi = \{r_i\}$ of $[0,T]$, we will write

$$\mathcal{D}_s Y_t^\pi = \sum_i \mathcal{D}_s Y_{r_i} \mathbb{1}_{[r_i, r_{i+1})}(t).$$

We will proceed to show that

(i) $\mathcal{D}Y^\pi$ lies in $\mathcal{H}_1^d \otimes \mathcal{H}_1^d$ almost surely, and under suitable regularity assumptions on $\mathcal{D}Y$, we have
(ii) $\|\mathcal{D}Y^\pi - \mathcal{D}Y\|_{\mathcal{H}_1^d \otimes \mathcal{H}_1^d} \to 0$ as $\|\pi\| \to 0$.

Coupled with the results in the previous subsections, this will mean that Y^π converges to Y in $\mathbb{D}^{1,2}(\mathcal{H}_1^d)$, and $\delta^X(Y)$ is then the $L^2(\Omega)$ limit of $\delta^X(Y^\pi)$.

A potential problem with (18) is the discontinuity at the diagonal $\{s = t\}$. The next two propositions show how to handle discontinuities of this form.

Proposition 3.9. *Given a Banach space $(E, \|\cdot\|_E)$, let $\psi : [0,T]^2 \to E$ be of the form*

$$\psi(u,v) = \mathbb{1}_{[0,v)}(u)\tilde{\psi}(u,v),$$

where $\tilde{\psi} : [0,T]^2 \to E$ is strongly λ-Hölder bi-continuous in the norm of E. Assume that K is a Volterra kernel which satisfies Condition 2 for some $\alpha \in [0, \frac{1}{4})$ and let $\mathcal{K}^* \otimes \mathcal{K}^*$ be the operator given in Definition 3.6. Then if $\lambda > \alpha$, $(\mathcal{K}^* \otimes \mathcal{K}^*)\psi$ is in $L^2([0,T]^2; E)$.

Proof. We will investigate the integrability of

$$\mathcal{K}^* \otimes \mathcal{K}^* \psi(u,v) = \psi(u,v)K(T,u)K(T,v) + K(T,v)A^K\big(\psi(\cdot,v)\big)(u) \\ + K(T,u)A^K\big(\psi(u,\cdot)\big)(v) + B^K(\psi)(u,v) \tag{19}$$

in the regions $\{u < v\}$ and $\{v < u\}$ separately (ignoring the diagonal as it has zero Lebesgue measure).

(i) $u < v$:

For the first term on the right of (19) we have

$$\psi(u,v)K(T,u)K(T,v) = \tilde{\psi}(u,v)K(T,u)K(T,v) \in L^2\big([0,T]^2; E\big),$$

and for the second term, we have

$$\begin{aligned} &\big\|K(T,v)A^K\big(\psi(\cdot,v)\big)(u)\big\|_E \\ &= \left\| K(T,v)\left(\int_u^v \big[\tilde{\psi}(r,v) - \tilde{\psi}(u,v)\big]K(\mathrm{d}r,u) - \int_v^T \tilde{\psi}(u,v)K(\mathrm{d}r,u)\right)\right\|_E \\ &\le C\big|K(T,v)\big|\left((v-u)^{\lambda-\alpha} + \left(\frac{1}{(v-u)^{\alpha}} - \frac{1}{(T-u)^{\alpha}}\right)\right) \in L^2\big([0,T]^2\big). \end{aligned}$$

The third term satisfies

$$\begin{aligned} \big\|K(T,u)A^K\big(\psi(u,\cdot)\big)(v)\big\|_E &= \left\|K(T,u)\int_v^T \big[\tilde{\psi}(u,r) - \tilde{\psi}(u,v)\big]K(\mathrm{d}r,v)\right\|_E \\ &\le C\big|K(T,u)\big|(T-v)^{\lambda-\alpha} \in L^2\big([0,T]^2\big). \end{aligned}$$

For the fourth term, given $r_1 \in (v,T]$, we have

$$(u,T] \times (v,T] = \big\{(u,v] \times (v,T]\big\} \sqcup \big\{(v,T] \times (v,r_1]\big\} \sqcup \big\{(v,T] \times (r_1,T]\big\},$$

and thus

$$\begin{aligned} \big\|B^K(\psi)(u,v)\big\|_E = \bigg\| &\int_u^v \left(\int_v^T \tilde{\psi}\begin{pmatrix} u & r_1 \\ v & r_2\end{pmatrix} K(\mathrm{d}r_2,v)\right) K(\mathrm{d}r_1,u) \\ &+ \int_v^T \left(\int_{r_1}^T \big[\tilde{\psi}(r_1,r_2) + \tilde{\psi}(u,v) - \tilde{\psi}(u,r_2)\big]K(\mathrm{d}r_2,v)\right)K(\mathrm{d}r_1,u) \\ &+ \int_v^T \left(\int_v^{r_1} \big[\tilde{\psi}(u,v) - \tilde{\psi}(u,r_2)\big]K(\mathrm{d}r_2,v)\right)K(\mathrm{d}r_1,u)\bigg\|_E. \end{aligned}$$

This expression is bounded above by

$$C\left((v-u)^{\lambda-\alpha}(T-v)^{\lambda-\alpha} + \int_v^T \frac{1}{(r_1-v)^{\alpha}(r_1-u)^{\alpha+1}}\,\mathrm{d}r_1 + \left(\frac{1}{(T-v)^{\alpha}}\right)\left(\frac{1}{(v-u)^{\alpha}} - \frac{1}{(T-u)^{\alpha}}\right)\right).$$

Since

$$\int_v^T \frac{1}{(r_1-v)^{\alpha}(r_1-u)^{\alpha+1}}\,\mathrm{d}r_1 = \int_v^T \frac{1}{(r_1-v)^{\alpha}(r_1-u)^{\alpha+\frac{1}{4}}(r_1-u)^{\frac{3}{4}}}\,\mathrm{d}r_1$$

$$\leq \frac{1}{(v-u)^{\alpha+\frac{1}{4}}} \int_v^T \frac{1}{(r_1-v)^{\alpha+\frac{3}{4}}} \, \mathrm{d}r_1, \tag{20}$$

and $\alpha < \frac{1}{4}$, the fourth term is also in $L^2([0,T]^2;E)$.

(ii) $v < u$:

The first two terms on the right of (19) vanish, and the third term obeys the estimate

$$\begin{aligned}\big\|K(T,u)A^K\big(\psi(u,\cdot)\big)(v)\big\|_E &= \left\|K(T,u)\int_u^T \tilde{\psi}(u,r)K(\mathrm{d}r,v)\right\|_E \quad \big(\psi(u,r)=0 \text{ when } v<r<u\big) \\ &\leq C\big|K(T,u)\big|\left(\frac{1}{(u-v)^\alpha} - \frac{1}{(T-v)^\alpha}\right),\end{aligned}$$

and hence it is in $L^2([0,T]^2;E)$. For the fourth term, note that

$$\psi\begin{pmatrix} u & r_1 \\ v & r_2\end{pmatrix} = 0 \quad \text{when } v < r_2 < u,$$

and thus we have

$$\begin{aligned}\big\|B^K(\psi)(u,v)\big\|_E &\leq \left\|\int_u^T \left(\int_u^{r_2} \big[\tilde{\psi}(r_1,r_2) - \tilde{\psi}(u,r_2)\big]K(\mathrm{d}r_1,u)\right)K(\mathrm{d}r_2,v)\right\|_E \\ &\quad + \left\|\int_u^T \left(\int_{r_2}^T \tilde{\psi}(u,r_2)K(\mathrm{d}r_1,u)\right)K(\mathrm{d}r_2,v)\right\|_E \\ &\leq C\left(\left(\frac{1}{(u-v)^\alpha} - \frac{1}{(T-v)^\alpha}\right) + \int_u^T \frac{1}{(r_2-u)^\alpha(r_2-v)^{\alpha+1}}\,\mathrm{d}r_2\right).\end{aligned}$$

Utilizing (20) again, we see that the fourth term is also in $L^2([0,T]^2;E)$. □

Proposition 3.10. *Let F denote either $\mathbb{R}^e$ or $L^2(\Omega;\mathbb{R}^e)$, and let $\psi : [0,T]^2 \to F$ be a function of the form $\psi(u,v) = \mathbb{1}_{[0,v)}(u)\tilde{\psi}(u,v)$, where $\tilde{\psi}$ is strongly λ-Hölder bi-continuous in the norm of F. Given a partition $\pi = \{r_i\}$ of $[0,T]$, denote*

$$\psi^\pi(s,t) := \sum_j \psi(s,r_j)\mathbb{1}_{[r_j,r_{j+1})}(t). \tag{21}$$

Moreover, let $\mathcal{K}^* \otimes \mathcal{K}^*$ be the operator given in Definition 3.6, where the Volterra kernel K satisfies Condition 2 for some $\alpha \in [0,\frac{1}{4})$. Then if $\lambda > \alpha$, we have

$$\int_{[0,T]^2} \big\|\mathcal{K}^* \otimes \mathcal{K}^*\big(\psi^\pi - \psi\big)(s,t)\big\|_F^2 \,\mathrm{d}s\,\mathrm{d}t \to 0.$$

Proof. We define

$$h(u,v) := \int_0^T \big\langle \mathcal{K}^*\big(\psi(\cdot,u)\big)(s), \mathcal{K}^*\big(\psi(\cdot,v)\big)(s)\big\rangle_F \,\mathrm{d}s,$$

and correspondingly,

$$\begin{aligned}h^\pi(u,v) &:= \int_0^T \big\langle \mathcal{K}^*\big(\psi^\pi(\cdot,u)\big)(s), \mathcal{K}^*\big(\psi^\pi(\cdot,v)\big)(s)\big\rangle_F \,\mathrm{d}s \\ &= \sum_{i,j}\left(\int_0^T \big\langle \mathcal{K}^*\big(\psi^\pi(\cdot,r_i)\big)(s), \mathcal{K}^*\big(\psi^\pi(\cdot,r_j)\big)(s)\big\rangle_F \,\mathrm{d}s\right)\mathbb{1}_{[r_i,r_{i+1})}(u)\mathbb{1}_{[r_j,r_{j+1})}(v) \\ &= \sum_{i,j} h(r_i,r_j)\mathbb{1}_{[r_i,r_{i+1})}(u)\mathbb{1}_{[r_j,r_{j+1})}(v).\end{aligned}$$

Let $\lambda' := \frac{1}{4} \wedge \lambda$. Since $\alpha < \frac{1}{4}$, λ' is greater than α, and note that any strongly λ-Hölder bi-continuous function is also strongly λ'-Hölder bi-continuous. We will begin by first showing that $h(u, v)$ is strongly λ'-Hölder bi-continuous.

For all $u, v, u_1, u_2, v_1, v_2 \in [0, T]$, we have

$$\big|h(u_1, v) - h(u_2, v)\big| \le \left(\int_0^T \big\|\mathcal{K}^*\big(\psi(\cdot, u_1) - \psi(\cdot, u_2)\big)(s)\big\|_F^2 \, \mathrm{d}s\right)^{\frac{1}{2}} \left(\int_0^T \big\|\mathcal{K}^*\big(\psi(\cdot, v)\big)(s)\big\|_F^2 \, \mathrm{d}s\right)^{\frac{1}{2}},$$

$$\big|h(u, v_1) - h(u, v_2)\big| \le \left(\int_0^T \big\|\mathcal{K}^*\big(\psi(\cdot, v_1) - \psi(\cdot, v_2)\big)(s)\big\|_F^2 \, \mathrm{d}s\right)^{\frac{1}{2}} \left(\int_0^T \big\|\mathcal{K}^*\big(\psi(\cdot, u)\big)(s)\big\|_F^2 \, \mathrm{d}s\right)^{\frac{1}{2}},$$

and $\big|h\big(\begin{smallmatrix} u_1 & u_2 \\ v_1 & v_2 \end{smallmatrix}\big)\big|$ is bounded above by

$$\left(\int_0^T \big\|\mathcal{K}^*\big(\psi(\cdot, u_1) - \psi(\cdot, u_2)\big)(s)\big\|_F^2 \, \mathrm{d}s\right)^{\frac{1}{2}} \left(\int_0^T \big\|\mathcal{K}^*\big(\psi(\cdot, v_1) - \psi(\cdot, v_2)\big)(s)\big\|_F^2 \, \mathrm{d}s\right)^{\frac{1}{2}}.$$

Note that for $p \ge 1$, using (13) and fixing $w \in [0, T]$, we have

$$\begin{aligned} \big\|\mathcal{K}^*\psi(\cdot, w)(s)\big\|_F^p &= \left\|\tilde{\psi}(s, w)K(w, s) + \int_s^w \big[\tilde{\psi}(r, w) - \tilde{\psi}(s, w)\big]K(\mathrm{d}r, s)\right\|_F^p \\ &\le C2^{p-1}\left(\frac{1}{s^{p\alpha}(w-s)^{p\alpha}} + (w-s)^{p(\lambda'-\alpha)}\right). \end{aligned} \tag{22}$$

Since $\alpha < \frac{1}{4}$, $\int_0^T \|\mathcal{K}^*\psi(\cdot, w)(s)\|_F^p \, \mathrm{d}s$ is finite as long as $p \le 4$.

Now, all we have to do is show that

$$\int_0^T \big\|\mathcal{K}^*\big(\psi(\cdot, w_2) - \psi(\cdot, w_1)\big)(s)\big\|_F^2 \, \mathrm{d}s \le C|w_2 - w_1|^{2\lambda'}, \tag{23}$$

for all $w_1, w_2 \in [0, T]$, where without loss of generality, we let $w_1 < w_2$. Observe that

$$\begin{aligned} &\int_0^T \big\|\mathcal{K}^*\big(\psi(\cdot, w_2) - \psi(\cdot, w_1)\big)(s)\big\|_F^2 \, \mathrm{d}s \\ &\quad = \int_0^{w_1} \big\|\mathcal{K}^*\big(\psi(\cdot, w_2) - \psi(\cdot, w_1)\big)(s)\big\|_F^2 \, \mathrm{d}s + \int_{w_1}^{w_2} \big\|\mathcal{K}^*\big(\psi(\cdot, w_2) - \psi(\cdot, w_1)\big)(s)\big\|_F^2 \, \mathrm{d}s \end{aligned} \tag{24}$$

as the integrand vanishes when $s \ge w_2$. For the first term above, for $s \in [0, w_1)$, we have (using (13))

$$\begin{aligned} \mathcal{K}^*\big(\psi(\cdot, w_2) - \psi(\cdot, w_1)\big)(s) &= \big(\psi(s, w_2) - \psi(s, w_1)\big)K(w_2, s) \\ &\quad + \int_s^{w_2} \big[\psi(r, w_2) - \psi(s, w_2) - \psi(r, w_1) + \psi(s, w_1)\big]K(\mathrm{d}r, s) \\ &= \big(\tilde{\psi}(s, w_2) - \tilde{\psi}(s, w_1)\big)K(w_2, s) + \int_s^{w_1} \tilde{\psi}\begin{pmatrix} s & r \\ w_1 & w_2 \end{pmatrix} K(\mathrm{d}r, s) \\ &\quad + \int_{w_1}^{w_2} \big[\tilde{\psi}(r, w_2) - \tilde{\psi}(s, w_2) + \tilde{\psi}(s, w_1)\big]K(\mathrm{d}r, s). \end{aligned} \tag{25}$$

Since $\tilde{\psi}$ is strongly λ'-Hölder bi-continuous, we have

$$\big\|\big(\tilde{\psi}(s, w_2) - \tilde{\psi}(s, w_1)\big)K(w_2, s)\big\|_F \le C|w_2 - w_1|^{\lambda'} s^{-\alpha}(w_2 - s)^{-\alpha}, \tag{26}$$

and

$$\left\|\int_s^{w_1} \tilde{\psi}\begin{pmatrix} s & r \\ w_1 & w_2 \end{pmatrix} K(\mathrm{d}r, s)\right\|_F \le C|w_2 - w_1|^{\lambda'}(w_1 - s)^{\lambda'-\alpha}. \tag{27}$$

For the last integral in (25), we let q_1 denote $\frac{1}{1-\lambda'}$ and use Hölder's inequality to derive

$$
\begin{aligned}
\left\| \int_{w_1}^{w_2} \left[\tilde{\psi}(r, w_2) - \tilde{\psi}(s, w_2) + \tilde{\psi}(s, w_1)\right] K(\mathrm{d}r, s) \right\|_F &\le C|w_2 - w_1|^{\lambda'} \left(\int_{w_1}^{w_2} \left| \frac{\partial K(r,s)}{\partial r} \right|^{q_1} \mathrm{d}r \right)^{\frac{1}{q_1}} \\
&\le C|w_2 - w_1|^{\lambda'} \left(\int_{w_1}^{w_2} \frac{1}{(r-s)^{q_1(\alpha+1)}} \mathrm{d}r \right)^{\frac{1}{q_1}} \\
&\le C|w_2 - w_1|^{\lambda'} (w_1 - s)^{-(\alpha+\lambda')}. \qquad (28)
\end{aligned}
$$

Putting estimates (26), (27) and (28) together, when $s < w_1$ we have

$$
\left\| \mathcal{K}^*\big(\psi(\cdot, w_2) - \psi(\cdot, w_1)\big)(s) \right\|_F \le C|w_2 - w_1|^{\lambda'} f(s), \qquad (29)
$$

for some $f(s) \in L^2([0,T])$ since $\lambda' > \alpha$ and $2(\alpha + \lambda') < 1$. This gives

$$
\int_0^{w_1} \left\| \mathcal{K}^*\big(\psi(\cdot, w_2) - \psi(\cdot, w_1)\big)(s) \right\|_F^2 \mathrm{d}s \le C|w_2 - w_1|^{2\lambda'}.
$$

Returning to the second term in (24), we let q_2 denote $\frac{1}{1-2\lambda'}$ and use Hölder's inequality again to obtain

$$
\int_{w_1}^{w_2} \left\| \mathcal{K}^*\big(\psi(\cdot, w_1) - \psi(\cdot, w_2)\big)(s) \right\|_F^2 \mathrm{d}s \le |w_2 - w_1|^{2\lambda'} \left(\int_0^T \left\| \mathcal{K}^*\big(\psi(\cdot, w_1) - \psi(\cdot, w_2)\big)(s) \right\|_F^{2q_2} \mathrm{d}s \right)^{\frac{1}{q_2}}.
$$

Since $\lambda' < \frac{1}{4}$, we have $2q_2 \le 4$ and this gives $(\int_0^T \|\mathcal{K}^*(\psi(\cdot, w_1) - \psi(\cdot, w_2))(s)\|_F^{2q_2} \mathrm{d}s)^{\frac{1}{q_2}} < \infty$ from (22). Now that we have shown that h is strongly λ'-Hölder bi-continuous, we will show that

$$
\int_{[0,T]^2} \left\| \mathcal{K}^* \otimes \mathcal{K}^*\big(\psi^\pi - \psi\big)(s,t) \right\|_F^2 \mathrm{d}s\,\mathrm{d}t = \int_0^T \big(\mathcal{K}^* \otimes \mathcal{K}^*\big(h^\pi - h\big)\big)(t,t)\,\mathrm{d}t,
$$

and then invoke Theorem 3.8 to complete the proof.

Let $g(s,t)$ denote $\mathcal{K}^*(\psi(\cdot, t))(s)$, and note that $g(s,t) = 0$ when $s \ge t$. We first compute

$$
\begin{aligned}
\mathcal{K}^* \otimes \mathcal{K}^* h(t,t) &= h(t,t)K(T,t)^2 + K(T,t)A^K\big(h(\cdot,t)\big)(t) + K(T,t)A^K\big(h(t,\cdot)\big)(t) + B^K(h)(t,t) \\
&= \int_0^T \langle g(s,t), g(s,t)\rangle_F K(T,t)^2 \,\mathrm{d}s \\
&\quad + 2K(T,t) \int_t^T \left(\int_0^T \langle g(s,r) - g(s,t), g(s,t)\rangle_F \,\mathrm{d}s \right) K(\mathrm{d}r, t) \\
&\quad + \int_t^T \int_t^T \left(\int_0^T \langle g(s,r_1) - g(s,t), g(s,r_2) - g(s,t)\rangle_F \,\mathrm{d}s \right) K(\mathrm{d}r_1, t) K(\mathrm{d}r_2, t). \qquad (30)
\end{aligned}
$$

The second term on the right vanishes when $s \ge t$, and when $s < t$, using (22) and (29) gives us

$$
\left| \langle g(s,r) - g(s,t), g(s,t)\rangle_F \right| \left| \frac{\partial K(r,t)}{\partial r} \right| \le C|r - t|^{\lambda' - \alpha - 1} \tilde{f}(s)
$$

for some $\tilde{f}(s) \in L^1([0,T])$, and thus we can swap the integral with respect to s outside the integral with respect to r. Similarly, the third term on the right of (30) is bounded by

$$
C\left(\int_s^T \frac{1}{(r-t)^{\alpha+1}} \mathrm{d}r \right)^2
$$

when $s > t$ since the integrand vanishes when $r_1 \le s$ or $r_2 \le s$. Furthermore, when $s < t$, its integrand is bounded by

$$
C|r_1 - t|^{\lambda' - \alpha - 1} |r_2 - t|^{\lambda' - \alpha - 1} f^2(s).
$$

Hence, we can also pull out the integral with respect to s, and we get

$$\mathcal{K}^* \otimes \mathcal{K}^* h(t,t) = \int_0^T \mathcal{K}^* \otimes \mathcal{K}^* \big(\langle g(s,\cdot), g(s,\cdot) \rangle_F \big)(t,t)\, \mathrm{d}s.$$

Observe that

$$\begin{aligned}\mathcal{K}^* \otimes \mathcal{K}^* \big(\langle g(s,\cdot), g(s,\cdot) \rangle_F \big)(t,t) &= \big\langle \mathcal{K}^* \big(g(s,\cdot) \big)(t), \mathcal{K}^* \big(g(s,\cdot) \big)(t) \big\rangle_F \\ &= \big\| \mathcal{K}^* \big(g(s,\cdot) \big)(t) \big\|_F^2,\end{aligned}$$

where here we use (16), and Fubini's theorem in the case when $F = L^2(\Omega; \mathbb{R}^e)$.

Fixing s, note that for all $t > s$, $g(s,\cdot)$ is λ'-Hölder continuous on $[t,T]$ (with the Hölder norm depending on t) from (29). Thus, $\mathcal{K}^*(g(s,\cdot))(t)$ is well defined for all $t > s$, vanishes when $t < s$, and we can apply Lemma 3.2 in [23] to obtain

$$\mathcal{K}^* \big(g(s,\cdot) \big)(t) = \mathcal{K}^* \otimes \mathcal{K}^* \psi(s,t)$$

for all $s \neq t$. This concludes the proof. □

We will use the previous proposition to show that $\mathcal{H}_1^d \otimes \mathcal{H}_1^d$ contains functions $\psi : [0,T]^2 \to \mathbb{R}^d \otimes \mathbb{R}^d$ of the form $\psi(u,v) = \mathbb{1}_{[0,v)}(u) \tilde{\psi}(u,v)$ whenever $\tilde{\psi}$ is strongly λ-Hölder bi-continuous.

Proposition 3.11. *Let $\psi : [0,T]^2 \to \mathbb{R}^d$ be of the form $\psi(u,v) = \mathbb{1}_{[0,v)}(u) \tilde{\psi}(u,v)$, where $\tilde{\psi}$ is strongly λ-Hölder bi-continuous, and let $\mathcal{K}^* \otimes \mathcal{K}^*$ be defined as in Definition 3.6, where the Volterra kernel K satisfies Condition 2 for some $\alpha \in [0, \frac{1}{4})$.*

Then if $\lambda > \alpha$, ψ is an element of $\mathcal{H}_1^d \otimes \mathcal{H}_1^d$, with norm given by

$$\|\psi\|_{\mathcal{H}_1^d \otimes \mathcal{H}_1^d} = \int_{[0,T]^2} \big| \mathcal{K}^* \otimes \mathcal{K}^* \psi(s,t) \big|^2_{\mathbb{R}^d \otimes \mathbb{R}^d} \,\mathrm{d}s\, \mathrm{d}t, \tag{31}$$

and with ψ^π defined as in (21), we have

$$\big\| \psi^\pi - \psi \big\|_{\mathcal{H}_1^d \otimes \mathcal{H}_1^d} \to 0 \tag{32}$$

as $\|\pi\| \to 0$.

Proof. Given a d-by-d matrix function $A(s)$, let $a_j^{(i)}(s)$ denote the i,jth entry of $A(s)$. Using the canonical identification

$$A(s)\mathbb{1}_{[a,b)}(t) \simeq \sum_{j=1}^d \sum_{i=1}^d a_j^{(i)}(s) e_i \otimes \mathbb{1}_{[a,b)}(t) e_j, \quad a,b \in [0,T], \tag{33}$$

it is clear that ψ^π is a member of $\Lambda_\alpha^d \otimes \mathcal{H}_1^d$, and thus lies in $\mathcal{H}_1^d \otimes \mathcal{H}_1^d$ by Proposition 3.5. Furthermore, $\|\psi^\pi\|^2_{\mathcal{H}_1^d \otimes \mathcal{H}_1^d}$ is equal to, using the notation $\mathbb{1}_{\Delta_i} = \mathbb{1}_{[r_i, r_{i+1})}$,

$$\begin{aligned}&\sum_{k,l} \int_0^T \sum_{j=1}^d \big\langle \mathcal{K}^* \big(\psi_j(\cdot, r_k) \big)(s), \mathcal{K}^* \big(\psi_j(\cdot, r_l) \big)(s) \big\rangle_{\mathbb{R}^d} \,\mathrm{d}s \int_0^T \mathcal{K}^*(\mathbb{1}_{\Delta_k})(t) \mathcal{K}^*(\mathbb{1}_{\Delta_l})(t)\, \mathrm{d}t \\ &\quad = \sum_{k,l} \int_{[0,T]^2} \big\langle \mathcal{K}^* \otimes \mathcal{K}^* \big(\psi(\cdot, r_k) \mathbb{1}_{\Delta_k}(\cdot) \big)(s,t), \mathcal{K}^* \otimes \mathcal{K}^* \big(\psi(\cdot, r_l) \mathbb{1}_{\Delta_l}(\cdot) \big)(s,t) \big\rangle_{\mathbb{R}^d \otimes \mathbb{R}^d} \,\mathrm{d}s\, \mathrm{d}t, \\ &\quad = \int_{[0,T]^2} \big| \mathcal{K}^* \otimes \mathcal{K}^* \psi^\pi(s,t) \big|^2_{\mathbb{R}^d \otimes \mathbb{R}^d} \,\mathrm{d}s\, \mathrm{d}t,\end{aligned}$$

which we know is Cauchy as $\|\pi\| \to 0$ by Proposition 3.10. We now take any sequence of partitions $\pi(n)$ with vanishing mesh and identify ψ with the limit of $\psi^{\pi(n)}$ in $\mathcal{H}_1^d \otimes \mathcal{H}_1^d$. Invoking Proposition 3.10 again then gives us (31) and (32). □

3.4. *The Itô–Skorohod isometry revisited*

We now give another formulation for the Itô-Skorohod isometry for Volterra processes (cf. Theorem A.6 in [12] for the specific case of fractional Brownian motion).

Theorem 3.12. *Let X be a Volterra process which satisfies Condition 1 for some $\rho \in [1,2)$, and assume that its kernel satisfies Condition 2 for $\alpha = \frac{1}{2} - \frac{1}{2\rho}$. Given $\lambda > \alpha$, let Y be a process which satisfies, almost surely,*

(i) $Y \in \mathcal{C}_{pw}^{\lambda\text{-Höl}}([0,T];\mathbb{R}^d)$,
(ii) $\mathcal{D}Y : [0,T]^2 \to \mathbb{R}^d \otimes \mathbb{R}^d$ *is a function of the form* $\mathbb{1}_{[0,t)}(s)g(s,t)$, *where g is strongly λ-Hölder bi-continuous (recall that $\mathcal{D}Y$ is the Malliavin derivative).*

Then $\lim_{\|\pi\|\to 0} Y^\pi = Y$ *in* $\mathbb{D}^{1,2}(\mathcal{H}_1^d)$ *if and only if*

$$\lim_{\|\pi\|\to 0} \mathbb{E}\left[\int_0^T \left|\mathcal{K}^*\left(Y^\pi - Y\right)(t)\right|_{\mathbb{R}^d}^2 \,\mathrm{d}t\right] = 0,$$

and

$$\lim_{\|\pi\|\to 0} \mathbb{E}\left[\int_{[0,T]^2} \left|\mathcal{K}^* \otimes \mathcal{K}^*\left(\mathcal{D}Y^\pi - \mathcal{D}Y\right)(s,t)\right|_{\mathbb{R}^d\otimes\mathbb{R}^d}^2 \,\mathrm{d}s\,\mathrm{d}t\right] = 0,$$

in which case $\lim_{\|\pi\|\to 0} \mathbb{E}[\delta^X(Y^\pi - Y)^2] = 0$ *and* $\mathbb{E}[\delta^X(Y)^2]$ *is equal to*

$$\mathbb{E}\left[\int_0^T \left|\mathcal{K}^* Y(t)\right|_{\mathbb{R}^d}^2 \,\mathrm{d}t\right] + \mathbb{E}\left[\int_{[0,T]^2} \mathrm{tr}\left(\mathcal{K}^* \otimes \mathcal{K}^* \mathcal{D}Y(s,t) \mathcal{K}^* \otimes \mathcal{K}^* \mathcal{D}Y(t,s)\right) \mathrm{d}s\,\mathrm{d}t\right].$$

Proof. Itô-Skorohod isometry (see [27]) gives us

$$\begin{aligned}
\mathbb{E}\left[\delta^X(Y)^2\right] &= \mathbb{E}\left[\|Y\|_{\mathcal{H}_1^d}^2\right] + \mathbb{E}\left[\mathrm{trace}(\mathcal{D}Y \circ \mathcal{D}Y)\right] \\
&= \lim_{\|\pi\|\to 0} \mathbb{E}\left[\left\|Y^\pi\right\|_{\mathcal{H}_1^d}^2\right] + \lim_{\|\pi\|\to 0} \mathbb{E}\left[\mathrm{trace}\left(\mathcal{D}Y^\pi \circ \mathcal{D}Y^\pi\right)\right].
\end{aligned}$$

The first term is equal to $\lim_{\|\pi\|\to 0} \mathbb{E}[\int_0^T |\mathcal{K}^* Y^\pi(t)|_{\mathbb{R}^d}^2 \,\mathrm{d}t]$, with limit $\mathbb{E}[\int_0^T |\mathcal{K}^* Y(t)|_{\mathbb{R}^d}^2 \,\mathrm{d}t]$.

For the second term, it can be shown that the trace $\mathbb{E}[\mathrm{trace}(\mathcal{D}Y^\pi \circ \mathcal{D}Y^\pi)]$ is equal to (see Theorem 4.8 of [7])

$$\mathbb{E}\left[\sum_{i,j} \sum_{k,l=1}^d \left\langle \mathcal{D}_\cdot^{(k)} Y_{t_j}^{(l)}, \mathbb{1}_{\Delta_i}(\cdot)\right\rangle_{\mathcal{H}_1} \left\langle \mathcal{D}_\cdot^{(l)} Y_{t_i}^{(k)}, \mathbb{1}_{\Delta_j}(\cdot)\right\rangle_{\mathcal{H}_1}\right],$$

which is the same as

$$\mathbb{E}\left[\sum_{i,j} \sum_{k,l=1}^d \int_0^T \mathcal{K}^*\left(\mathcal{D}_\cdot^{(k)} Y_{t_j}^{(l)}\right)(s) K(\Delta_i, s)\,\mathrm{d}s \int_0^T \mathcal{K}^*\left(\mathcal{D}_\cdot^{(l)} Y_{t_i}^{(k)}\right)(t) K(\Delta_j, t)\,\mathrm{d}t\right]$$

for a Volterra process. Using Lemma 3.2 in [23], this last expression equals

$$\mathbb{E}\left[\int_{[0,T]^2} \mathrm{tr}\left(\mathcal{K}^* \otimes \mathcal{K}^* \mathcal{D}Y^\pi(s,t) \mathcal{K}^* \otimes \mathcal{K}^* \mathcal{D}Y^\pi(t,s)\right) \mathrm{d}s\,\mathrm{d}t\right],$$

and it converges as $\|\pi\| \to 0$ to

$$\mathbb{E}\left[\int_{[0,T]^2} \mathrm{tr}\left(\mathcal{K}^* \otimes \mathcal{K}^* \mathcal{D}Y(s,t) \mathcal{K}^* \otimes \mathcal{K}^* \mathcal{D}Y(t,s)\right) \mathrm{d}s\,\mathrm{d}t\right].$$

□

In the case of Brownian motion both $\mathcal{K}^*$ and $\mathcal{K}^* \otimes \mathcal{K}^*$ are identity operators and Theorem 3.12 recovers the usual Itô-Skorohod isometry:

$$\mathbb{E}\left[\delta^X(Y)^2\right] = \mathbb{E}\left[\int_0^T |Y_t|^2 \,\mathrm{d}t\right] + \mathbb{E}\left[\int_{[0,T]^2} \mathrm{tr}(\mathcal{D}_t Y_s \mathcal{D}_s Y_t)\,\mathrm{d}s\,\mathrm{d}t\right].$$

3.5. *Approximation of the Skorohod integral*

We will now put together the results of the previous subsections to show that the Skorohod integral of the discrete approximations to the solution of an RDE converge. Before we proceed, let $Y \in \mathcal{C}^{p\text{-var}}([0,T];\mathcal{L}(\mathbb{R}^d;\mathbb{R}^m))$ denote the path-level solution to

$$\mathrm{d}Y_t = V(Y_t) \circ \mathrm{d}\mathbf{X}_t, \qquad Y_0 = y_0,$$

where $V \in \mathcal{C}_b^{\lfloor p \rfloor+1}(\mathbb{R}^m \otimes \mathbb{R}^d; \mathbb{R}^m \otimes \mathbb{R}^d \otimes \mathbb{R}^d)$.

Given a Hilbert space H, we will denote an element of y of $\mathbb{R}^m \otimes H$ as

$$y = \sum_{j=1}^m e_j \otimes [y]_j, \tag{34}$$

where $[y]_j \in H$ for $j = 1, \dots, m$. (Note that there may be several ways to perform the decomposition.)

Now fix $0 \le s < t \le T$. Since $V(Y_s) \in \mathbb{R}^{md} \otimes \mathbb{R}^d \simeq \mathbb{R}^m \otimes \mathbb{R}^d \otimes \mathbb{R}^d$, we can decompose $V(Y_s)$ as

$$V(Y_s) = \sum_{j=1}^m e_j \otimes \big[V(Y_s)\big]_j,$$

where

$$\big[V(Y_s)\big]_j := \sum_{i,k=1}^d V_k^{(d(j-1)+i)}(Y_s) e_i \otimes e_k.$$

If we canonically identify $\mathbb{R}^{md} \otimes \mathbb{R}^d$ with the space of md-by-d matrices, then $[V(Y_s)]_j$ simply denotes the d-by-d sub-matrix of $V(Y_s)$ which starts at the $(d(j-1)+1)$th row and ends at the djth row. Contrast this with $V_j(Y_s)$, which denotes the jth column of $V(Y_s)$.

Similarly, we can do the same with $Y_s \in \mathbb{R}^{md} \simeq \mathbb{R}^m \otimes \mathbb{R}^d$ and $J^{\mathbf{X}}_{t\leftarrow s} V(Y_s) \in \mathbb{R}^{md} \otimes \mathbb{R}^d \simeq \mathbb{R}^m \otimes \mathbb{R}^d \otimes \mathbb{R}^d$, $J^{\mathbf{X}}_{t\leftarrow s} := J^{\mathbf{X}}_t (J^{\mathbf{X}}_s)^{-1}$.

Proposition 3.13. *Let $\mathbf{X} \in \mathcal{C}^{0,p\text{-var}}([0,T]; G^{\lfloor p \rfloor}(\mathbb{R}^d))$, $1 \le p < 4$, be a Volterra rough path which satisfies Condition 1, and assume that its kernel satisfies Condition 2 with $\alpha < \frac{1}{p}$. Let $Y \in \mathcal{C}^{p\text{-var}}([0,T];\mathcal{L}(\mathbb{R}^d;\mathbb{R}^m))$ denote the path-level solution to*

$$\mathrm{d}Y_t = V(Y_t) \circ \mathrm{d}\mathbf{X}_t, Y_0 = y_0,$$

where $V \in \mathcal{C}_b^{\lfloor p \rfloor+1}(\mathbb{R}^m \otimes \mathbb{R}^d; \mathbb{R}^m \otimes \mathbb{R}^d \otimes \mathbb{R}^d)$. Then $Y \in \mathbb{D}^{1,2}(\mathbb{R}^m \otimes \mathcal{H}_1^d)$ and

$$\int_0^T Y_r \,\mathrm{d}X_r = \lim_{\|\pi = \{r_i\}\| \to 0} \sum_i \Bigg[Y_{r_i}(X_{r_i,r_{i+1}}) - \sum_{j=1}^m \bigg(\int_0^{r_i} \mathrm{tr}\big[J^{\mathbf{X}}_{r_i \leftarrow s} V(Y_s)\big]_j R(\Delta_i, \mathrm{d}s) \bigg) e_j \Bigg],$$

where the limit is taken in $L^2(\Omega)$.

Proof. We first use integration-by-parts to obtain

$$\big\langle \delta^X\big(Y^\pi\big), e_j \big\rangle_{\mathbb{R}^m} = \sum_i \bigg[\big\langle [Y_{r_i}]_j, X_{r_i,r_{i+1}} \big\rangle_{\mathbb{R}^d} - \int_0^{r_i} \mathrm{tr}\big[J^{\mathbf{X}}_{r_i \leftarrow s} V(Y_s)\big]_j R(\Delta_i, \mathrm{d}s) \bigg],$$

for all $j = 1, \dots, m$. Next, we invoke Theorem 3.12, which requires us to prove that

$$\mathbb{E}\bigg[\int_0^T \big| \mathcal{K}^*\big(Y^\pi - Y\big)(t) \big|^2_{\mathbb{R}^m \otimes \mathbb{R}^d} \,\mathrm{d}t \bigg] \to 0, \tag{35}$$

and

$$\mathbb{E}\bigg[\int_{[0,T]^2}\big|\mathcal{K}^*\otimes\mathcal{K}^*\big(\mathcal{D}_sY_t^\pi-\mathcal{D}_sY_t\big)(s,t)\big|^2_{\mathbb{R}^m\otimes\mathbb{R}^d\otimes\mathbb{R}^d}\,\mathrm{d}s\,\mathrm{d}t\bigg]\to 0. \tag{36}$$

We will show that Y is $\frac{1}{p}$-Hölder continuous in $L^2(\Omega;\mathbb{R}^m\otimes\mathbb{R}^d)$, and then invoke Proposition 3.2 to obtain (35). We have (see equation (10.15) in [16])

$$|Y_{s,t}|\le C\big(\|\mathbf{X}\|_{p\text{-var};[s,t]}\vee\|\mathbf{X}\|^p_{p\text{-var};[s,t]}\big)\le C\|\mathbf{X}\|_{\frac{1}{p}\text{-Höl};[0,T]}\big(T^{1-\frac{1}{p}}\vee 1\big)(t-s)^{\frac{1}{p}} \tag{37}$$

almost surely, and thus

$$\sqrt{\mathbb{E}\big[|Y_{s,t}|^2\big]}\le C_1|t-s|^{\frac{1}{p}}$$

since $\|\mathbf{X}\|_{\frac{1}{p}\text{-Höl};[0,T]}$ has moments of all orders (corollary 66 in [15]).

To show (36), we will apply Proposition 3.10 with $\psi(s,t)=\mathcal{D}_sY_t=\mathbb{1}_{[0,t)}(s)J_t^{\mathbf{X}}(J_s^{\mathbf{X}})^{-1}V(Y_s)$. To do so, we have to show that $\tilde{\psi}(s,t):=J_t^{\mathbf{X}}(J_s^{\mathbf{X}})^{-1}V(Y_s)$ is strongly $\frac{1}{p}$-Hölder bi-continuous in $L^2(\Omega;\mathbb{R}^{md}\otimes\mathbb{R}^d)$. By Lemma 3.6 in [23], this is equivalent to showing that $J_\cdot^{\mathbf{X}}$ and $(J_\cdot^{\mathbf{X}})^{-1}Y_\cdot$ are both $\frac{1}{p}$-Hölder continuous.

Using equation (4.10) in [8], we have

$$\big|J^{\mathbf{X}}_{s,t}\big|\le C_1\|\mathbf{X}\|_{p\text{-var};[s,t]}\exp\big(C_2N^{\mathbf{X}}_{1;[s,t]}\big)\le C_1(t-s)^{\frac{1}{p}}\|\mathbf{X}\|_{\frac{1}{p}\text{-Höl};[0,T]}\exp\big(C_2N^{\mathbf{X}}_{1;[0,T]}\big),$$

where $N_1^{\mathbf{x}}$ is defined in (44) of [7], and essentially gives a count of the number of times the p-variation of $\mathbf{x}$ increases by 1 in $[0,T]$. This yields $\frac{1}{p}$-Hölder continuity for $J_\cdot^{\mathbf{X}}$ as the expression to the right of $(t-s)^{\frac{1}{p}}$ is in $L^q(\Omega)$ for all $q>0$ (Theorem 6.3 in [8]). The same is true for $(J_\cdot^{\mathbf{X}})^{-1}$ since the inverse obeys the same bound.

Finally, $(J_\cdot^{\mathbf{X}})^{-1}V(Y_\cdot)$ is also $\frac{1}{p}$-Hölder continuous, since V is $\mathcal{C}^1$ smooth and both Y and $(J_\cdot^{\mathbf{X}})^{-1}$ are $\frac{1}{p}$-Hölder continuous. □

4. Augmenting the Skorohod integral with higher-level terms

The main purpose of this section is to show that the usual Riemann-sum approximation to the Skorohod integral can be augmented with (suitably corrected) second-level and third-level rough path terms which vanish in $L^2(\Omega)$ as the mesh of the partition goes to zero.

Before we do so, we will consider the theory of controlled rough paths [17,18] in the case $3\le p<4$, and give bounds on the higher-directional derivatives of a controlled rough path satisfying an RDE.

4.1. *Estimates for controlled rough paths of lower regularity*

To construct the rough integral of controlled rough paths for $3\le p<4$, we need the following definition. In what follows, let $\mathcal{U}$, $\mathcal{V}$ denote finite-dimensional vector spaces.

Definition 4.1. Let $\mathbf{x}=(1,x,\mathbf{x}^2,\mathbf{x}^3)\in\mathcal{C}^{p\text{-var}}([0,T];G^3(\mathbb{R}^d))$, where $3\le p<4$, and let q be such that $\frac{1}{p}+\frac{1}{q}>1$. Let (ϕ,ϕ',ϕ'') be such that $\phi\in\mathcal{C}^{p\text{-var}}([0,T];\mathcal{U})$, $\phi'\in\mathcal{C}^{p\text{-var}}([0,T];\mathcal{L}(\mathbb{R}^d;\mathcal{U}))$ and $\phi''\in\mathcal{C}^{p\text{-var}}([0,T];\mathcal{L}(\mathbb{R}^d\otimes\mathbb{R}^d;\mathcal{U}))$.

Then we say that (ϕ,ϕ',ϕ'') (or ϕ) is controlled by $\mathbf{x}$ if for all $s,t\in[0,T]$ we have

$$\phi_{s,t}=\phi'_sx_{s,t}+\phi''_s\mathbf{x}^2_{s,t}+R^\phi_{s,t},\qquad \phi'_{s,t}=\phi''_sx_{s,t}+R^{\phi'}_{s,t}, \tag{38}$$

where the remainder terms satisfy

$$R^\phi\in\mathcal{C}^{q\text{-var}}\big([0,T];\mathcal{U}\big),\qquad R^{\phi'}\in\mathcal{C}^{\frac{p}{2}\text{-var}}\big([0,T];\mathcal{U}\big).$$

Thus, ϕ is controlled by $\mathbf{x}$ if $\|\phi\|_{p,q\text{-cvar}}<\infty$, where the controlled variation norm is defined as

$$\|\phi\|_{p,q\text{-cvar}}:=\|\phi\|_{\mathcal{V}^p;[0,T]}+\big\|\phi'\big\|_{\mathcal{V}^p;[0,T]}+\big\|\phi''\big\|_{\mathcal{V}^p;[0,T]}+\big\|R^\phi\big\|_{q\text{-var};[0,T]}+\big\|R^{\phi'}\big\|_{\frac{p}{2}\text{-var};[0,T]}.$$

Before we continue, note that $3 \le p < 4$ implies that $\frac{p}{3} < \frac{p}{p-1} \le \frac{p}{2}$. Since $\|R^\phi\|_{q\text{-var}} \ge \|R^\phi\|_{q'\text{-var}}$ for $q' \ge q$, we can replace q with q' where

$$\frac{p}{3} \le q' < \frac{p}{p-1} \le \frac{p}{2} \tag{39}$$

in the subsequent analysis when working with p in the interval $[3,4)$ without affecting the results.

The following theorem and the next two propositions are the lower-regularity analogues of Theorem 2.20, Proposition 2.22 and Proposition 2.21 respectively from [7]. The proofs are routine and can be found in Section 5.1 of [22].

Theorem 4.2. *Let* $\mathbf{x} = (1, x, \mathbf{x}^2, \mathbf{x}^3) \in \mathcal{C}^{p\text{-}var}([0,T]; G^3(\mathbb{R}^d))$, *where* $3 \le p < 4$, *and let* q *be such that* $\frac{1}{p} + \frac{1}{q} > 1$. *Let* (ϕ, ϕ', ϕ'') *satisfy*

$$\begin{aligned}
&\phi \in \mathcal{C}^{p\text{-}var}\big([0,T]; \mathcal{L}\big(\mathbb{R}^d; \mathbb{R}^e\big)\big),\\
&\phi' \in \mathcal{C}^{p\text{-}var}\big([0,T]; \mathcal{L}\big(\mathbb{R}^d; \mathcal{L}\big(\mathbb{R}^d; \mathbb{R}^e\big)\big)\big) \quad \text{and}\\
&\phi'' \in \mathcal{C}^{p\text{-}var}\big([0,T]; \mathcal{L}\big(\mathbb{R}^d \otimes \mathbb{R}^d; \mathcal{L}\big(\mathbb{R}^d; \mathbb{R}^e\big)\big)\big).
\end{aligned}$$

If (ϕ, ϕ', ϕ'') *is controlled by* $\mathbf{x}$ *with remainder terms* R^ϕ *and* $R^{\phi'}$ *of bounded* q*-variation and* $\frac{p}{2}$*-variation respectively, we can define the rough integral*

$$\int_0^t \phi_r \circ \mathrm{d}\mathbf{x}_r := \lim_{\|\pi\|\to 0, \pi=\{0=r_0<\cdots<r_n=t\}} \sum_{i=0}^{n-1} \big(\phi_{r_i} x_{r_i,r_{i+1}} + \phi'_{r_i} \mathbf{x}^2_{r_i,r_{i+1}} + \phi''_{r_i} \mathbf{x}^3_{r_i,r_{i+1}}\big), \tag{40}$$

where we have made use of the canonical identification $\mathcal{L}(\mathbb{R}^d; \mathcal{L}(\mathbb{R}^d; \mathbb{R}^e)) \simeq \mathcal{L}(\mathbb{R}^d \otimes \mathbb{R}^d; \mathbb{R}^e)$ *and* $\mathcal{L}(\mathbb{R}^d \otimes \mathbb{R}^d; \mathcal{L}(\mathbb{R}^d; \mathbb{R}^e)) \simeq \mathcal{L}(\mathbb{R}^d \otimes \mathbb{R}^d \otimes \mathbb{R}^d; \mathbb{R}^e)$. *Furthermore, if* $q \ge \frac{p}{3}$, *then denoting*

$$z_t := \int_0^t \phi_r \circ \mathrm{d}\mathbf{x}_r, \qquad z'_t := \phi_t, \qquad z''_t := \phi'_t,$$

(z, z', z'') *is again controlled by* $\mathbf{x}$, *and we have*

$$\|z\|_{p,q} \le C_{p,q} \|\phi\|_{p,q\text{-}cvar} \big(1 + \|x\|_{p\text{-}var;[0,T]} + \big\|\mathbf{x}^2\big\|_{\frac{p}{2}\text{-}var;[0,T]} + \big\|\mathbf{x}^3\big\|_{\frac{p}{3}\text{-}var;[0,T]}\big). \tag{41}$$

For the next proposition, given maps $A \in \mathcal{L}(\mathbb{R}^d; \mathcal{L}(\mathcal{U}; \mathcal{V}))$ and $B \in \mathcal{L}(\mathbb{R}^d; \mathcal{U})$, we will identify them as tensors (either $\mathcal{L}(\mathcal{U}; \mathcal{V})$-valued or $\mathcal{U}$-valued)

$$A = \sum_{j=1}^d a_j \,\mathrm{d}e_j, \quad a_j \in \mathcal{L}(\mathcal{U}; \mathcal{V}), \qquad B = \sum_{j=1}^d b_j \,\mathrm{d}e_j, \quad b_j \in \mathcal{U},$$

and use the following notation

$$\begin{aligned}
&AB := a_i(b_j)\,\mathrm{d}e_i \otimes \mathrm{d}e_j, \qquad BA := a_j(b_i)\,\mathrm{d}e_i \otimes \mathrm{d}e_j,\\
&\mathrm{Sym}(AB) := \frac{1}{2}(AB + BA).
\end{aligned}$$

Proposition 4.3 (Leibniz rule). *For* $3 \le p < 4$, *let*

$$\begin{aligned}
&\phi \in \mathcal{C}^{p\text{-}var}\big([0,T]; \mathcal{L}(\mathcal{U}; \mathcal{V})\big),\\
&\phi' \in \mathcal{C}^{p\text{-}var}\big([0,T]; \mathcal{L}\big(\mathbb{R}^d; \mathcal{L}(\mathcal{U}; \mathcal{V})\big)\big) \quad \text{and}\\
&\phi'' \in \mathcal{C}^{p\text{-}var}\big([0,T]; \mathcal{L}\big(\mathbb{R}^d \otimes \mathbb{R}^d; \mathcal{L}(\mathcal{U}; \mathcal{V})\big)\big).
\end{aligned}$$

Assume that (ϕ, ϕ', ϕ'') *is controlled by* $\mathbf{x} \in \mathcal{C}^{p\text{-}var}([0,T]; G^3(\mathbb{R}^d))$, *with remainder terms* R^ϕ *and* $R^{\phi'}$ *of bounded* q*-variation and* $\frac{p}{2}$*-variation respectively, where* $\frac{1}{p} + \frac{1}{q} > 1$ *and* $q \ge \frac{p}{3}$.

(i) *Let* $\psi \in \mathcal{C}^{p\text{-}var}([0,T];\mathcal{U})$, $\psi' \in \mathcal{C}^{p\text{-}var}([0,T];\mathcal{L}(\mathbb{R}^d;\mathcal{U}))$ *and* $\psi'' \in \mathcal{C}^{p\text{-}var}([0,T];\mathcal{L}(\mathbb{R}^d \otimes \mathbb{R}^d;\mathcal{U}))$. *If* (ψ, ψ', ψ'') *is controlled by* $\mathbf{x}$, *then the path* $\phi\psi \in \mathcal{C}^{p\text{-}var}([0,T];\mathcal{V})$ *given by the composition of* ϕ *and* ψ *is also controlled by* $\mathbf{x}$, *with first and second derivative processes*

$$(\phi\psi)' = \phi'\psi + \phi\psi' \quad \text{and} \quad (\phi\psi)'' = \phi''\psi + 2\mathrm{Sym}(\phi'\psi') + \phi\psi''.$$

In addition, we have the bound

$$\|\phi\psi\|_{p,q\text{-}cvar} \le 4\|\phi\|_{p,q\text{-}cvar}\|\psi\|_{p,q\text{-}cvar}\left(1 + \|x\|_{p\text{-}var;[0,T]} + \left\|\mathbf{x}^2\right\|_{\frac{p}{2}\text{-}var;[0,T]}\right). \tag{42}$$

(ii) *Suppose that* $\psi \in \mathcal{C}^{q\text{-}var}([0,T];\mathcal{U})$. *Then* $\phi\psi \in \mathcal{C}^{p\text{-}var}([0,T];\mathcal{V})$ *is also controlled by* $\mathbf{x}$, *with first and second derivative processes*

$$(\phi\psi)' = \phi'\psi \quad \text{and} \quad (\phi\psi)'' = \phi''\psi.$$

Moreover, we have the bound

$$\|\phi\psi\|_{p,q\text{-}cvar} \le \|\phi\|_{p,q\text{-}cvar}\|\psi\|_{\mathcal{V}^q;[0,T]}. \tag{43}$$

Proposition 4.4. *Let* $\mathbf{x} \in \mathcal{C}^{p\text{-}var}([0,T];G^3(\mathbb{R}^d))$ *where* $3 \le p < 4$. *We assume that*

$$y \in \mathcal{C}^{p\text{-}var}\big([0,T];\mathcal{U}\big), \qquad y' \in \mathcal{C}^{p\text{-}var}\big([0,T];\mathcal{L}\big(\mathbb{R}^d,\mathcal{U}\big)\big) \quad \text{and} \quad y'' \in \mathcal{C}^{p\text{-}var}\big([0,T];\mathcal{L}\big(\mathbb{R}^d \otimes \mathbb{R}^d;\mathcal{U}\big)\big).$$

We also assume that (y, y', y'') *is controlled by* $\mathbf{x}$ *with remainder terms* R^y *and* $R^{y'}$ *of bounded* q*-variation and* $\frac{p}{2}$*-variation respectively, where* $\frac{1}{p} + \frac{1}{q} > 1$ *and* $q \ge \frac{p}{3}$. *Let* $\phi \in \mathcal{C}_b^3(\mathcal{U},\mathcal{V})$ *and define*

$$\left(z_s, z_s', z_s''\right) := \left(\phi(y_s), \nabla\phi(y_s)y_s', \nabla\phi(y_s)y_s'' + \nabla^2\phi(y_s)\left(y_s', y_s'\right)\right)$$

for all $s \in [0,T]$. *Then* (z, z', z'') *is controlled by* $\mathbf{x}$, *and we have the following bounds*

$$\|z\|_{p\text{-}var;[0,T]} \le \|\phi\|_{\mathcal{C}_b^3}\|y\|_{p\text{-}var;[0,T]},$$
$$\left\|z'\right\|_{p\text{-}var;[0,T]} \le \|\phi\|_{\mathcal{C}_b^3}\|y\|_{\mathcal{V}^p;[0,T]}\left\|y'\right\|_{\mathcal{V}^p;[0,T]},$$
$$\left\|z''\right\|_{p\text{-}var;[0,T]} \le \|\phi\|_{\mathcal{C}_b^3}\|y\|_{\mathcal{V}^p;[0,T]}\left(\left\|y''\right\|_{\mathcal{V}^p;[0,T]} + \left\|y'\right\|^2_{\mathcal{V}^p;[0,T]}\right),$$

and

$$\left\|R^z\right\|_{q\text{-}var;[0,T]}, \left\|R^{z'}\right\|_{\frac{p}{2}\text{-}var;[0,T]} \le \|\phi\|_{\mathcal{C}_b^3}\left(1 + \|y\|_{p,q\text{-}cvar}\right)^3\left(1 + \|x\|_{p\text{-}var;[0,T]} + \left\|\mathbf{x}^2\right\|_{\frac{p}{2}\text{-}var;[0,T]}\right)^2. \tag{44}$$

The following theorem extends Theorem 3.1 in [7].

Theorem 4.5. *Consider the system of RDEs*

$$\mathrm{d}y_t = V(y_t) \circ \mathrm{d}\mathbf{x}_t, \qquad y_0 = a \in \mathbb{R}^e,$$
$$\mathrm{d}J_t^{\mathbf{x}} = \nabla V(y_t)(\circ\, \mathrm{d}\mathbf{x}_t)J_t^{\mathbf{x}}, \qquad J_0^{\mathbf{x}} = \mathcal{I}_e,$$

where $V = (V_1, \dots, V_d)$ *is a collection of* $\mathbb{R}^e$*-valued vector fields. If* $\mathbf{x} = (1, x, \mathbf{x}^2, \mathbf{x}^3) \in \mathcal{C}^{p\text{-}var}([0,T];G^3(\mathbb{R}^d))$, $3 \le p < 4$, *and* V *is in* $\mathcal{C}_b^4$, *then both* $(y, V(y), V^2(y))$ *and* $(J^{\mathbf{x}}, (J^{\mathbf{x}})', (J^{\mathbf{x}})'')$ *are controlled by* $\mathbf{x}$. *In addition,*

$$\|y\|_{p,\frac{p}{3}\text{-}cvar} \le C_p\left(1 + \|V\|_{\mathcal{C}_b^3}\right)^{10}\left(1 + \|\mathbf{x}\|_{p\text{-}var;[0,T]}\right)^8, \tag{45}$$

and

$$\left\|J^{\mathbf{x}}\right\|_{p,\frac{p}{3}\text{-}cvar} \le C_1\left(1 + \exp\left(C_2 N^{\mathbf{x}}_{1;[0,T]}\right)\right)^{10}\left(1 + \|\mathbf{x}\|_{p\text{-}var;[0,T]}\right)^8, \tag{46}$$

where C_1, C_2 *depend on* p *and* $\|V\|_{\mathcal{C}_b^4}$.

4.2. *Upper bounds on the high-order Malliavin derivatives*

We now use the results from the proceeding section to obtain upper bounds on the directional derivative. The following result extends Proposition 3.5 in [7].

Proposition 4.6. *Let $p \in [2,4)$ and q be such that $\frac{1}{p}+\frac{1}{q}>1$. Let $\mathbf{x} \in \mathcal{C}^{0,p\text{-}var}([0,T]; G^{\lfloor p \rfloor}(\mathbb{R}^d))$, and y be the solution to the RDE*

$$dy_t = V(y_t) \circ d\mathbf{x}_t, \qquad y_0 \in \mathbb{R}^e \text{ given},$$

where $V \in \mathcal{C}_b^{\lfloor p \rfloor + n}(\mathbb{R}^e; \mathbb{R}^e \otimes \mathbb{R}^d)$. Then there exists a polynomial $P_{d(n)} : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}_+$ of finite degree $d(n)$ for which

$$\big\| \mathcal{D}^n_{g_1,\dots,g_n} y_\cdot \big\|_{\mathcal{V}^p;[0,T]} \le P_{d(n)}\big(\|\mathbf{x}\|_{p\text{-}var;[0,T]}, \exp\big(CN^{\mathbf{x}}_{1;[0,T]}\big)\big) \prod_{i=1}^n \|g_i\|_{q\text{-}var;[0,T]}, \tag{47}$$

for any $g_1,\dots,g_n \in \mathcal{C}^{q\text{-}var}([0,T];\mathbb{R}^d)$. The constant C as well as the coefficients of $P_{d(n)}$ depend only on $\|V\|_{\mathcal{C}_b^{\lfloor p \rfloor + n}}$, p and q ($=\frac{p}{2}$ when $2 \le p < 3$).

Proof. We shall omit the details of the proof as it proceeds in virtually the same manner as in Proposition 3.5 in [7], for which the reader is invited to consult. The key again is to use the explicit expression of $\mathcal{D}^n_{g_1,\dots,g_n} y_t$ as the sum of the integrals

$$\sum_{i=2}^n \int_0^t J_t^{\mathbf{x}}\big(J_s^{\mathbf{x}}\big)^{-1} \nabla^i V(y_s) A_i^n(s) \circ d\mathbf{x}_s + \sum_{i=1}^{n-1}\sum_{j=1}^n \int_0^t J_t^{\mathbf{x}}\big(J_s^{\mathbf{x}}\big)^{-1} \nabla^i V(y_s) B_{i,j}^n(s)\, dg_j(s) \tag{48}$$

as derived in Corollary 3.4 of [7]. Here, A_i^n and $B_{i,j}^n$ are respectively defined as

$$A_i^n(t) := \sum_{\pi=\{\pi_1,\dots,\pi_i\} \in \mathcal{P}(\{g_1,\dots,g_n\})} \mathcal{D}^{|\pi_1|}_{\pi_1} y_t \tilde{\otimes} \cdots \tilde{\otimes} \mathcal{D}^{|\pi_i|}_{\pi_i} y_t, \quad t \in [0,T], \tag{49}$$

and

$$B_{i,j}^n(t) := \sum_{\pi=\{\pi_1,\dots,\pi_i\} \in \mathcal{P}(\{g_1,\dots,g_{j-1},g_{j+1},\dots,g_n\})} \mathcal{D}^{|\pi_1|}_{\pi_1} y_t \tilde{\otimes} \cdots \tilde{\otimes} \mathcal{D}^{|\pi_i|}_{\pi_i} y_t. \tag{50}$$

The only difference is that when $p \ge 3$, we use Theorems 5.2 and 5.5, as well as Propositions 5.3 and 5.4 in lieu of Theorem 2.20 and 3.5, and Propositions 2.22 and 2.21 respectively in [7]. □

4.3. *Augmenting the higher-order iterated integrals*

For this section, we will use $\pi(n) := \{t_i^n\}$ to denote the nth dyadic partition of $[0,T]$, i.e. $t_i^n = \frac{iT}{2^n}$ for $i = 0,\dots,2^n$, and Δ_i^n to denote the interval $[t_i^n, t_{i+1}^n]$. In addition, ρ' will denote the Hölder conjugate of ρ, i.e. $\frac{1}{\rho}+\frac{1}{\rho'}=1$.

One of the main results (Proposition 5.1) in [7] is the following: given a Gaussian process X and a stochastic process ψ, under certain conditions on the covariance of X and bounds on the Malliavin derivatives of ψ, one can show that

$$\lim_{n\to\infty} \left\| \sum_{i=0}^{2^n-1} \psi_{t_i^n}\left(\mathbf{X}^2_{t_i^n,t_{i+1}^n} - \frac{1}{2}\sigma^2\big(t_i^n,t_{i+1}^n\big)\mathcal{I}_d \right) \right\|_{L^2(\Omega)} = 0. \tag{51}$$

Recall that $\sigma^2(s,t)$ denotes $\mathbb{E}[X^{(1)}_{s,t}]^2$, and that the presence of this compensating term was critical in the proof of the correction formula in [7]. We will now proceed to show that for the third-order terms, this is not needed for the L^2 norm to vanish. We begin with the following lemma.

Lemma 4.7. *For any $h_1,\dots,h_6 \in \mathcal{H}_1^d$, we have*

$$\mathbb{E}\left[\psi_{t_i^n}\psi_{t_j^n} \prod_{k=1}^6 I_1(h_k) \right] = \mathbb{E}\big[\mathcal{D}^6_{h_1,h_2,h_3,h_4,h_5,h_6} \psi_{t_i^n}\psi_{t_j^n}\big] + \sum_{\sigma \in \mathcal{S}_6} C_{\sigma,1}A_{\sigma,1} + C_{\sigma,2}A_{\sigma,2} + C_{\sigma,3}A_{\sigma,3},$$

where

$$A_{\sigma,1} := \mathbb{E}\big[\mathcal{D}^4_{h_{\sigma(1)},h_{\sigma(2)},h_{\sigma(3)},h_{\sigma(4)}}\psi_{t_i^n}\psi_{t_j^n}\big]\langle h_{\sigma(5)}, h_{\sigma(6)}\rangle_{\mathcal{H}_1^d},$$
$$A_{\sigma,2} := \mathbb{E}\big[\mathcal{D}^2_{h_{\sigma(1)},h_{\sigma(2)}}\psi_{t_i^n}\psi_{t_j^n}\big]\langle h_{\sigma(3)}, h_{\sigma(4)}\rangle_{\mathcal{H}_1^d}\langle h_{\sigma(5)}, h_{\sigma(6)}\rangle_{\mathcal{H}_1^d},$$
$$A_{\sigma,3} := \mathbb{E}[\psi_{t_i^n}\psi_{t_j^n}]\langle h_{\sigma(1)}, h_{\sigma(2)}\rangle_{\mathcal{H}_1^d}\langle h_{\sigma(3)}, h_{\sigma(4)}\rangle_{\mathcal{H}_1^d}\langle h_{\sigma(5)}, h_{\sigma(6)}\rangle_{\mathcal{H}_1^d},$$

$\mathcal{S}_6$ denotes the symmetric group of permutations on $\{1,\dots,6\}$, and the $C_{\sigma,k}$'s are constants that depend on the permutation σ.

Proof. Using the product formula (cf. [27]), we have

$$\begin{aligned}\prod_{k=1}^{6} I_1(h_k) = {} & I_6(h_1\otimes h_2\otimes h_3\otimes h_4\otimes h_5\otimes h_6)\\ & + \sum_{\sigma\in\mathcal{S}_6} C_{\sigma,1} I_4(h_{\sigma(1)}\otimes h_{\sigma(2)}\otimes h_{\sigma(3)}\otimes h_{\sigma(4)})\langle h_{\sigma(5)}, h_{\sigma(6)}\rangle_{\mathcal{H}_1^d}\\ & + C_{\sigma,2} I_2(h_{\sigma(1)}\otimes h_{\sigma(2)})\langle h_{\sigma(3)}, h_{\sigma(4)}\rangle_{\mathcal{H}_1^d}\langle h_{\sigma(5)}, h_{\sigma(6)}\rangle_{\mathcal{H}_1^d}\\ & + C_{\sigma,3}\langle h_{\sigma(1)}, h_{\sigma(2)}\rangle_{\mathcal{H}_1^d}\langle h_{\sigma(3)}, h_{\sigma(4)}\rangle_{\mathcal{H}_1^d}\langle h_{\sigma(5)}, h_{\sigma(6)}\rangle_{\mathcal{H}_1^d}.\end{aligned} \tag{52}$$

Applying integration-by-parts finishes the proof. □

Proposition 4.8. *Let $\mathbf{X}\in\mathcal{C}^{0,p\text{-}var}([0,T];G^3(\mathbb{R}^d))$, $3\le p<4$, be a geometric Gaussian rough path which satisfies Condition 1, and assume that its covariance function satisfies, for all $s,t\in[0,T]$,*

$$\big\|R(t,\cdot)-R(s,\cdot)\big\|_{q\text{-}var;[0,T]}\le C|t-s|^{\frac{1}{\rho}},$$

for some finite constant C and $\rho\in[1,2)$.

Let $\psi:\Omega\times[0,T]\to\mathbb{R}^d\otimes\mathbb{R}^d\otimes\mathbb{R}^d$ be a stochastic process satisfying $\psi_t=\sum_{a,b,c=1}^d\psi_t^{(a,b,c)}\,\mathrm{d}e_a\otimes\mathrm{d}e_b\otimes\mathrm{d}e_c\in\mathbb{D}^{6,2}(\mathbb{R}^d\otimes\mathbb{R}^d\otimes\mathbb{R}^d)$ for all $t\in[0,T]$. Furthermore, assume there exists $C<\infty$ such that we have

$$\big|\mathbb{E}\big[\psi_s^{(a,b,c)}\psi_t^{(a,b,c)}\big]\big|\le C, \tag{53}$$

and for $k=2,4,6$, we have

$$\big|\mathbb{E}\big[\mathcal{D}^k_{h_1,\dots,h_k}\big(\psi_s^{(a,b,c)}\psi_t^{(a,b,c)}\big)\big]\big|\le C\prod_{i=1}^{k}\big\|\Phi(h_i)\big\|_{q\text{-}var;[0,T]}, \tag{54}$$

for all $h_1,\dots,h_k\in\mathcal{H}_1^d$, $s,t\in[0,T]$ and $a,b,c=1,\dots,d$. Then

$$\lim_{n\to\infty}\bigg\|\sum_{i=0}^{2^n-1}\psi_{t_i^n}\big(\mathbf{X}^3_{t_i^n,t_{i+1}^n}\big)\bigg\|_{L^2(\Omega)}=0. \tag{55}$$

Proof. First note that

$$\bigg\|\sum_{i=0}^{2^n-1}\psi_{t_i^n}\big(\mathbf{X}^3_{t_i^n,t_{i+1}^n}\big)\bigg\|_{L^2(\Omega)}\le\bigg\|\sum_{i=0}^{2^n-1}\psi_{t_i^n}\big(\big(\mathbf{X}^3_{t_i^n,t_{i+1}^n}\big)^S\big)\bigg\|_{L^2(\Omega)}+\bigg\|\sum_{i=0}^{2^n-1}\psi_{t_i^n}\big(\big(\mathbf{X}^3_{t_i^n,t_{i+1}^n}\big)^{NS}\big)\bigg\|_{L^2(\Omega)}, \tag{56}$$

where $(\mathbf{X}^3)^S$ denotes the symmetric part of $\mathbf{X}^3$ and

$$\big(\mathbf{X}^3_{s,t}\big)^{NS}=\mathbf{X}^3_{s,t}-\big(\mathbf{X}^3_{s,t}\big)^S$$

denotes the non-symmetric part. The two parts will be tackled separately, and since

$$\left\| \sum_{i=0}^{2^n-1} \psi_{t_i^n}\big((\mathbf{X}^3_{t_i^n,t_{i+1}^n})^S\big) \right\|_{L^2(\Omega)} \le \sum_{a,b,c=1}^{d} \left\| \sum_{i=0}^{2^n-1} \psi_{t_i^n}^{(a,b,c)}\big((\mathbf{X}^3_{t_i^n,t_{i+1}^n})^S\big)^{(a,b,c)} \right\|_{L^2(\Omega)},$$

and similarly for the non-symmetric part, we will study the convergence of each fixed (a,b,c)th tensor component individually and henceforth suppress the notation for the component in the superscript of ψ.

(a) To begin, we will prove that

$$\lim_{n\to\infty} \left\| \sum_{i=0}^{2^n-1} \psi_{t_i^n}\big(\mathbf{X}^3_{t_i^n,t_{i+1}^n}\big)^S \right\|_{L^2(\Omega)} = 0. \tag{57}$$

Since

$$\big(\mathbf{X}^3_{s,t}\big)^S = \frac{1}{6} X_{s,t} \otimes X_{s,t} \otimes X_{s,t},$$

this is equivalent to showing that

$$\mathbb{E}\left[\left(\sum_{i=0}^{2^n-1} \psi_{t_i^n}\big((\mathbf{X}^3_{t_i^n,t_{i+1}^n})^S\big)^{(a,b,c)}\right)^2\right] = \frac{1}{36}\sum_{i,j=0}^{2^n-1} \mathbb{E}\big[\psi_{t_i^n}\psi_{t_j^n} X^{(a)}_{\Delta_i^n} X^{(b)}_{\Delta_i^n} X^{(c)}_{\Delta_i^n} X^{(a)}_{\Delta_j^n} X^{(b)}_{\Delta_j^n} X^{(c)}_{\Delta_j^n}\big]$$

converges to zero as $n \to \infty$. First define

$$h_1 := \mathbb{1}^{(a)}_{\Delta_i^n}, \qquad h_2 := \mathbb{1}^{(b)}_{\Delta_i^n}, \qquad h_3 := \mathbb{1}^{(c)}_{\Delta_i^n}, \qquad h_4 := \mathbb{1}^{(a)}_{\Delta_j^n}, \qquad h_5 := \mathbb{1}^{(b)}_{\Delta_j^n} \quad \text{and} \quad h_6 := \mathbb{1}^{(c)}_{\Delta_j^n}.$$

Note that for $k = 1, \dots, 6$, we have

$$\begin{aligned}\|\Phi(h_k)\|_{q\text{-var};[0,T]} &= \|R(t_{i+1}^n,\cdot) - R(t_i^n,\cdot)\|_{q\text{-var};[0,T]} \text{ or } \|R(t_{j+1}^n,\cdot) - R(t_j^n,\cdot)\|_{q\text{-var};[0,T]}\\ &\le C2^{-\frac{n}{\rho}}\end{aligned}$$

and

$$\begin{aligned}\|h_k\|_{\mathcal{H}_1^d} &= \sqrt{\sigma^2(t_i^n, t_{i+1}^n)} \text{ or } \sqrt{\sigma^2(t_j^n, t_{j+1}^n)}\\ &\le \sqrt{\|R(t_{i+1}^n,\cdot) - R(t_i^n,\cdot)\|_{q\text{-var};[0,T]}} \text{ or } \sqrt{\|R(t_{j+1}^n,\cdot) - R(t_j^n,\cdot)\|_{q\text{-var};[0,T]}}\\ &\le C2^{-\frac{n}{2\rho}}.\end{aligned}$$

Recall from Lemma 4.7 that

$$\mathbb{E}\left[\psi_{t_i^n}\psi_{t_j^n}\prod_{k=1}^{6} I_1(h_k)\right] =: \mathbb{E}\big[\mathcal{D}^6_{h_1,h_2,h_3,h_4,h_5,h_6}\psi_{t_i^n}\psi_{t_j^n}\big] + \sum_{\sigma\in\mathcal{S}_6} C_{\sigma,1}A_{\sigma,1} + C_{\sigma,2}A_{\sigma,2} + C_{\sigma,3}A_{\sigma,3}.$$

For the first term on the right side of the above expression, we have

$$\sum_{i,j=0}^{2^n-1}\mathbb{E}\big[\mathcal{D}^6_{h_1,h_2,h_3,h_4,h_5,h_6}\psi_{t_i^n}\psi_{t_j^n}\big] \le C\sum_{i,j=0}^{2^n-1}\prod_{k=1}^{6}\|\Phi(h_k)\|_{q\text{-var};[0,T]} \le C2^{-2n(\frac{3}{\rho}-1)},$$

which vanishes as $n \to \infty$ since $\rho < 2$.

For the $A_{\sigma,1}$ terms we have

$$\sum_{i,j=0}^{2^n-1}\mathbb{E}\big[\mathcal{D}^4_{h_{\sigma(1)},h_{\sigma(2)},h_{\sigma(3)},h_{\sigma(4)}}\psi_{t_i^n}\psi_{t_j^n}\big]\langle h_{\sigma(5)}, h_{\sigma(6)}\rangle_{\mathcal{H}_1^d}$$

$$\leq \sum_{i,j=0}^{2^n-1} \prod_{k=1}^{4} \left\| \Phi(h_{\sigma(k)}) \right\|_{q\text{-var};[0,T]} \|h_{\sigma(5)}\|_{\mathcal{H}_1^d} \|h_{\sigma(6)}\|_{\mathcal{H}_1^d}$$

$$\leq C 2^{-2n(\frac{5}{2\rho}-1)} \to 0,$$

and similarly for the $A_{\sigma,2}$ terms we have

$$\sum_{i,j=0}^{2^n-1} \mathbb{E}\big[\mathcal{D}^2_{h_{\sigma(1)},h_{\sigma(2)}} \psi_{t_i^n} \psi_{t_j^n}\big] \langle h_{\sigma(3)}, h_{\sigma(4)} \rangle_{\mathcal{H}_1^d} \langle h_{\sigma(5)}, h_{\sigma(6)} \rangle_{\mathcal{H}_1^d}$$

$$\leq \sum_{i,j=0}^{2^n-1} \left\| \Phi(h_{\sigma(1)}) \right\|_{q\text{-var};[0,T]} \left\| \Phi(h_{\sigma(2)}) \right\|_{q\text{-var};[0,T]} \prod_{k=3}^{6} \|h_{\sigma(k)}\|_{\mathcal{H}_1^d}$$

$$\leq C 2^{-2n(\frac{2}{\rho}-1)} \to 0.$$

Finally for the $A_{\sigma,3}$ terms we have two cases: either

$$\langle h_{\sigma(1)}, h_{\sigma(2)} \rangle_{\mathcal{H}_1^d} \langle h_{\sigma(3)}, h_{\sigma(4)} \rangle_{\mathcal{H}_1^d} \langle h_{\sigma(5)}, h_{\sigma(6)} \rangle_{\mathcal{H}_1^d} = R \begin{pmatrix} t_i^n & t_{i+1}^n \\ t_j^n & t_{j+1}^n \end{pmatrix}^3,$$

or

$$\langle h_{\sigma(1)}, h_{\sigma(2)} \rangle_{\mathcal{H}_1^d} \langle h_{\sigma(3)}, h_{\sigma(4)} \rangle_{\mathcal{H}_1^d} \langle h_{\sigma(5)}, h_{\sigma(6)} \rangle_{\mathcal{H}_1^d} = R \begin{pmatrix} t_i^n & t_{i+1}^n \\ t_j^n & t_{j+1}^n \end{pmatrix} \sigma^2\big(t_i^n, t_{i+1}^n\big) \sigma^2\big(t_j^n, t_{j+1}^n\big).$$

In either case, since

$$\left| R \begin{pmatrix} t_i^n & t_{i+1}^n \\ t_j^n & t_{j+1}^n \end{pmatrix} \right|, \sigma^2\big(t_i^n, t_{i+1}^n\big), \sigma^2\big(t_j^n, t_{j+1}^n\big) \leq \frac{C}{2^{\frac{n}{\rho}}},$$

we have

$$\sum_{i,j=0}^{2^n-1} \mathbb{E}[\psi_{t_i^n} \psi_{t_j^n}] \langle h_{\sigma(1)}, h_{\sigma(2)} \rangle_{\mathcal{H}_1^d} \langle h_{\sigma(3)}, h_{\sigma(4)} \rangle_{\mathcal{H}_1^d} \langle h_{\sigma(5)}, h_{\sigma(6)} \rangle_{\mathcal{H}_1^d}$$

$$\leq C \left(\sum_{i,j=0}^{2^n-1} R \begin{pmatrix} t_i^n & t_{i+1}^n \\ t_j^n & t_{j+1}^n \end{pmatrix}^{\rho} \right)^{\frac{1}{\rho}} \left(\sum_{i,j=0}^{2^n-1} 2^{\frac{-2n\rho'}{\rho}} \right)^{\frac{1}{\rho'}}$$

$$\leq C \|R\|_{\rho\text{-var};[0,T]^2} 2^{-2n(\frac{1}{\rho}-\frac{1}{\rho'})}$$

$$\leq C \|R\|_{\rho\text{-var};[0,T]^2} 2^{-2n(\frac{2}{\rho}-1)} \to 0.$$

(b) We will now move on to show that

$$\lim_{n\to\infty} \left\| \sum_{i=0}^{2^n-1} \psi_{t_i^n} \big(\mathbf{X}^3_{t_i^n, t_{i+1}^n}\big)^{NS} \right\|_{L^2(\Omega)} = 0. \tag{58}$$

Let $X^{\pi(k)}$ denote the piece-wise linear approximation of X over $\pi(k)$, and let $\mathbf{X}^{\pi(k)} = (1, \mathbf{X}^1(\pi(k)), \mathbf{X}^2(\pi(k)), \mathbf{X}^3(\pi(k))) = S_3(X^{\pi(k)})$ denote its canonical lift to a geometric rough path.

Next, define

$$\big(\mathbf{X}^3_{s,t}\big)^{NS}(\Delta_{l+1}) := \big(\mathbf{X}^3_{s,t}\big)^{NS}\big(\pi(l+1)\big) - \big(\mathbf{X}^3_{s,t}\big)^{NS}\big(\pi(l)\big).$$

Since $(\mathbf{X}^3_{t_i^n,t_{i+1}^n})^{NS}(\pi(n)) = 0$, we have

$$\big(\mathbf{X}^3_{t_i^n,t_{i+1}^n}\big)^{NS} = \lim_{m\to\infty}\sum_{k=1}^{m}\big(\mathbf{X}^3_{t_i^n,t_{i+1}^n}\big)^{NS}(\Delta_{n+k}) \quad \text{for every } n\in\mathbb{N} \text{ and } i = 0,1,\dots,2^n-1,$$

where the limit is taken in $L^2(\Omega)$.

We want to show that

$$\left\| \sum_{i=0}^{2^n-1} \psi_{t_i^n}\big(\big(\mathbf{X}^3_{t_i^n,t_{i+1}^n}\big)^{NS}\big(\pi(n+m)\big)\big)^{(a,b,c)}\right\|_{L^2(\Omega)} \to 0$$

uniformly for all m as $n\to\infty$. To begin, let

$$s_u^{k,i} := t_i^n + \frac{u}{2^{n+k}} = t^{n+k}_{u+i2^k}, \tag{59}$$

and we will denote the intervals

$$\begin{aligned} &\Delta^i_{uL} := \big[s^{k+1,i}_{2u}, s^{k+1,i}_{2u+1}\big], \qquad \Delta^i_{uR} := \big[s^{k+1,i}_{2u+1}, s^{k+1,i}_{2u+2}\big],\\ &\Delta^i_u := \Delta^i_{uL}\cup\Delta^i_{uR} = \big[s^{k,i}_u, s^{k,i}_{u+1}\big]\subset\big[t_i^n, t_{i+1}^n\big], \quad \forall u = 0,\dots 2^k-1. \end{aligned} \tag{60}$$

Note that we suppress the dependence on k and n in the notation for the variables on the left. The following computation on $G^3(\mathbb{R}^d)$ can be verified easily; for $f, g\in\mathbb{R}^d$ we have

$$\exp(f)\otimes\exp(g) = \left(1, f+g, \frac{(f+g)^{\otimes 2}}{2} + \frac{1}{2}[f,g], \frac{(f+g)^{\otimes 3}}{6} + N(f,g)\right),$$

where

$$N(f,g) := \frac{1}{4}\big((f+g)\otimes[f,g] + [f,g]\otimes(f+g)\big) + \frac{1}{12}\big(\big[f,[f,g]\big] + \big[g,[g,f]\big]\big).$$

Using the above expression with $f = X_{\Delta^i_{uL}}$ and $g = X_{\Delta^i_{uR}}$, for $k = 1,\dots,m$ we obtain the following identity on $T^3(\mathbb{R}^d)$:

$$\begin{aligned} &\bigotimes_{u=0}^{2^k-1}\exp(X_{\Delta^i_{uL}})\otimes\exp(X_{\Delta^i_{uR}}) - \bigotimes_{u=0}^{2^k-1}\exp(X_{\Delta^i_u})\\ &\quad = \bigotimes_{u=0}^{2^k-1}\left[\left(1, X_{\Delta^i_u}, \frac{1}{2}X^{\otimes 2}_{\Delta^i_u}, \frac{1}{6}X^{\otimes 3}_{\Delta^i_u}\right) + \left(0,0,\frac{1}{2}[X_{\Delta^i_{uL}}, X_{\Delta^i_{uR}}], 0\right) + \big(0,0,0,N(X_{\Delta^i_{uL}}, X_{\Delta^i_{uR}})\big)\right]\\ &\qquad - \bigotimes_{u=0}^{2^k-1}\left(1, X_{\Delta^i_u}, \frac{1}{2}X^{\otimes 2}_{\Delta^i_u}, \frac{1}{6}X^{\otimes 3}_{\Delta^i_u}\right)\\ &\quad = \sum_{u=0}^{2^k-1}\left(0,0,\frac{1}{2}[X_{\Delta^i_{uL}}, X_{\Delta^i_{uR}}], M(X_{\Delta^i_{uL}}, X_{\Delta^i_{uR}}) + N(X_{\Delta^i_{uL}}, X_{\Delta^i_{uR}})\right), \end{aligned}$$

where

$$\begin{aligned} M(X_{\Delta^i_{uL}}, X_{\Delta^i_{uR}}) &:= \sum_{r=0}^{u-1} X_{\Delta^i_r}\otimes\frac{1}{2}[X_{\Delta^i_{uL}}, X_{\Delta^i_{uR}}] + \frac{1}{2}[X_{\Delta^i_{uL}}, X_{\Delta^i_{uR}}]\otimes\sum_{r=u+1}^{2^k-1}X_{\Delta^i_r}\\ &= X_{t_i^n, s_u^{k,i}}\otimes\frac{1}{2}[X_{\Delta^i_{uL}}, X_{\Delta^i_{uR}}] + \frac{1}{2}[X_{\Delta^i_{uL}}, X_{\Delta^i_{uR}}]\otimes X_{s^{k,i}_{u+1}, t^n_{i+1}}. \end{aligned}$$

This means that

$$\mathbf{X}^3_{t_i^n,t_{i+1}^n}\big(\pi(n+k+1)\big)-\mathbf{X}^3_{t_i^n,t_{i+1}^n}\big(\pi(n+k)\big)=\sum_{u=0}^{2^k-1}M_u+N_u,$$

where we use M_u and N_u as short-hand for $M(X_{\Delta^i_{uL}},X_{\Delta^i_{uR}})$ and $N(X_{\Delta^i_{uL}},X_{\Delta^i_{uR}})$ respectively. This in turn gives us

$$\big(\mathbf{X}^3_{t_i^n,t_{i+1}^n}\big)^{NS}(\Delta_{n+k})=\sum_{u=0}^{2^k-1}M_u+N_u,$$

since

$$\begin{aligned}
&\mathbf{X}^3_{t_i^n,t_{i+1}^n}\big(\pi(n+k+1)\big)-\mathbf{X}^3_{t_i^n,t_{i+1}^n}\big(\pi(n+k)\big)\\
&\quad=\big(\mathbf{X}^3_{t_i^n,t_{i+1}^n}\big)^{S}\big(\pi(n+k+1)\big)-\big(\mathbf{X}^3_{t_i^n,t_{i+1}^n}\big)^{S}\big(\pi(n+k)\big)+\big(\mathbf{X}^3_{t_i^n,t_{i+1}^n}\big)^{NS}(\Delta_{n+k})\\
&\quad=\exp(X_{t_i^n,t_{i+1}^n})-\exp(X_{t_i^n,t_{i+1}^n})+\big(\mathbf{X}^3_{t_i^n,t_{i+1}^n}\big)^{NS}(\Delta_{n+k})\\
&\quad=\big(\mathbf{X}^3_{t_i^n,t_{i+1}^n}\big)^{NS}(\Delta_{n+k}).
\end{aligned}$$

Thus, we obtain

$$\begin{aligned}
&\mathbb{E}\Bigg[\Bigg(\sum_{i=0}^{2^n-1}\psi_{t_i^n}\big(\big(\mathbf{X}^3_{t_i^n,t_{i+1}^n}\big)^{NS}\big(\pi(n+m)\big)\big)^{(a,b,c)}\Bigg)^2\Bigg]\\
&\quad=\mathbb{E}\Bigg[\Bigg(\sum_{i=0}^{2^n-1}\psi_{t_i^n}\sum_{k=1}^{m}\big(\big(\mathbf{X}^3_{t_i^n,t_{i+1}^n}\big)^{NS}(\Delta_{n+k})\big)^{(a,b,c)}\Bigg)^2\Bigg]\\
&\quad=\sum_{i,j=0}^{2^n-1}\mathbb{E}\Bigg[\psi_{t_i^n}\psi_{t_j^n}\sum_{k=1}^{m}\big(\big(\mathbf{X}^3_{t_i^n,t_{i+1}^n}\big)^{NS}(\Delta_{n+k})\big)^{(a,b,c)}\sum_{l=1}^{m}\big(\big(\mathbf{X}^3_{t_j^n,t_{j+1}^n}\big)^{NS}(\Delta_{n+l})\big)^{(a,b,c)}\Bigg]\\
&\quad=\sum_{i,j=0}^{2^n-1}\sum_{k,l=1}^{m}\sum_{u=0}^{2^k-1}\sum_{v=0}^{2^l-1}\mathbb{E}\big[\psi_{t_i^n}\psi_{t_j^n}(M_u+N_u)^{(a,b,c)}(M_v+N_v)^{(a,b,c)}\big].
\end{aligned}\tag{61}$$

In what follows, it does not matter to the analysis which particular subinterval of Δ^i_u, Δ^j_v, Δ^n_i or Δ^n_j is present in the terms. Hence we will use the notation

$$\begin{aligned}
&\Delta_{u^*}=\Delta^i_{uL},\Delta^i_{uR}\quad\text{or}\quad\Delta^i_u,\qquad \Delta_{v^*}=\Delta^j_{vL},\Delta^j_{vR}\quad\text{or}\quad\Delta^j_v,\\
&\Delta_{i^*}=\big[t_i^n,s_u^{k,i}\big]\quad\text{or}\quad\big[s_{u+1}^{k,i},t_{i+1}^n\big],\qquad \Delta_{j^*}=\big[t_j^n,s_v^{l,j}\big]\quad\text{or}\quad\big[s_{v+1}^{l,j},t_{j+1}^n\big],
\end{aligned}$$

and

$$R\begin{pmatrix}\Delta_{u^*}\\ \Delta_{v^*}\end{pmatrix}:=\langle\mathbb{1}_{\Delta_u^*},\mathbb{1}_{\Delta_v^*}\rangle_{\mathcal{H}_1}=R\begin{pmatrix}a_1 & a_2\\ b_1 & b_2\end{pmatrix},$$

where $[a_1, a_2] = \Delta^i_{u^L}, \Delta^i_{u^R}$ or Δ^i_u, and $[b_1, b_2] = \Delta^j_{v^L}, \Delta^j_{v^R}$ or Δ^j_v. $R\binom{\Delta_{u^*}}{\Delta_{i^*}}$, $R\binom{\Delta_{u^*}}{\Delta_{j^*}}$, $R\binom{\Delta_{u^*}}{\Delta_{u^*}}$, $R\binom{\Delta_{v^*}}{\Delta_{i^*}}$, $R\binom{\Delta_{v^*}}{\Delta_{j^*}}$ and $R\binom{\Delta_{v^*}}{\Delta_{v^*}}$ can be defined in the same manner, and we have the bounds

$$\begin{aligned}
&\left|R\begin{pmatrix}\Delta_{u^*}\\ \Delta_{v^*}\end{pmatrix}\right|, \left|R\begin{pmatrix}\Delta_{u^*}\\ \Delta_{u^*}\end{pmatrix}\right|, \left|R\begin{pmatrix}\Delta_{u^*}\\ \Delta_{i^*}\end{pmatrix}\right|, \left|R\begin{pmatrix}\Delta_{u^*}\\ \Delta_{j^*}\end{pmatrix}\right| \le \left\|R(\Delta_{u^*}, \cdot)\right\|_{q\text{-var};[0,T]}\\
&\quad = \left\|\Phi(\mathbb{1}_{\Delta_{u^*}})\right\|_{q\text{-var};[0,T]} \le \frac{C}{2^{\frac{n+k}{\rho}}},\\
&\left|R\begin{pmatrix}\Delta_{u^*}\\ \Delta_{v^*}\end{pmatrix}\right|, \left|R\begin{pmatrix}\Delta_{v^*}\\ \Delta_{v^*}\end{pmatrix}\right|, \left|R\begin{pmatrix}\Delta_{v^*}\\ \Delta_{i^*}\end{pmatrix}\right|, \left|R\begin{pmatrix}\Delta_{v^*}\\ \Delta_{j^*}\end{pmatrix}\right| \le \left\|R(\Delta_{v^*}, \cdot)\right\|_{q\text{-var};[0,T]}\\
&\quad = \left\|\Phi(\mathbb{1}_{\Delta_{v^*}})\right\|_{q\text{-var};[0,T]} \le \frac{C}{2^{\frac{n+l}{\rho}}}.
\end{aligned} \tag{62}$$

Using the notation

$$R_{\Delta^i_u\times\Delta^j_v} := \left|R\begin{pmatrix}s^{k,i}_{2u} & s^{k,i}_{2u+1}\\ s^{l,j}_{2v} & s^{l,j}_{2v+1}\end{pmatrix}\right| + \left|R\begin{pmatrix}s^{k,i}_{2u+1} & s^{k,i}_{2u+2}\\ s^{l,j}_{2v} & s^{l,j}_{2v+1}\end{pmatrix}\right| + \left|R\begin{pmatrix}s^{k,i}_{2u} & s^{k,i}_{2u+1}\\ s^{l,j}_{2v+1} & s^{l,j}_{2v+2}\end{pmatrix}\right| + \left|R\begin{pmatrix}s^{k,i}_{2u+1} & s^{k,i}_{2u+2}\\ s^{l,j}_{2v+1} & s^{l,j}_{2v+2}\end{pmatrix}\right|,$$

note that

$$\sum_{i,j=0}^{2^n-1}\sum_{u=0}^{2^k-1}\sum_{v=0}^{2^l-1}\left|R\begin{pmatrix}\Delta_{u^*}\\ \Delta_{v^*}\end{pmatrix}\right|^\rho \le \sum_{i,j=0}^{2^n-1}\sum_{u=0}^{2^k-1}\sum_{v=0}^{2^l-1} R^\rho_{\Delta^i_u\times\Delta^j_v} \le 4^\rho \|R\|^\rho_{\rho\text{-var};[0,T]^2}$$

for all $k, l \in \mathbb{N}$.

For $k = 1, \dots 6$, let y_k denote a, b or c. Returning to (61), we see that the last line can be expanded to include terms of the type $M_u^{(a,b,c)} M_v^{(a,b,c)}$:

$$\mathbb{E}\big[\psi_{t_i^n}\psi_{t_j^n} X^{(y_1)}_{\Delta_{u^*}} X^{(y_2)}_{\Delta_{u^*}} X^{(y_3)}_{\Delta_{i^*}} X^{(y_4)}_{\Delta_{v^*}} X^{(y_5)}_{\Delta_{v^*}} X^{(y_6)}_{\Delta_{j^*}}\big], \tag{63}$$

terms coming from $N_u^{(a,b,c)} N_v^{(a,b,c)}$:

$$\mathbb{E}\big[\psi_{t_i^n}\psi_{t_j^n} X^{(y_1)}_{\Delta_{u^*}} X^{(y_2)}_{\Delta_{u^*}} X^{(y_3)}_{\Delta_{u^*}} X^{(y_4)}_{\Delta_{v^*}} X^{(y_5)}_{\Delta_{v^*}} X^{(y_6)}_{\Delta_{v^*}}\big], \tag{64}$$

and terms arising from $M_u^{(a,b,c)} N_v^{(a,b,c)}$:

$$\mathbb{E}\big[\psi_{t_i^n}\psi_{t_j^n} X^{(y_1)}_{\Delta_{u^*}} X^{(y_2)}_{\Delta_{u^*}} X^{(y_3)}_{\Delta_{i^*}} X^{(y_4)}_{\Delta_{v^*}} X^{(y_5)}_{\Delta_{v^*}} X^{(y_6)}_{\Delta_{v^*}}\big]. \tag{65}$$

To account for the remaining $N_u^{(a,b,c)} M_v^{(a,b,c)}$ terms, we simply swap u and v in the third case. Note also that with our short-hand notation, as an example, $X^{(y_1)}_{\Delta_{u^*}}$ may not be equal to $X^{(y_2)}_{\Delta_{u^*}}$ even if $y_1 = y_2$ since Δ_{u^*} may be one of several intervals.

Since M_u is anti-symmetric with respect to $X_{\Delta_{u^L}}$ and $X_{\Delta_{u^R}}$, we can assume that $y_1 \ne y_2$ in (63) and (65), and $y_4 \ne y_5$ in (63). In each of the three cases, we will use $I_1(h_k)$ to denote $X^{(y_k)}$ for $k = 1, \dots, 6$; for example in (63), $h_1 := \mathbb{1}^{(y_1)}_{\Delta^*_u}$ and $I_1(h_1) = X^{(y_1)}_{\Delta_{u^*}}$. Now applying Lemma 4.7, we have

$$\begin{aligned}
\mathbb{E}\left[\psi_{t_i^n}\psi_{t_j^n}\prod_{k=1}^6 I_1(h_k)\right] &= \mathbb{E}\big[\mathcal{D}^6_{h_1,h_2,h_3,h_4,h_5,h_6}\psi_{t_i^n}\psi_{t_j^n}\big]\\
&\quad + \sum_{\sigma\in\mathcal{S}_6} C_{\sigma,1}A_{\sigma,1} + C_{\sigma,2}A_{\sigma,2} + C_{\sigma,3}A_{\sigma,3},
\end{aligned}$$

where we recall that

$$A_{\sigma,1} := \mathbb{E}\big[\mathcal{D}^4_{h_{\sigma(1)},h_{\sigma(2)},h_{\sigma(3)},h_{\sigma(4)}}\psi_{t_i^n}\psi_{t_j^n}\big]\langle h_{\sigma(5)}, h_{\sigma(6)}\rangle_{\mathcal{H}_1^d},$$

$$A_{\sigma,2} := \mathbb{E}\big[\mathcal{D}^2_{h_{\sigma(1)},h_{\sigma(2)}}\psi_{t_i^n}\psi_{t_j^n}\big]\langle h_{\sigma(3)}, h_{\sigma(4)}\rangle_{\mathcal{H}_1^d}\langle h_{\sigma(5)}, h_{\sigma(6)}\rangle_{\mathcal{H}_1^d},$$
$$A_{\sigma,3} := \mathbb{E}[\psi_{t_i^n}\psi_{t_j^n}]\langle h_{\sigma(1)}, h_{\sigma(2)}\rangle_{\mathcal{H}_1^d}\langle h_{\sigma(3)}, h_{\sigma(4)}\rangle_{\mathcal{H}_1^d}\langle h_{\sigma(5)}, h_{\sigma(6)}\rangle_{\mathcal{H}_1^d}.$$

We will show that for each of these terms, the sum over all the sub-intervals converges to zero as $n \to \infty$.

For the first term, from (54) we have

$$\mathbb{E}\big[\mathcal{D}^6_{h_1,h_2,h_3,h_4,h_5,h_6}\psi_{t_i^n}\psi_{t_j^n}\big] \le C\prod_{i=1}^{6}\big\|\Phi(h_i)\big\|_{q\text{-var};[0,T]}.$$

Looking at each of the three types of terms (63), (64) and (65), we see that at least two of the h_i's are $\mathbb{1}_{\Delta_{u^*}}$, and another two of the h_i's are $\mathbb{1}_{\Delta_{v^*}}$. Thus we obtain

$$\begin{aligned}\sum_{i,j=0}^{2^n-1}\sum_{k,l=1}^{m}\sum_{u=0}^{2^k-1}\sum_{v=0}^{2^l-1}\mathbb{E}\big[\mathcal{D}^6_{h_1,h_2,h_3,h_4,h_5,h_6}\psi_{t_i^n}\psi_{t_j^n}\big] &\le C\sum_{i,j=0}^{2^n-1}\sum_{k,l=1}^{m}\sum_{u=0}^{2^k-1}\sum_{v=0}^{2^l-1}\frac{1}{2^{(n+k)\frac{2}{\rho}}}\frac{1}{2^{(n+l)\frac{2}{\rho}}}\\ &\le C\sum_{i,j=0}^{2^n-1}\frac{1}{2^{2n(\frac{2}{\rho})}}\sum_{k,l=1}^{\infty}\frac{1}{2^{k(\frac{2}{\rho}-1)}}\frac{1}{2^{l(\frac{2}{\rho}-1)}}\\ &\le \frac{C}{2^{2n(\frac{2}{\rho}-1)}} \to 0\end{aligned}$$

since $\rho < 2$.

For the $A_{\sigma,1}$ terms, we have two cases:

(i) $\langle h_{\sigma(5)}, h_{\sigma(6)}\rangle_{\mathcal{H}_1^d} = R\binom{\Delta_{u^*}}{\Delta_{v^*}}$:

In all three types of terms (63), (64) and (65), at least one of $\{h_{\sigma(1)}, h_{\sigma(2)}, h_{\sigma(3)}, h_{\sigma(4)}\}$ equals $\mathbb{1}_{\Delta_{u^*}}$, and another one in the set equals $\mathbb{1}_{\Delta_{v^*}}$. Applying the bounds in (62), we get

$$\begin{aligned}&\sum_{i,j=0}^{2^n-1}\sum_{k,l=1}^{m}\sum_{u=0}^{2^k-1}\sum_{v=0}^{2^l-1}\mathbb{E}\big[\mathcal{D}^4_{h_{\sigma(1)},h_{\sigma(2)},h_{\sigma(3)},h_{\sigma(4)}}\psi_{t_i^n}\psi_{t_j^n}\big]\langle h_{\sigma(5)}, h_{\sigma(6)}\rangle_{\mathcal{H}_1^d},\\ &\quad\le C\sum_{i,j=0}^{2^n-1}\sum_{k,l=1}^{m}\sum_{u=0}^{2^k-1}\sum_{v=0}^{2^l-1}\prod_{r=1}^{4}\big\|\Phi(h_{\sigma(r)})\big\|_{q\text{-var};[0,T]}\big|\langle h_{\sigma(5)}, h_{\sigma(6)}\rangle_{\mathcal{H}_1^d}\big|\\ &\quad\le C\sum_{i,j=0}^{2^n-1}\sum_{k,l=1}^{m}\sum_{u=0}^{2^k-1}\sum_{v=0}^{2^l-1}2^{\frac{-(n+k)}{\rho}}2^{\frac{-(n+l)}{\rho}}\left|R\begin{pmatrix}\Delta_{u^*}\\ \Delta_{v^*}\end{pmatrix}\right|\\ &\quad\le C\sum_{k,l=1}^{m}\left(\sum_{i,j=0}^{2^n-1}\sum_{u=0}^{2^k-1}\sum_{v=0}^{2^l-1}R\begin{pmatrix}\Delta_{u^*}\\ \Delta_{v^*}\end{pmatrix}^{\rho}\right)^{\frac{1}{\rho}}\left(\sum_{i,j=0}^{2^n-1}\sum_{u=0}^{2^k-1}\sum_{v=0}^{2^l-1}2^{-(n+k)(\frac{\rho'}{\rho})}2^{-(n+l)(\frac{\rho'}{\rho})}\right)^{\frac{1}{\rho'}}\\ &\quad\le C2^{-2n(\frac{1}{\rho}-\frac{1}{\rho'})}\|R\|_{\rho\text{-var};[0,T]^2}\sum_{k,l=1}^{m}2^{-k(\frac{1}{\rho}-\frac{1}{\rho'})}2^{-l(\frac{1}{\rho}-\frac{1}{\rho'})}\\ &\quad\le C2^{-2n(\frac{2}{\rho}-1)}\|R\|_{\rho\text{-var};[0,T]^2}\sum_{k,l=1}^{\infty}2^{-k(\frac{2}{\rho}-1)}2^{-l(\frac{2}{\rho}-1)} \to 0.\end{aligned}$$

(ii) $\langle h_{\sigma(5)}, h_{\sigma(6)}\rangle_{\mathcal{H}_1^d} \ne R\binom{\Delta_{u^*}}{\Delta_{v^*}}$:

We will go through each of the three types of terms (63), (64) and (65) to count the number of quantities with increments Δ_{u^*} or Δ_{v^*}, which yield the factors $2^{\frac{-(n+k)}{\rho}}$ and $2^{\frac{-(n+l)}{\rho}}$ respectively.

(a) $M_u^{(a,b,c)}M_v^{(a,b,c)}$ terms:

We have five possibilities:

$$\langle h_{\sigma(5)}, h_{\sigma(6)}\rangle_{\mathcal{H}_1^d} = R\begin{pmatrix}\Delta_{u^*}\\ \Delta_{i^*}\end{pmatrix}, R\begin{pmatrix}\Delta_{u^*}\\ \Delta_{j^*}\end{pmatrix}, R\begin{pmatrix}\Delta_{v^*}\\ \Delta_{i^*}\end{pmatrix}, R\begin{pmatrix}\Delta_{v^*}\\ \Delta_{j^*}\end{pmatrix} \text{ or } R\begin{pmatrix}\Delta_{i^*}\\ \Delta_{j^*}\end{pmatrix}; \tag{66}$$

we need not consider the cases $R\binom{\Delta_{u^*}}{\Delta_{u^*}}$ or $R\binom{\Delta_{v^*}}{\Delta_{v^*}}$ since $y_1 \neq y_2$ and $y_4 \neq y_5$ in (63).

If $\langle h_{\sigma(5)}, h_{\sigma(6)}\rangle_{\mathcal{H}_1^d}$ is equal to either of the first two quantities on the right of (66), then one of $\{h_{\sigma(1)}, h_{\sigma(2)}, h_{\sigma(3)}, h_{\sigma(4)}\}$ must be equal to $\mathbb{1}_{\Delta_{u^*}}$ and another two in the set must be equal to $\mathbb{1}_{\Delta_{v^*}}$. If $\langle h_{\sigma(5)}, h_{\sigma(6)}\rangle_{\mathcal{H}_1^d}$ is equal to the third or the fourth quantity in (66), we have the same count with u and v switched.

If $\langle h_{\sigma(5)}, h_{\sigma(6)}\rangle_{\mathcal{H}_1^d} = R\binom{\Delta_{i^*}}{\Delta_{j^*}}$, then without loss of generality,

$$h_{\sigma(1)} \text{ and } h_{\sigma(2)} = \mathbb{1}_{\Delta_{u^*}}, h_{\sigma(3)} \text{ and } h_{\sigma(4)} = \mathbb{1}_{\Delta_{v^*}}.$$

(b) $N_u^{(a,b,c)} N_v^{(a,b,c)}$ terms:

If $\langle h_{\sigma(5)}, h_{\sigma(6)}\rangle_{\mathcal{H}_1^d} = R\binom{\Delta_{u^*}}{\Delta_{u^*}}$, then one of $\{h_{\sigma(1)}, h_{\sigma(2)}, h_{\sigma(3)}, h_{\sigma(4)}\}$ must equal $\mathbb{1}_{\Delta_{u^*}}$ and another two in the set must equal $\mathbb{1}_{\Delta_{v^*}}$. By switching u and v, we can resolve the only other case $\langle h_{\sigma(5)}, h_{\sigma(6)}\rangle_{\mathcal{H}_1^d} = R\binom{\Delta_{v^*}}{\Delta_{v^*}}$ similarly.

(c) $M_u^{(a,b,c)} N_v^{(a,b,c)}$ terms:

There are only three possibilities

$$\langle h_{\sigma(5)}, h_{\sigma(6)}\rangle_{\mathcal{H}_1^d} = R\begin{pmatrix}\Delta_{u^*}\\ \Delta_{i^*}\end{pmatrix}, R\begin{pmatrix}\Delta_{v^*}\\ \Delta_{v^*}\end{pmatrix} \text{ or } R\begin{pmatrix}\Delta_{v^*}\\ \Delta_{i^*}\end{pmatrix},$$

and we need not consider the case $R\binom{\Delta_{u^*}}{\Delta_{u^*}}$ since $y_1 \neq y_2$ in (65). If $\langle h_{\sigma(5)}, h_{\sigma(6)}\rangle_{\mathcal{H}_1^d}$ is equal to $R\binom{\Delta_{u^*}}{\Delta_{i^*}}$, then one of $\{h_{\sigma(1)}, h_{\sigma(2)}, h_{\sigma(3)}, h_{\sigma(4)}\}$ must be equal to $\mathbb{1}_{\Delta_{u^*}}$ and another two in the set must be equal to $\mathbb{1}_{\Delta_{v^*}}$. If $\langle h_{\sigma(5)}, h_{\sigma(6)}\rangle_{\mathcal{H}_1^d}$ is equal to the second or third quantity, the same count applies with u and v switched.

Thus in each case, applying the bounds in (62) yields

$$\prod_{r=1}^{4} \left\|\Phi(h_{\sigma(r)})\right\|_{q\text{-var};[0,T]} \left|\langle h_{\sigma(5)}, h_{\sigma(6)}\rangle_{\mathcal{H}_1^d}\right| \leq C 2^{\frac{-2(n+k)}{\rho}} 2^{\frac{-2(n+l)}{\rho}}, \tag{67}$$

which gives us

$$\begin{aligned}
&\sum_{i,j=0}^{2^n-1}\sum_{k,l=1}^{m}\sum_{u=0}^{2^k-1}\sum_{v=0}^{2l-1} \mathbb{E}\big[\mathcal{D}^4_{h_{\sigma(1)},h_{\sigma(2)},h_{\sigma(3)},h_{\sigma(4)}} \psi_{t_i^n}\psi_{t_j^n}\big]\langle h_{\sigma(5)}, h_{\sigma(6)}\rangle_{\mathcal{H}_1^d} \\
&\quad \leq C \sum_{i,j=0}^{2^n-1}\sum_{k,l=1}^{m}\sum_{u=0}^{2^k-1}\sum_{v=0}^{2l-1} 2^{-(n+k)\frac{2}{\rho}} 2^{-(n+l)\frac{2}{\rho}} \\
&\quad \leq C 2^{-2n(\frac{2}{\rho}-1)} \sum_{k,l=1}^{\infty} 2^{-k(\frac{2}{\rho}-1)} 2^{-l(\frac{2}{\rho}-1)} \to 0.
\end{aligned}$$

For the $A_{\sigma,2}$ terms, when we consider $\langle h_{\sigma(3)}, h_{\sigma(4)}\rangle_{\mathcal{H}_1^d}$ and $\langle h_{\sigma(5)}, h_{\sigma(6)}\rangle_{\mathcal{H}_1^d}$, we have three cases: either both, one, or none of them are equal to $R\binom{\Delta_{u^*}}{\Delta_{v^*}}$.

(i) $\langle h_{\sigma(3)}, h_{\sigma(4)}\rangle_{\mathcal{H}_1^d}$ and $\langle h_{\sigma(5)}, h_{\sigma(6)}\rangle_{\mathcal{H}_1^d} = R\binom{\Delta_{u^*}}{\Delta_{v^*}}$:

(Note that this does not imply that they are equal to one another since Δ_{u^*} and Δ_{v^*} can be one of several intervals.) Observe that

$$\left|\langle h_{\sigma(3)}, h_{\sigma(4)}\rangle_{\mathcal{H}_1^d}\right| \leq \left|R\begin{pmatrix}\Delta_{u^*}\\ \Delta_{v^*}\end{pmatrix}\right|^{\frac{\rho}{2}} \left|R\begin{pmatrix}\Delta_{u^*}\\ \Delta_{v^*}\end{pmatrix}\right|^{1-\frac{\rho}{2}} \leq C R^{\frac{\rho}{2}}_{\Delta_u^i \times \Delta_v^j} 2^{\frac{-(n+k)}{\rho}(1-\frac{\rho}{2})},$$

and

$$\left|\langle h_{\sigma(5)}, h_{\sigma(6)}\rangle_{\mathcal{H}_1^d}\right| \leq \left|R\begin{pmatrix}\Delta_{u^*}\\ \Delta_{v^*}\end{pmatrix}\right|^{\frac{\rho}{2}} \left|R\begin{pmatrix}\Delta_{u^*}\\ \Delta_{v^*}\end{pmatrix}\right|^{1-\frac{\rho}{2}} \leq C R^{\frac{\rho}{2}}_{\Delta_u^i\times\Delta_v^j} 2^{\frac{-(n+l)}{\rho}(1-\frac{\rho}{2})}.$$

Thus we obtain

$$\sum_{i,j=0}^{2^n-1}\sum_{k,l=1}^{m}\sum_{u=0}^{2^k-1}\sum_{v=0}^{2^l-1}\mathbb{E}\big[\mathcal{D}^2_{h_{\sigma(1)},h_{\sigma(2)}}\psi_{t_i^n}\psi_{t_j^n}\big]\langle h_{\sigma(3)}, h_{\sigma(4)}\rangle_{\mathcal{H}_1^d}\langle h_{\sigma(5)}, h_{\sigma(6)}\rangle_{\mathcal{H}_1^d}$$

$$\leq C\sum_{k,l=1}^{m} 2^{\frac{-(n+k)}{\rho}(1-\frac{\rho}{2})}2^{\frac{-(n+l)}{\rho}(1-\frac{\rho}{2})}\sum_{i,j=0}^{2^n-1}\sum_{u=0}^{2^k-1}\sum_{v=0}^{2^l-1} R^{\rho}_{\Delta_u^i\times\Delta_v^j}$$

$$\leq C 2^{-2n(\frac{1}{\rho}-\frac{1}{2})}\sum_{k,l=1}^{\infty} 2^{-k(\frac{1}{\rho}-\frac{1}{2})}2^{-l(\frac{1}{\rho}-\frac{1}{2})}\|R\|^{\rho}_{\rho\text{-var};[0,T]^2}\to 0,$$

since $\frac{1}{\rho}-\frac{1}{2}>0$.

(ii) WLOG, assume $\langle h_{\sigma(3)}, h_{\sigma(4)}\rangle_{\mathcal{H}_1^d} = R\big(\begin{smallmatrix}\Delta_{u^*}\\ \Delta_{v^*}\end{smallmatrix}\big)$, $\langle h_{\sigma(5)}, h_{\sigma(6)}\rangle_{\mathcal{H}_1^d} \neq R\big(\begin{smallmatrix}\Delta_{u^*}\\ \Delta_{v^*}\end{smallmatrix}\big)$:

As before, we will use the bounds in (62) to show that

$$\left\|\Phi(h_{\sigma(1)})\right\|_{q\text{-var};[0,T]}\left\|\Phi(h_{\sigma(2)})\right\|_{q\text{-var};[0,T]}\left|\langle h_{\sigma(5)}, h_{\sigma(6)}\rangle_{\mathcal{H}_1^d}\right| \leq C2^{\frac{-(n+k)}{\rho}}2^{\frac{-(n+l)}{\rho}}. \tag{68}$$

(a) $M_u^{(a,b,c)}M_v^{(a,b,c)}$ terms:

Again we have five possibilities,

$$\langle h_{\sigma(5)}, h_{\sigma(6)}\rangle_{\mathcal{H}_1^d} = R\begin{pmatrix}\Delta_{u^*}\\ \Delta_{i^*}\end{pmatrix}, R\begin{pmatrix}\Delta_{u^*}\\ \Delta_{j^*}\end{pmatrix}, R\begin{pmatrix}\Delta_{v^*}\\ \Delta_{i^*}\end{pmatrix}, R\begin{pmatrix}\Delta_{v^*}\\ \Delta_{j^*}\end{pmatrix} \text{ or } R\begin{pmatrix}\Delta_{i^*}\\ \Delta_{j^*}\end{pmatrix}, \tag{69}$$

and we need not consider the cases $R\big(\begin{smallmatrix}\Delta_{u^*}\\ \Delta_{u^*}\end{smallmatrix}\big)$ or $R\big(\begin{smallmatrix}\Delta_{v^*}\\ \Delta_{v^*}\end{smallmatrix}\big)$ since $y_1\neq y_2$ and $y_4\neq y_5$ in (63).

If $\langle h_{\sigma(5)}, h_{\sigma(6)}\rangle_{\mathcal{H}_1^d}$ is equal to either of the first two quantities on the right of (69), then either $h_{\sigma(1)}$ or $h_{\sigma(2)}$ is equal to $\mathbb{1}_{\Delta_{v^*}}$. If $\langle h_{\sigma(5)}, h_{\sigma(6)}\rangle_{\mathcal{H}_1^d}$ is equal to the third or fourth quantity, then either $h_{\sigma(1)}$ or $h_{\sigma(2)}$ is equal to $\mathbb{1}_{\Delta_{u^*}}$.

If $\langle h_{\sigma(5)}, h_{\sigma(6)}\rangle_{\mathcal{H}_1^d} = R\big(\begin{smallmatrix}\Delta_{i^*}\\ \Delta_{j^*}\end{smallmatrix}\big)$, then we must have $h_{\sigma(1)}=\mathbb{1}_{\Delta_{u^*}}$ and $h_{\sigma(2)}=\mathbb{1}_{\Delta_{v^*}}$, or vice versa.

(b) $N_u^{(a,b,c)}N_v^{(a,b,c)}$ terms:

If $\langle h_{\sigma(5)}, h_{\sigma(6)}\rangle_{\mathcal{H}_1^d} = R\big(\begin{smallmatrix}\Delta_{u^*}\\ \Delta_{u^*}\end{smallmatrix}\big)$ (resp. $R\big(\begin{smallmatrix}\Delta_{v^*}\\ \Delta_{v^*}\end{smallmatrix}\big)$), then both $h_{\sigma(1)}$ and $h_{\sigma(2)}$ must be equal to $\mathbb{1}_{\Delta_{v^*}}$ (resp. $\mathbb{1}_{\Delta_{u^*}}$).

(c) $M_u^{(a,b,c)}N_v^{(a,b,c)}$ terms:

There are only three possibilities,

$$\langle h_{\sigma(5)}, h_{\sigma(6)}\rangle_{\mathcal{H}_1^d} = R\begin{pmatrix}\Delta_{u^*}\\ \Delta_{i^*}\end{pmatrix}, R\begin{pmatrix}\Delta_{v^*}\\ \Delta_{v^*}\end{pmatrix} \text{ or } R\begin{pmatrix}\Delta_{v^*}\\ \Delta_{i^*}\end{pmatrix},$$

and we need not consider the case $R\big(\begin{smallmatrix}\Delta_{u^*}\\ \Delta_{u^*}\end{smallmatrix}\big)$ since $y_1\neq y_2$ in (65). If $\langle h_{\sigma(5)}, h_{\sigma(6)}\rangle_{\mathcal{H}_1^d}$ is equal to $R\big(\begin{smallmatrix}\Delta_{u^*}\\ \Delta_{i^*}\end{smallmatrix}\big)$, then both $h_{\sigma(1)}$ and $h_{\sigma(2)}$ are equal to $\mathbb{1}_{\Delta_{v^*}}$. If $\langle h_{\sigma(5)}, h_{\sigma(6)}\rangle_{\mathcal{H}_1^d}$ is equal to $R\big(\begin{smallmatrix}\Delta_{v^*}\\ \Delta_{v^*}\end{smallmatrix}\big)$ or $R\big(\begin{smallmatrix}\Delta_{v^*}\\ \Delta_{i^*}\end{smallmatrix}\big)$, then either $h_{\sigma(1)}$ or $h_{\sigma(2)}$ is equal to $\mathbb{1}_{\Delta_{u^*}}$.

Thus we obtain

$$\sum_{i,j=0}^{2^n-1}\sum_{k,l=1}^{m}\sum_{u=0}^{2^k-1}\sum_{v=0}^{2^l-1}\mathbb{E}\big[\mathcal{D}^2_{h_{\sigma(1)},h_{\sigma(2)}}\psi_{t_i^n}\psi_{t_j^n}\big]\langle h_{\sigma(3)}, h_{\sigma(4)}\rangle_{\mathcal{H}_1^d}\langle h_{\sigma(5)}, h_{\sigma(6)}\rangle_{\mathcal{H}_1^d}$$

$$\leq \sum_{k,l=1}^{m}\left(\sum_{i,j=0}^{2^n-1}\sum_{u=0}^{2^k-1}\sum_{v=0}^{2^l-1}\left|R\begin{pmatrix}\Delta_{u^*}\\ \Delta_{v^*}\end{pmatrix}\right|^{\rho}\right)^{\frac{1}{\rho}}\left(\sum_{i,j=0}^{2^n-1}\sum_{u=0}^{2^k-1}\sum_{v=0}^{2^l-1}2^{-(n+k)(\frac{\rho'}{\rho})}2^{-(n+l)(\frac{\rho'}{\rho})}\right)^{\frac{1}{\rho'}}$$

$$\le \|R\|_{\rho\text{-var};[0,T]^2} \sum_{k,l=1}^{m} 2^{-2n(\frac{1}{\rho}-\frac{1}{\rho'})} 2^{-k(\frac{1}{\rho}-\frac{1}{\rho'})} 2^{-l(\frac{1}{\rho}-\frac{1}{\rho'})}$$

$$\le C\|R\|_{\rho\text{-var};[0,T]^2} 2^{-2n(\frac{2}{\rho}-1)} \to 0.$$

(iii) $\langle h_{\sigma(3)}, h_{\sigma(4)}\rangle_{\mathcal{H}_1^d}$ and $\langle h_{\sigma(5)}, h_{\sigma(6)}\rangle_{\mathcal{H}_1^d} \neq R\binom{\Delta_{u^*}}{\Delta_{v^*}}$:

We will show that

$$\|\Phi(h_{\sigma(1)})\|_{q\text{-var};[0,T]} \|\Phi(h_{\sigma(2)})\|_{q\text{-var};[0,T]} \left|\langle h_{\sigma(3)}, h_{\sigma(4)}\rangle_{\mathcal{H}_1^d}\right| \left|\langle h_{\sigma(5)}, h_{\sigma(6)}\rangle_{\mathcal{H}_1^d}\right| \tag{70}$$

is bounded above by $C2^{\frac{-2(n+k)}{\rho}} 2^{\frac{-2(n+l)}{\rho}}$, which gives us

$$\sum_{i,j=0}^{2^n-1} \sum_{k,l=1}^{m} \sum_{u=0}^{2^k-1} \sum_{v=0}^{2^l-1} \mathbb{E}\big[\mathcal{D}^2_{h_{\sigma(1)},h_{\sigma(2)}} \psi_{t_i^n} \psi_{t_j^n}\big] \langle h_{\sigma(3)}, h_{\sigma(4)}\rangle_{\mathcal{H}_1^d} \langle h_{\sigma(5)}, h_{\sigma(6)}\rangle_{\mathcal{H}_1^d}$$

$$\le C \sum_{i,j=0}^{2^n-1} \sum_{k,l=1}^{m} \sum_{u=0}^{2^k-1} \sum_{v=0}^{2^l-1} 2^{-(n+k)\frac{2}{\rho}} 2^{-(n+l)\frac{2}{\rho}}$$

$$\le C 2^{-2n(\frac{2}{\rho}-1)} \sum_{k,l=1}^{m} 2^{-k(\frac{2}{\rho}-1)} 2^{-l(\frac{2}{\rho}-1)} \to 0.$$

(a) $M_u^{(a,b,c)} M_v^{(a,b,c)}$ terms:

Note that in this scenario, neither $h_{\sigma(1)}$ nor $h_{\sigma(2)}$ can be equal to $\mathbb{1}_{\Delta_{i^*}}$ or $\mathbb{1}_{\Delta_{j^*}}$, so we essentially have two cases.

If $h_{\sigma(1)}$ and $h_{\sigma(2)} = \mathbb{1}_{\Delta_{u^*}}$, we must have

$$\langle h_{\sigma(3)}, h_{\sigma(4)}\rangle_{\mathcal{H}_1^d} = R\begin{pmatrix}\Delta_{v^*}\\ \Delta_{i^*}\end{pmatrix} \quad \text{and} \quad \langle h_{\sigma(5)}, h_{\sigma(6)}\rangle_{\mathcal{H}_1^d} = R\begin{pmatrix}\Delta_{v^*}\\ \Delta_{j^*}\end{pmatrix},$$

or vice versa. (The case $h_{\sigma(1)}$ and $h_{\sigma(2)} = \mathbb{1}_{\Delta_{v^*}}$ can be resolved similarly by swapping u and v.)

If instead we have $h_{\sigma(1)} = \mathbb{1}_{\Delta_{u^*}}$ and $h_{\sigma(2)} = \mathbb{1}_{\Delta_{v^*}}$, or vice versa, then without loss of generality, it must be the case that

$$\langle h_{\sigma(3)}, h_{\sigma(4)}\rangle_{\mathcal{H}_1^d} = R\begin{pmatrix}\Delta_{u^*}\\ \Delta_{i^*}\end{pmatrix} \text{ or } R\begin{pmatrix}\Delta_{u^*}\\ \Delta_{j^*}\end{pmatrix}, \qquad \langle h_{\sigma(5)}, h_{\sigma(6)}\rangle_{\mathcal{H}_1^d} = R\begin{pmatrix}\Delta_{v^*}\\ \Delta_{j^*}\end{pmatrix} \text{ or } R\begin{pmatrix}\Delta_{v^*}\\ \Delta_{i^*}\end{pmatrix}.$$

(b) $N_u^{(a,b,c)} N_v^{(a,b,c)}$ terms:

Without loss of generality, we have

$$h_{\sigma(1)} = \mathbb{1}_{\Delta_{u^*}}, \qquad h_{\sigma(2)} = \mathbb{1}_{\Delta_{v^*}}, \qquad \langle h_{\sigma(3)}, h_{\sigma(4)}\rangle_{\mathcal{H}_1^d} = R\begin{pmatrix}\Delta_{u^*}\\ \Delta_{u^*}\end{pmatrix}, \qquad \langle h_{\sigma(5)}, h_{\sigma(6)}\rangle_{\mathcal{H}_1^d} = R\begin{pmatrix}\Delta_{v^*}\\ \Delta_{v^*}\end{pmatrix}.$$

(c) $M_u^{(a,b,c)} N_v^{(a,b,c)}$ terms:

Without loss of generality, either

$$h_{\sigma(1)}, h_{\sigma(2)} = \mathbb{1}_{\Delta_{u^*}}, \qquad \langle h_{\sigma(3)}, h_{\sigma(4)}\rangle_{\mathcal{H}_1^d} = R\begin{pmatrix}\Delta_{v^*}\\ \Delta_{i^*}\end{pmatrix} \quad \text{and} \quad \langle h_{\sigma(5)}, h_{\sigma(6)}\rangle_{\mathcal{H}_1^d} = R\begin{pmatrix}\Delta_{v^*}\\ \Delta_{v^*}\end{pmatrix},$$

or

$$h_{\sigma(1)} = \mathbb{1}_{\Delta_{u^*}}, \qquad h_{\sigma(2)} = \mathbb{1}_{\Delta_{v^*}},$$

$$\langle h_{\sigma(3)}, h_{\sigma(4)}\rangle_{\mathcal{H}_1^d} = R\begin{pmatrix}\Delta_{u^*}\\ \Delta_{i^*}\end{pmatrix} \quad \text{and} \quad \langle h_{\sigma(5)}, h_{\sigma(6)}\rangle_{\mathcal{H}_1^d} = R\begin{pmatrix}\Delta_{v^*}\\ \Delta_{v^*}\end{pmatrix}.$$

For the $A_{\sigma,3}$ terms, when we consider the three inner-products $\langle h_{\sigma(1)}, h_{\sigma(2)}\rangle_{\mathcal{H}_1^d}$, $\langle h_{\sigma(3)}, h_{\sigma(4)}\rangle_{\mathcal{H}_1^d}$ and $\langle h_{\sigma(5)}, h_{\sigma(6)}\rangle_{\mathcal{H}_1^d}$, we have two cases: either one of them is equal to $R\binom{\Delta_{u^*}}{\Delta_{v^*}}$, or two or more of them are. Observe that it is not possible for none of them to equal $R\binom{\Delta_{u^*}}{\Delta_{v^*}}$.

(i) If two or more of the inner-products are equal to $R\binom{\Delta_{u^*}}{\Delta_{v^*}}$, then we can use the same computation as in the first case for the $A_{\sigma,2}$ terms to show that

$$\left|\langle h_{\sigma(1)}, h_{\sigma(2)}\rangle_{\mathcal{H}_1^d}\right|\left|\langle h_{\sigma(3)}, h_{\sigma(4)}\rangle_{\mathcal{H}_1^d}\right|\left|\langle h_{\sigma(5)}, h_{\sigma(6)}\rangle_{\mathcal{H}_1^d}\right| \le C R^{\rho}_{\Delta_u^i\times\Delta_v^j} 2^{\frac{-(n+k)}{\rho}(1-\frac{\rho}{2})} 2^{\frac{-(n+l)}{\rho}(1-\frac{\rho}{2})},$$

and this gives us

$$\sum_{i,j=0}^{2^n-1}\sum_{k,l=1}^{m}\sum_{u=0}^{2^k-1}\sum_{v=0}^{2^l-1}\mathbb{E}[\psi_{t_i^n}\psi_{t_j^n}]\langle h_{\sigma(1)}, h_{\sigma(2)}\rangle_{\mathcal{H}_1^d}\langle h_{\sigma(3)}, h_{\sigma(4)}\rangle_{\mathcal{H}_1^d}\langle h_{\sigma(5)}, h_{\sigma(6)}\rangle_{\mathcal{H}_1^d}$$
$$\le C 2^{-2n(\frac{1}{\rho}-\frac{1}{2})}\sum_{k,l=1}^{\infty} 2^{-k(\frac{1}{\rho}-\frac{1}{2})}2^{-l(\frac{1}{\rho}-\frac{1}{2})}\|R\|^{\rho}_{\rho\text{-var};[0,T]^2} \to 0.$$

(ii) Assume that $\langle h_{\sigma(1)}, h_{\sigma(2)}\rangle_{\mathcal{H}_1^d} = R\binom{\Delta_{u^*}}{\Delta_{v^*}}$, and $\langle h_{\sigma(3)}, h_{\sigma(4)}\rangle_{\mathcal{H}_1^d}, \langle h_{\sigma(5)}, h_{\sigma(6)}\rangle_{\mathcal{H}_1^d} \ne R\binom{\Delta_{u^*}}{\Delta_{v^*}}$.
Then without loss of generality, we have:

(a) $M_u^{(a,b,c)}M_v^{(a,b,c)}$ terms:

$$\langle h_{\sigma(3)}, h_{\sigma(4)}\rangle_{\mathcal{H}_1^d} = R\begin{pmatrix}\Delta_{u^*}\\ \Delta_{i^*}\end{pmatrix} \text{ or } R\begin{pmatrix}\Delta_{u^*}\\ \Delta_{j^*}\end{pmatrix}, \qquad \langle h_{\sigma(5)}, h_{\sigma(6)}\rangle_{\mathcal{H}_d^1} = R\begin{pmatrix}\Delta_{v^*}\\ \Delta_{i^*}\end{pmatrix} \text{ or } R\begin{pmatrix}\Delta_{v^*}\\ \Delta_{j^*}\end{pmatrix}.$$

(b) $N_u^{(a,b,c)}N_v^{(a,b,c)}$ terms:

$$\langle h_{\sigma(3)}, h_{\sigma(4)}\rangle_{\mathcal{H}_1^d} = R\begin{pmatrix}\Delta_{u^*}\\ \Delta_{u^*}\end{pmatrix}, \qquad \langle h_{\sigma(5)}, h_{\sigma(6)}\rangle_{\mathcal{H}_1^d} = R\begin{pmatrix}\Delta_{v^*}\\ \Delta_{v^*}\end{pmatrix}.$$

(c) $M_u^{(a,b,c)}N_v^{(a,b,c)}$ terms:

$$\langle h_{\sigma(3)}, h_{\sigma(4)}\rangle_{\mathcal{H}_1^d} = R\begin{pmatrix}\Delta_{u^*}\\ \Delta_{i^*}\end{pmatrix}, \qquad \langle h_{\sigma(5)}, h_{\sigma(6)}\rangle_{\mathcal{H}_1^d} = R\begin{pmatrix}\Delta_{v^*}\\ \Delta_{v^*}\end{pmatrix}.$$

In each case, applying the bounds in (62) gives us

$$\left|\langle h_{\sigma(3)}, h_{\sigma(4)}\rangle_{\mathcal{H}_1^d}\right|\left|\langle h_{\sigma(5)}, h_{\sigma(6)}\rangle_{\mathcal{H}_1^d}\right| \le C 2^{\frac{-(n+k)}{\rho}} 2^{\frac{-(n+l)}{\rho}},$$

which in turn yields

$$\sum_{i,j=0}^{2^n-1}\sum_{k,l=1}^{m}\sum_{u=0}^{2^k-1}\sum_{v=0}^{2^l-1}\mathbb{E}[\psi_{t_i^n}\psi_{t_j^n}]\langle h_{\sigma(1)}, h_{\sigma(2)}\rangle_{\mathcal{H}_1^d}\langle h_{\sigma(3)}, h_{\sigma(4)}\rangle_{\mathcal{H}_1^d}\langle h_{\sigma(5)}, h_{\sigma(6)}\rangle_{\mathcal{H}_1^d}$$
$$\le \sum_{k,l=1}^{m}\left(\sum_{i,j=0}^{2^n-1}\sum_{u=0}^{2^k-1}\sum_{v=0}^{2^l-1}\left|R\begin{pmatrix}\Delta_{u^*}\\ \Delta_{v^*}\end{pmatrix}\right|^{\rho}\right)^{\frac{1}{\rho}}\left(\sum_{i,j=0}^{2^n-1}\sum_{u=0}^{2^k-1}\sum_{v=0}^{2^l-1}2^{-(n+k)(\frac{\rho'}{\rho})}2^{-(n+l)(\frac{\rho'}{\rho})}\right)^{\frac{1}{\rho'}}$$
$$\le \|R\|_{\rho\text{-var};[0,T]^2}\sum_{k,l=1}^{m}2^{-2n(\frac{1}{\rho}-\frac{1}{\rho'})}2^{-k(\frac{1}{\rho}-\frac{1}{\rho'})}2^{-l(\frac{1}{\rho}-\frac{1}{\rho'})}$$
$$\le C\|R\|_{\rho\text{-var};[0,T]^2}2^{-2n(\frac{2}{\rho}-1)} \to 0.$$

□

Corollary 4.9. *For $2 \le p < 4$, let $Y \in \mathcal{C}^{p\text{-var}}([0,T];\mathcal{L}(\mathbb{R}^d;\mathbb{R}^m))$ denote the path-level solution to*

$$\mathrm{d}Y_t = V(Y_t)\circ \mathrm{d}\mathbf{X}_t, \qquad Y_0 = y_0,$$

where $\mathbf{X} \in \mathcal{C}^{0,p\text{-}var}([0,T]; G^{\lfloor p \rfloor}(\mathbb{R}^d))$ *satisfies Condition* 1 *and its covariance function satisfies*

$$\left\| R(t,\cdot) - R(s,\cdot) \right\|_{q\text{-}var;[0,T]} \le C|t-s|^{\frac{1}{\rho}}, \quad \forall s,t \in [0,T].$$

Then if $V \in \mathcal{C}_b^{\lfloor p \rfloor + 4}(\mathbb{R}^{md}; \mathbb{R}^{md} \otimes \mathbb{R}^d)$, *we have*

$$\lim_{\|\pi(n)\| \to 0} \left\| \sum_i V(Y_{t_i^n}) \left(\mathbf{X}^2_{t_i^n, t_{i+1}^n} - \frac{1}{2} \sigma^2 (t_i^n, t_{i+1}^n) \mathcal{I}_d \right) \right\|_{L^2(\Omega)} = 0. \tag{71}$$

Furthermore, if $3 \le p < 4$ *and* $V \in \mathcal{C}_b^9(\mathbb{R}^{md}; \mathbb{R}^{md} \otimes \mathbb{R}^d)$, *we have*

$$\lim_{\|\pi(n)\| \to 0} \left\| \sum_i \nabla V(Y_{t_i^n}) \big(V(Y_{t_i^n}) \big) \big(\mathbf{X}^3_{t_i^n, t_{i+1}^n} \big) \right\|_{L^2(\Omega)} = 0. \tag{72}$$

Proof. We have to show that bounds (91) and (92) of Proposition 5.1 in [7] are satisfied with

$$\psi_t = \big[V(Y_t) \big]_j \in \mathbb{R}^d \otimes \mathbb{R}^d, \quad j = 1, \dots, m,$$

to show (71). Similarly, proving bounds (53) and (54) in Proposition 4.8 are satisfied with

$$\psi_t = \big[\nabla V(Y_t) \big(V(Y_t) \big) \big]_j \in \mathbb{R}^d \otimes \mathbb{R}^d \otimes \mathbb{R}^d, \quad j = 1, \dots, m,$$

will yield (72). (53), as well as (91) in Proposition 5.1 of [7], is trivially true since $V \in \mathcal{C}_b^1$. To show that the bounds hold for the higher Malliavin derivatives, recall Proposition 4.6, which states that almost surely we have

$$\left\| \mathcal{D}^n_{h_1,\dots,h_n} Y_\cdot \right\|_\infty \le P_{d(n)} \big(\|\mathbf{X}\|_{p\text{-var};[0,T]}, \exp\big(C N^{\mathbf{X}}_{1;[0,T]} \big) \big) \prod_{i=1}^n \left\| \Phi(h_i) \right\|_{q\text{-var};[0,T]}. \tag{73}$$

As both $\|\mathbf{X}\|_{p\text{-var};[0,T]}$ and $\exp(C N^{\mathbf{X}}_{1;[0,T]})$ belong to $\bigcap_{r>0} L^r(\Omega)$, we have

$$\left\| \mathcal{D}^n_{h_1,\dots,h_n} Y_t \right\|_{L^r(\Omega)} \le C_{n,q} \prod_{i=1}^n \left\| \Phi(h_i) \right\|_{q\text{-var};[0,T]} \tag{74}$$

for any $r > 0$. Now we simply use the product and chain rule of Malliavin differentiation in conjunction with the fact that V has bounded derivatives up to the appropriate order. □

5. Correction formula

We are now ready to prove the main result of the paper. As before, $\pi(n) := \{t_i^n\}$, $t_i^n := \frac{iT}{2^n}$, denotes the sequence of dyadic partitions on $[0,T]$.

5.1. *Main theorem*

Theorem 5.1. *For* $3 \le p < 4$, *let* $Y \in \mathcal{C}^{p\text{-}var}([0,T]; \mathcal{L}(\mathbb{R}^d; \mathbb{R}^m))$ *denote the path-level solution to*

$$\mathrm{d}Y_t = V(Y_t) \circ \mathrm{d}\mathbf{X}_t, \qquad Y_0 = y_0,$$

where $V \in \mathcal{C}_b^9(\mathbb{R}^{md}; \mathbb{R}^{md} \otimes \mathbb{R}^d)$, *and* $\mathbf{X} \in \mathcal{C}^{0,p\text{-}var}([0,T]; G^{\lfloor p \rfloor}(\mathbb{R}^d))$ *is a Volterra process which satisfies Condition* 1, *and whose kernel satisfies Condition* 2 *with* $\alpha < \frac{1}{p}$. *Furthermore, we assume the covariance function satisfies*

$$\left\| R(t,\cdot) - R(s,\cdot) \right\|_{q\text{-}var;[0,T]} \le C|t-s|^{\frac{1}{\rho}}, \tag{75}$$

for all $s,t \in [0,T]$, *and* $\|R(\cdot)\|_{q\text{-}var;[0,T]} < \infty$. *Then almost surely, we have*

$$\int_0^T Y_t \circ \mathrm{d}\mathbf{X}_t = \int_0^T Y_t \, \mathrm{d}X_t + \sum_{j=1}^m \left(\frac{1}{2} \int_0^T \mathrm{tr}\big[V(Y_s) \big]_j \, \mathrm{d}R(s) + U_T^{(j)} \right) e_j, \tag{76}$$

where for $j = 1, \ldots, m$, $U_T^{(j)}$ *is the limit in* $L^2(\Omega)$ *of*

$$\sum_i \int_0^{t_i^n} \mathrm{tr}\big[J^{\mathbf{X}}_{t_i^n \leftarrow s} V(Y_s) - V(Y_{t_i^n})\big]_j R\big(\Delta_i^n, \mathrm{d}s\big) \tag{77}$$

along the dyadic partitions $\{t_i^n\}$ *of* $[0, T]$.

Proof. Using bounds (41), (45) together with the integrability of $\mathbf{X}$, we can apply dominated convergence theorem to (40) in Theorem 4.2 to show that $\int_0^T Y_t \circ \mathrm{d}\mathbf{X}_t$ is the $L^2(\Omega)$ limit of

$$\lim_{n\to\infty} \sum_i Y_{t_i^n}(X_{t_i^n, t_{i+1}^n}) + V(Y_{t_i^n})\big(\mathbf{X}^2_{t_i^n, t_{i+1}^n}\big) + \nabla V(Y_{t_i^n})\big(V(Y_{t_i^n})\big)\big(\mathbf{X}^3_{t_i^n, t_{i+1}^n}\big).$$

Now applying Proposition 3.13 in conjunction with Corollary 4.9 gives us

$$\int_0^T Y_t \, \mathrm{d}X_t = \lim_{n\to\infty} \sum_i \left[Y_{t_i^n}(X_{t_i^n, t_{i+1}^n}) - \sum_{j=1}^m \left(\int_0^{t_i^n} \mathrm{tr}\big[J^{\mathbf{X}}_{t_i^n \leftarrow s} V(Y_s)\big]_j R\big(\Delta_i^n, \mathrm{d}s\big)\right) e_j + A_i \right],$$

where the limit is also in $L^2(\Omega)$ and

$$A_i := V(Y_{t_i^n})\left(\big(\mathbf{X}^2_{t_i^n, t_{i+1}^n}\big) - \frac{1}{2}\sigma^2\big(t_i^n, t_{i+1}^n\big)\mathcal{I}_d \right) + \nabla V(Y_{t_i^n})\big(V(Y_{t_i^n})\big)\big(\mathbf{X}^3_{t_i^n, t_{i+1}^n}\big).$$

Following the procedure in Theorem 6.1 of [7], subtracting the two integrals and re-balancing the terms gives us

$$\begin{aligned}
&\int_0^T Y_t \circ \mathrm{d}\mathbf{X}_t - \int_0^T Y_t \, \mathrm{d}X_t \\
&= \sum_{j=1}^m \left(\lim_{n\to\infty} \sum_i \int_0^{t_i^n} \mathrm{tr}\big[J^{\mathbf{X}}_{t_i^n \leftarrow s} V(Y_s)\big]_j R\big(\Delta_i^n, \mathrm{d}s\big) + \frac{1}{2}\sigma^2\big(t_i^n, t_{i+1}^n\big)\, \mathrm{tr}\big[V(Y_{t_i^n})\big]_j \right) e_j \\
&= \sum_{j=1}^m \left(\lim_{n\to\infty} \sum_i \int_0^{t_i^n} \mathrm{tr}\big[J^{\mathbf{X}}_{t_i^n \leftarrow s} V(Y_s) - V(Y_{t_i^n})\big]_j R\big(\Delta_i^n, \mathrm{d}s\big) \right. \\
&\quad \left. + \frac{1}{2} \lim_{n\to\infty} \sum_i \mathrm{tr}\big[V(Y_{t_i^n})\big]_j \Big(R\big(t_{i+1}^n, t_{i+1}^n\big) - R\big(t_i^n, t_i^n\big)\Big) \right) e_j.
\end{aligned} \tag{78}$$

The second term in the last line of the expression above is dominated by

$$C \big\| V(Y_\cdot) \big\|_{p\text{-var};[0,T]} \big\| R(\cdot) \big\|_{q\text{-var};[0,T]}$$

by Young's inequality, and thus converges in $L^2(\Omega)$ to

$$\frac{1}{2} \int_0^T \mathrm{tr}\big[V(Y_s)\big]_j \, \mathrm{d}R(s).$$

This in turn guarantees the convergence of the first term in $L^2(\Omega)$ to the random variable $U_T^{(j)}$. Now extracting an almost sure subsequence allows us to equate both sides of (76) almost surely, and the proof is thus complete. □

In the more regular case $2 \le p < 3$, we can be more precise in identifying the second term (77) when X is a Volterra process (cf. Theorem 6.1 in [7]).

Proposition 5.2. *For* $2 \le p < 3$, *let* $Y \in \mathcal{C}^{p\text{-}var}([0, T]; \mathcal{L}(\mathbb{R}^d; \mathbb{R}^m))$ *denote the path-level solution to* $\mathbf{X}$

$$\mathrm{d}Y_t = V(Y_t) \circ \mathrm{d}\mathbf{X}_t, \qquad Y_0 = y_0,$$

where $\mathbf{X} \in \mathcal{C}^{0,p\text{-}var}([0,T]; G^{\lfloor p \rfloor}(\mathbb{R}^d))$ *is a Volterra process satisfying Condition 1 with some* $\rho \in [1, \frac{3}{2})$, *and whose kernel satisfies Condition 2 with* $\alpha < \frac{1}{2p}$.

Furthermore, assume that $V \in \mathcal{C}_b^6(\mathbb{R}^{md}; \mathbb{R}^{md} \otimes \mathbb{R}^d)$, *and the covariance function satisfies*

$$\left\| R(t,\cdot) - R(s,\cdot) \right\|_{\rho\text{-}var;[0,T]} \le C|t-s|^{\frac{1}{\rho}}, \tag{79}$$

for all $s,t \in [0,T]$, *and* $\|R(\cdot)\|_{q\text{-}var;[0,T]} < \infty$. *Then almost surely we have*

$$\int_0^T Y_t \circ \mathrm{d}\mathbf{X}_t = \int_0^T Y_t \,\mathrm{d}X_t + Z_T,$$

where the correction term is given by

$$\begin{aligned} Z_T^{(j)} &= \frac{1}{2}\int_0^T \operatorname{tr}\big[V(Y_s)\big]_j \,\mathrm{d}R(s) + \int_{[0,T]^2} h_j(s,t)\,\mathrm{d}R(s,t) \\ &= \frac{1}{2}\int_0^T \operatorname{tr}\big[V(Y_s)\big]_j \,\mathrm{d}R(s) + \int_0^T \mathcal{K}^* \otimes \mathcal{K}^* h_j(r,r)\,\mathrm{d}r, \quad j = 1,\dots,m. \end{aligned} \tag{80}$$

with

$$h_j(s,t) := \mathbb{1}_{[0,t)}(s)\operatorname{tr}\big[J^{\mathbf{X}}_{t\leftarrow s} V(Y_s) - V(Y_t)\big]_j.$$

Proof. Under the conditions of the theorem, we can invoke Theorem 6.1 from [7] to obtain the first line of (80). To obtain the second line, we will use Proposition 4.3 from [23], which states that if $\phi : [0,T]^2 \to \mathbb{R}$ is a λ-Hölder bi-continuous function (one that satisfies Definition 3.7 without necessarily satisfying (17)) with $\lambda > 2\alpha$, then

$$\int_{[0,T]^2} \phi(s,t)\,\mathrm{d}R(s,t) = \int_0^T \mathcal{K}^* \otimes \mathcal{K}^* \phi(r,r)\,\mathrm{d}r. \tag{81}$$

Thus, the proof is complete once we show that $h_j(s,t)$ is $\frac{1}{p}$-Hölder bi-continuous for all $j = 1,\dots,m$ since $\frac{1}{p} > 2\alpha$. Using the fact that

$$\tilde{h}_j(s,t) := \operatorname{tr}\big[J^{\mathbf{X}}_{t\leftarrow s} V(Y_s) - V(Y_t)\big]_j$$

is $\frac{1}{p}$-Hölder bi-continuous, we have, assuming $v_2 > v_1$ without loss of generality,

$$\begin{aligned} \big|h_j(u,v_2) - h_j(u,v_1)\big| &= \big|\tilde{h}_j(u, v_2 \vee u) - \tilde{h}_j(u, v_1 \vee u)\big|, \quad u, v_1, v_2 \in [0,T], \\ &\le C_1 |v_2 \vee u - v_1 \vee u|^{\frac{1}{p}} \\ &\le C_2 |v_2 - v_1|^{\frac{1}{p}}, \end{aligned}$$

and similarly,

$$\big|h_j(u_2,v) - h_j(u_1,v)\big| \le C|u_2 - u_1|^{\frac{1}{p}}, \quad v, u_1, u_2 \in [0,T].$$

□

Remark 5.3. In the case $3 \le p < 4$, due to the lack of complementary regularity, we cannot apply the standard criterion to ensure the 2D integral exists even though $\mathbb{1}_{[0,t)}(s)\operatorname{tr}[J^{\mathbf{X}}_{t\leftarrow s} V(Y_s) - V(Y_t)]_j$ is continuous almost surely on $[0,T]^2$. Furthermore, although the integrand is strongly $\frac{1}{p}$-Hölder bi-continuous away from the diagonal, one can check that in general, (17) fails at the diagonal, which means that we cannot employ (81) from Proposition 4.3 in [23] (it can also be verified that there would be insufficient Hölder regularity in the weaker sense). Hence, we can only show convergence in $L^2(\Omega)$ rather than almost surely. The question of whether the second part of the correction term can be identified as a proper 2D Young–Stieltjes integral requires further investigation.

An interesting special case of Theorem 5.1 is when the vector fields defining the RDE commute. In this situation the U_T terms in the correction formula (76) disappear.

Corollary 5.4. *Under the conditions of Theorem* 5.1, *if in addition the vector fields commute*, *i.e.* $[V_i, V_j] = 0$ *for all* $i, j = 1, \ldots, d$, *then*

$$\int_0^T Y_t \circ \mathrm{d}\mathbf{X}_t = \int_0^T Y_t \,\mathrm{d}X_t + \frac{1}{2}\sum_{j=1}^m \left(\int_0^T \mathrm{tr}\big[V(Y_s)\big]_j \,\mathrm{d}R(s)\right) e_j,$$

Proof. For any vector field $W \in \mathcal{C}^1(\mathbb{R}^{md}; \mathbb{R}^{md})$, we have

$$\left(J_t^{\mathbf{X}}\right)^{-1} W(Y_t) = W(y_0) + \sum_{i=1}^d \int_0^t \left(J_s^{\mathbf{X}}\right)^{-1} [V_i, W] \circ \mathrm{d}\mathbf{X}_s^{(i)},$$

which can be computed using the RDEs satisfied by Y and $(J^{\mathbf{X}})^{-1}$, cf. Chapter 20 (Section 4.2) in [16]. Hence, if the V_i's commute, then each V_i is invariant under the flow of Y, and we have

$$J_{t\leftarrow s}^{\mathbf{X}} V(Y_s) = V(Y_t), \quad 0 \le s < t \le T. \qquad \square$$

5.2. *Applications of the correction formula*

We present applications of the main theorem to two important special cases. The first is to fractional Brownian motion in the regime $H > \frac{1}{4}$. The second is to use the commuting case discussed in Corollary 5.4 to obtain Itô formulas for Gaussian processes.

Theorem 5.5 (Correction formula fBM, $H > \frac{1}{4}$). *For* $1 \le p < 4$, *let* $Y \in \mathcal{C}^{p\text{-}var}([0, T]; \mathcal{L}(\mathbb{R}^d; \mathbb{R}^m))$ *denote the path-level solution to*

$$\mathrm{d}Y_t = V(Y_t) \circ \mathrm{d}\mathbf{X}_t, \quad Y_0 = y_0,$$

where we assume that $V \in \mathcal{C}_b^k(\mathbb{R}^{md}; \mathbb{R}^{md} \otimes \mathbb{R}^d)$, *with*

$$k = \begin{cases} 2, & 1 \le p < 2, \\ 6, & 2 \le p < 3, \\ 9, & 3 \le p < 4, \end{cases} \tag{82}$$

and $\mathbf{X} \in \mathcal{C}^{0,p\text{-}var}([0, T]; G^{\lfloor p \rfloor}(\mathbb{R}^d))$ *is the geometric rough path constructed from the limit of the piecewise-linear approximations of standard fractional Brownian motion with Hurst parameter* $H > \frac{1}{4}$. *Then almost surely, we have*

$$\int_0^T Y_t \circ \mathrm{d}\mathbf{X}_t = \int_0^T Y_t \,\mathrm{d}X_t + Z_T,$$

where the correction term $Z_T = (Z_T^{(1)}, \ldots, Z_T^{(m)})$ *is given by*

$$\begin{aligned} Z_T^{(j)} &= H \int_0^T \mathrm{tr}\big[V(Y_s)\big]_j s^{2H-1} \,\mathrm{d}s + \int_{[0,T]^2} h_j(s, t) \,\mathrm{d}R(s, t), \quad j = 1, \ldots, m, \\ &= H \int_0^T \mathrm{tr}\big[V(Y_s)\big]_j s^{2H-1} \,\mathrm{d}s + \int_0^T \mathcal{K}^* \otimes \mathcal{K}^* h_j(r, r) \,\mathrm{d}r, \quad \left(\textit{when } \frac{1}{3} < H \le \frac{1}{2}\right), \end{aligned} \tag{83}$$

with

$$h_j(s, t) := \mathbb{1}_{[0,t)}(s)\, \mathrm{tr}\big[J_{t\leftarrow s}^{\mathbf{X}} V(Y_s) - V(Y_t)\big]_j, \quad j = 1, \ldots, m.$$

Remark 5.6. For simplicity, we use the same notation for the second term of $Z_T^{(j)}$ for all $H > \frac{1}{4}$, with the understanding that it denotes the $L^2(\Omega)$ limit of (77) when $\frac{1}{4} < H \le \frac{1}{3}$.

Proof. The proof rests entirely on the following Proposition 5.7, which tells us that fractional Brownian motion fulfills all the requirements needed to apply Theorem 6.1 of [7] when $H > \frac{1}{3}$, and Theorem 5.1 when $\frac{1}{4} < H \le \frac{1}{3}$. $\square$

Proposition 5.7. *Let B^H be standard fractional Brownian motion with Hurst index $H \in (\frac{1}{4}, 1)$, and let K be the square-integrable kernel associated with it* [11]. *We have*:

(i) *For any $p > \frac{1}{H}$ the sample paths of B^H are almost surely $\frac{1}{p}$-Hölder continuous. Furthermore, there exists a geometric rough path $\mathbf{X} \in \mathcal{C}^{0,p\text{-}var}([0,T]; G^{\lfloor p \rfloor}(\mathbb{R}^d))$ which is the $d_{p\text{-}var}$-limit of the paths $S_{\lfloor p \rfloor}(X^\pi)$ as $||\pi|| \to 0$.*

(ii) *B^H satisfies Condition 1 with $\rho = \frac{1}{2H}$ and*

$$q = \begin{cases} \frac{2\rho}{\rho+1} & \text{if } \frac{1}{4} < H \le \frac{1}{3}, \\ \rho \vee 1 & \text{if } \frac{1}{3} < H \le 1. \end{cases}$$

(iii) *If $\frac{1}{4} < H \le \frac{1}{2}$, then the kernel K satisfies Condition 2 with $\alpha = \frac{1}{2} - H$.*

(iv) *The covariance function, $R(s,t) := \frac{1}{2}(s^{2H} + t^{2H} - |t-s|^{2H})$, of B^H satisfies*:

   (a) $\|R(t,\cdot) - R(s,\cdot)\|_{q\text{-}var;[0,T]} \le C|t-s|^{\frac{1}{\rho}}$, *if $\frac{1}{4} < H \le \frac{1}{2}$,*

   (b) *$R(t) = t^{2H}$ is of bounded variation and thus of finite q-variation for any $q \ge 1$.*

Proof. The proof uses standard arguments and can be found in Proposition 1.2.5 of [22]. □

We now show that we can recover Itô's formulas; cf. [2,3,20,28] and [29].

Theorem 5.8 (Itô formulas for Gaussian processes). *For $1 \le p < 4$, let $\mathbf{X} \in \mathcal{C}^{0,p\text{-}var}([0,T]; G^{\lfloor p \rfloor}(\mathbb{R}^d))$ satisfy Condition 1. Depending on p, we further impose the following conditions*:

(i) $1 \le p < 2$: *$\sigma^2(s,t) \le C|t-s|^\theta$ for some $\theta > 1$ and $\|R(\cdot)\|_{q\text{-}var;[0,T]} < \infty$.*

(ii) $2 \le p < 3$: *The covariance function satisfies*

$$\left\| R(t,\cdot) - R(s,\cdot) \right\|_{q\text{-}var;[0,T]} \le C|t-s|^{\frac{1}{\rho}}, \tag{84}$$

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

(iii) $3 \le p < 4$: *$\mathbf{X}$ is a Volterra process whose kernel satisfies Condition 2 with $\alpha < \frac{1}{p}$. Furthermore, its covariance function satisfies* (84) *and $\|R(\cdot)\|_{q\text{-}var;[0,T]} < \infty$.*

Then almost surely, for $f \in \mathcal{C}_b^{k+2}(\mathbb{R}^d; \mathbb{R})$, k defined as in (82), *we have*

$$f(X_T) - f(0) = \int_0^T \langle \nabla f(X_t), \circ \mathrm{d}\mathbf{X}_t \rangle = \int_0^T \langle \nabla f(X_t), \mathrm{d}X_t \rangle + \frac{1}{2}\int_0^T \Delta f(X_t)\, \mathrm{d}R(t).$$

Proof. Let $Y_t = (Y_t^{(1)}, \dots, Y_t^{(2d)})$ denote the augmented process

$$\left(\frac{\partial f}{\partial e_1}(X_t), \dots, \frac{\partial f}{\partial e_d}(X_t), X_t^{(1)}, \dots X_t^{(d)} \right).$$

In this case Y satisfies the RDE

$$\mathrm{d}Y_t = V(Y_t) \circ \mathrm{d}\mathbf{X}_t, \; Y_0 = (y_0, 0),$$

where $V(Y) \in \mathbb{R}^{2d} \otimes \mathbb{R}^d$ is represented by the $2d$-by-d matrix

$$V(Y_t) = \begin{bmatrix} \nabla^2 f(Y_t) \\ \hline \mathcal{I}_d \end{bmatrix},$$

and note that $\nabla^2 f(Y_t) = \nabla^2 f(Y_t^{(d+1)}, \dots, Y_t^{(2d)}) = \nabla^2 f(X_t)$.

Now one can check that $[V_i, V_j] = 0$ for all $i, j = 1, \dots, d$, apply Corollary 5.4, and project back onto the first d components to obtain the result. □

Acknowledgements

The first-named author acknowledges the support of EPSRC Grants EP/M00516X/1 and EP/S026347/1.

References

[1] E. Alòs, O. Mazet and D. Nualart. Stochastic calculus with respect to fractional Brownian motion with Hurst parameter lesser than $\frac{1}{2}$. *Stochastic Process. Appl.* **86** (1) (2000) 121–139. MR1741199 https://doi.org/10.1016/S0304-4149(99)00089-7

[2] E. Alòs, O. Mazet and D. Nualart. Stochastic calculus with respect to Gaussian processes. *Ann. Probab.* **29** (2) (2001) 766–801. MR1849177 https://doi.org/10.1214/aop/1008956692

[3] P. Carmona, L. Coutin and G. Montseny. Stochastic integration with respect to fractional Brownian motion. *Ann. Inst. Henri Poincaré Probab. Stat.* **39** (1) (2003) 27–68. MR1959841 https://doi.org/10.1016/S0246-0203(02)01111-1

[4] T. Cass and P. Friz. Malliavin calculus and rough paths. *Bull. Sci. Math.* **135** (6–7) (2011) 542–556. MR2838089 https://doi.org/10.1016/j.bulsci.2011.07.003

[5] T. Cass, P. Friz and N. Victoir. Non-degeneracy of Wiener functionals arising from rough differential equations. *Trans. Amer. Math. Soc.* **361** (6) (2009) 3359–3371. MR2485431 https://doi.org/10.1090/S0002-9947-09-04677-7

[6] T. Cass, M. Hairer, C. Litterer and S. Tindel. Smoothness of the density for solutions to Gaussian rough differential equations. *Ann. Probab.* **43** (1) (2015) 188–239. MR3298472 https://doi.org/10.1214/13-AOP896

[7] T. Cass and N. Lim. A Stratonovich–Skorohod integral formula for Gaussian rough paths. *Ann. Probab.* **47** (1) (2019) 1–60. MR3909965 https://doi.org/10.1214/18-AOP1254

[8] T. Cass, C. Litterer and T. Lyons. Integrability and tail estimates for Gaussian rough differential equations. *Ann. Probab.* **41** (4) (2013) 3026–3050. MR3112937 https://doi.org/10.1214/12-AOP821

[9] L. Coutin and Z. Qian. Stochastic analysis, rough path analysis and fractional Brownian motions. *Probab. Theory Related Fields* **122** (2002) 108–140. MR1883719 https://doi.org/10.1007/s004400100158

[10] L. Decreusefond. Stochastic integration with respect to Volterra processes. *Ann. Inst. Henri Poincaré Probab. Stat.* **41** (2) (2005) 123–149. MR2124078 https://doi.org/10.1016/j.anihpb.2004.03.004

[11] L. Decreusefond and A. S. Üstünel. Stochastic analysis of the fractional Brownian motion. *Potential Anal.* **10** (2) (1999) 177–214. MR1677455 https://doi.org/10.1023/A:1008634027843

[12] R. J. Elliott and J. Van Der Hoek. A general fractional white noise theory and applications to finance. *Math. Finance* **13** (2) (2003) 301–330. MR1967778 https://doi.org/10.1111/1467-9965.00018

[13] P. Friz, B. Gess, A. Gulisashvili and S. Riedel. The jain-monrad criterion for rough paths and applications. *Ann. Probab.* **44** (1) (2016) 684–738. MR3456349 https://doi.org/10.1214/14-AOP986

[14] P. Friz and M. Hairer. *A Course on Rough Paths*. Springer, Berlin, 2014. MR3289027 https://doi.org/10.1007/978-3-319-08332-2

[15] P. Friz and N. Victoir. Differential equations driven by Gaussian signals. *Ann. Inst. Henri Poincaré Probab. Stat.* **46** (2) (2010) 369–413. MR2667703 https://doi.org/10.1214/09-AIHP202

[16] P. Friz and N. Victoir. *Multidimensional Stochastic Processes as Rough Paths: Theory and Applications*, 1st edition. *Cambridge Studies in Advanced Mathematics* **120**. Cambridge University Press, Cambridge, 2010. MR2604669 https://doi.org/10.1017/CBO9780511845079

[17] M. Gubinelli. Controlling rough paths. *J. Funct. Anal.* **216** (1) (2004) 86–140. MR2091358 https://doi.org/10.1016/j.jfa.2004.01.002

[18] M. Gubinelli. Ramification of rough paths. *J. Differential Equations* **248** (4) (2010) 693–721. MR2578445 https://doi.org/10.1016/j.jde.2009.11.015

[19] M. Hairer and N. S. Pillai. Regularity of laws and ergodicity of hypoelliptic SDES driven by rough paths. *Ann. Probab.* **41** (4) (2013) 2544–2598. MR3112925 https://doi.org/10.1214/12-AOP777

[20] Y. Hu, M. Jolis and S. Tindel. On Stratonovich and Skorohod stochastic calculus for Gaussian processes. *Ann. Probab.* **41** (3A) (2013) 1656–1693. MR3098687 https://doi.org/10.1214/12-AOP751

[21] Y. Inahama. Malliavin differentiability of solutions of rough differential equations. *J. Funct. Anal.* **267** (5) (2014) 1566–1584. MR3229800 https://doi.org/10.1016/j.jfa.2014.06.011

[22] N. Lim. A Stratonovich–Skorohod integral formula for Gaussian rough paths. Ph.D. thesis, Imperial College, London, 2016.

[23] N. Lim Young–Stieltjes integrals with respect to Volterra covariance functions, 2018. Preprint. Available at arXiv:1806.02214.

[24] T. Lyons. Differential equations driven by rough signals. *Rev. Mat. Iberoam.* **14** (2) (1998) 215–310. MR1654527 https://doi.org/10.4171/RMI/240

[25] T. Lyons, M. Caruana and T. Lévy. *Differential Equations Driven by Rough Paths*. Springer, Berlin, 2007. MR2314753

[26] T. Lyons and Z. Qian. *System Control and Rough Paths*. *Oxford Mathematical Monographs*. Oxford University Press, London, 2003. MR2036784 https://doi.org/10.1093/acprof:oso/9780198506485.001.0001

[27] D. Nualart. *The Malliavin Calculus and Related Topics*, 2nd edition. *Probability and Its Applications*. Springer, Berlin, 2006. MR2200233

[28] D. Nualart and S. Ortiz-Latorre. *Multidimensional Wick–Itô Formula for Gaussian Processes*. *Stochastic Analysis, Stochastic Systems and Applications to Finance*. World Scientific, Singapore, 2011. MR2884564 https://doi.org/10.1142/9789814355711_0001

[29] N. Privault. Skorohod stochastic integration with respect to non-adapted processes on Wiener space. *Stoch. Stoch. Rep.* **65** (1998) 13–39. MR1708428 https://doi.org/10.1080/17442509808834172

[30] N. Towghi. Multidimensional extension of L.C. Young's inequality. *JIPAM. J. Inequal. Pure Appl. Math.* **4** (1) (2002) 1–7. MR1906391