

**Original citation:**

Riedel, Sebastian and Xu, Weijun. (2013) A simple proof of distance bounds for Gaussian rough paths. *Electronic Journal of Probability*, Volume 18. pp. 1-18.

**Permanent WRAP url:**

<http://wrap.warwick.ac.uk/61970>

**Copyright and reuse:**

The Warwick Research Archive Portal (WRAP) makes this work of researchers of the University of Warwick available open access under the following conditions.

This article is made available under the Creative Commons Attribution- 3.0 Unported (CC BY 3.0) license and may be reused according to the conditions of the license. For more details see <http://creativecommons.org/licenses/by/3.0/>

**A note on versions:**

The version presented in WRAP is the published version, or, version of record, and may be cited as it appears here.

For more information, please contact the WRAP Team at: [publications@warwick.ac.uk](mailto:publications@warwick.ac.uk)

warwick**publications**wrap

highlight your research

<http://wrap.warwick.ac.uk/>

## A simple proof of distance bounds for Gaussian rough paths

Sebastian Riedel\*      Weijun Xu<sup>†</sup>

### Abstract

We derive explicit distance bounds for Stratonovich iterated integrals along two Gaussian processes (also known as signatures of Gaussian rough paths) based on the regularity assumption of their covariance functions. Similar estimates have been obtained recently in [Friz-Riedel, AIHP, to appear]. One advantage of our argument is that we obtain the bound for the third level iterated integrals merely based on the first two levels, and this reflects the intrinsic nature of rough paths. Our estimates are sharp when both covariance functions have finite 1-variation, which includes a large class of Gaussian processes.

Two applications of our estimates are discussed. The first one gives the a.s. convergence rates for approximated solutions to rough differential equations driven by Gaussian processes. In the second example, we show how to recover the optimal time regularity for solutions of some rough SPDEs.

**Keywords:** Gaussian rough paths, iterated integrals, signatures.

**AMS MSC 2010:** 60G15; 60H05; 60H10; 60H35.

Submitted to EJP on October 22, 2012, final version accepted on December 7, 2013.

## 1 Introduction

The intersection between rough path theory and Gaussian processes has been an active research area in recent years ([5], [6], [10]). The central idea of rough paths, as realized by Lyons ([11]), is that the key properties needed for defining integration against an irregular path do not only come from the path itself, but from the path together with a sequence of iterated integrals along the path, namely

$$\mathbf{X}_{s,t}^n = \int_{s < u_1 < \dots < u_n < t} dX_{u_1} \otimes \dots \otimes dX_{u_n}. \tag{1.1}$$

In particular, Lyons extension theorem shows that for paths of finite  $p$ -variation, the first  $\lfloor p \rfloor$  levels iterated integrals determine all higher levels. For instance, if  $p = 1$ , the path has bounded variation and the higher iterated integrals coincide with the usual

---

\*Technische Universität Berlin, Germany. E-mail: riedel@math.tu-berlin.de

<sup>†</sup>University of Warwick, UK. E-mail: Weijun.Xu@warwick.ac.uk

Riemann-Stieltjes integrals. However, for  $p \geq 2$ , this is not true anymore and one has to say what the second (and possibly higher) order iterated integrals should be before they determine the whole rough path.

Lyons and Zeitouni ([15]) were the first to study iterated Wiener integrals in the sense of rough paths. They provide sharp exponential bounds on the iterated integrals of all levels by controlling the variation norm of the Lévy area. The case of more general Gaussian processes were studied by Friz and Victoir in [5] and [6]. They showed that if  $X$  is a Gaussian process with covariance of finite  $\rho$ -variation for some  $\rho \in [1, 2)$ , then its iterated integrals in the sense of (1.1) can be defined in a natural way and we can lift  $X$  to a Gaussian rough path  $\mathbf{X}$ .

In the recent work [4], Friz and the first author compared the two lift maps  $\mathbf{X}$  and  $\mathbf{Y}$  for the joint process  $(X, Y)$ . It was shown that their average distance in rough paths topology can be controlled by the value  $\sup_t |X_t - Y_t|_{L^2}^\zeta$  for some  $\zeta > 0$ , and a sharp quantitative estimate for  $\zeta$  was given. In particular, it was shown that considering both rough paths in a larger rough paths space (and therefore in a different topology) allows for larger choices of  $\zeta$ . Using this, the authors derived essentially optimal convergence rates for  $\mathbf{X}^\epsilon \rightarrow \mathbf{X}$  in rough paths topology when  $\epsilon \rightarrow 0$  where  $\mathbf{X}^\epsilon$  is a suitable approximation of  $\mathbf{X}$ .

In order to prove this result, sharp estimates of  $|\mathbf{X}_{s,t}^n - \mathbf{Y}_{s,t}^n|$  need to be calculated on every level  $n$ . Under the assumption  $\rho \in [1, \frac{3}{2})$ , the sample paths of  $\mathbf{X}$  and  $\mathbf{Y}$  are  $p$ -rough paths for any  $p > 2\rho$ , hence we can always choose  $p < 3$  and therefore the first two levels determine the entire rough path. Lyons' continuity theorem then suggests that one only needs to give sharp estimates on level 1 and 2; the estimates on the higher levels can be obtained from the lower levels through induction. On the other hand, interestingly, one additional level was estimated "by hand" in [4] before performing the induction. To understand the necessity of computing this additional term, let us note from [11] that the standard distance for two deterministic  $p$ -rough paths takes the form of the smallest constant  $C_n$  such that

$$|\mathbf{X}_{s,t}^n - \mathbf{Y}_{s,t}^n| \leq C_n \epsilon \omega(s, t)^{\frac{n}{p}}, \quad n = 1, \dots, [p]$$

holds for all  $s < t$  where  $\omega$  is a control function to be defined later. The exponent on the control for the next level is expected to be

$$\frac{n+1}{p} = \frac{[p]+1}{p} > 1, \tag{1.2}$$

so when one repeats Young's trick of dropping points in the induction argument (the key idea of the extension theorem), condition (1.2) will ensure that one can establish a maximal inequality for the next level. However, in the current problem where Gaussian randomness is involved, the  $L^2$  distance for the first  $[2\rho]$  iterated integrals takes the form

$$|\mathbf{X}_{s,t}^n - \mathbf{Y}_{s,t}^n|_{L^2} < C_n \epsilon \omega(s, t)^{\frac{1}{2\gamma} + \frac{n-1}{2\rho}}, \quad n = 1, 2, \quad \rho \in [1, \frac{3}{2}),$$

where  $\gamma$  might be much larger than  $\rho$ . Thus, the ' $n-1$ ' in the exponent leaves condition (1.2) unsatisfied, and one needs to compute the third level by hand before starting induction on  $n$ .

In this article, we resolve the difficulty by moving part of  $\epsilon$  to fill in the gap in the control so that the exponent for the third level control reaches 1. In this way, we obtain the third level estimate merely based on the first two levels, and it takes the form

$$|\mathbf{X}_{s,t}^3 - \mathbf{Y}_{s,t}^3|_{L^2} < C_3 \epsilon^\eta \omega(s, t),$$

where  $\eta \in (0, 1]$ , and its exact value depends on  $\gamma$  and  $\rho$ . We see that there is a  $1 - \eta$  reduction in the exponent of  $\epsilon$ , which is due to the fact that it is used to compensate the control exponent. This interplay between the rate exponent and the control exponent can be viewed as an analogy to the relationship between time and space regularities for solutions to SPDEs. We will make the above heuristic argument rigorous in section 4. We also refer to the recent work [14] for the situation of deterministic rough paths.

Our main theorem is the following.

**Theorem 1.1.** *Let  $(X, Y) = (X^1, Y^1, \dots, X^d, Y^d) : [0, T] \rightarrow \mathbb{R}^{d+d}$  be a centered Gaussian process on a probability space  $(\Omega, \mathcal{F}, P)$  where  $(X^i, Y^i)$  and  $(X^j, Y^j)$  are independent for  $i \neq j$ , with continuous sample paths and covariance function  $R_{(X,Y)} : [0, T]^2 \mapsto \mathbb{R}^{2d \times 2d}$ . Assume further that there is a  $\rho \in [1, \frac{3}{2})$  such that the  $\rho$ -variation of  $R_{(X,Y)}$  is bounded by a finite constant  $K$ . Choose  $\sigma > 2\rho$  and  $\beta < \min\{1 - \frac{2\rho}{\sigma}, 3 - 2\rho\}$ . Then for every  $N \in \mathbb{N}$  and  $q \geq 1$  there is a constant  $C = C(\beta, \sigma, \rho, K, q, N)$  such that*

$$|\varrho_{\sigma\text{-var}}^N(\mathbf{X}, \mathbf{Y})|_{L^q} \leq C \sup_{t \in [0, T]} |X_t - Y_t|_{L^2}^\beta.$$

The proof of this theorem will be postponed to section 4.2 after we have established all the estimates needed. We first give two remarks.

**Remark 1.2.** *We emphasize that the constant  $C$  in the above theorem depends on the process  $(X, Y)$  only through the parameters  $\rho$  and  $K$ .*

**Remark 1.3.** *For  $N = \lfloor \sigma \rfloor$ ,  $\varrho_{\sigma\text{-var}}^N$  denotes an inhomogeneous rough paths metric. See section 2 for the precise definition.*

Our paper is structured as follows. In section 2, we provide some important concepts and notations from rough path theory that are necessary for our problem. In section 3, we introduce the class of Gaussian processes which possess a lift to Gaussian rough paths and estimate the difference of two Gaussian rough paths on level one and two. Section 4 is devoted to the proof of the main theorem. We first obtain the third level estimate directly from the first two levels, which requires a technical extension of Lyons' continuity theorem, and justify the heuristic argument above rigorously. All higher level estimates are then obtained with the induction procedure in [11], and the claim of the main theorem follows. In section 5, we give two applications of our main theorem. The first one deals with convergence rates for Wong-Zakai approximations in the context of rough differential equations. The second example shows how to derive optimal time regularity for the solution of a modified stochastic heat equation seen as an evolution in rough paths space.

*Notations.* Throughout the paper,  $C, C_n, C_n(\rho, \gamma)$  will denote constants depending on certain parameters only, and their actual values may change from line to line.

## 2 Elements from Rough path theory

In this section, we introduce the concepts and definitions from rough path theory that are necessary for our current application. For a detailed account of the theory, we refer readers to [6], [12] and [13].

Fix the time interval  $[0, T]$ . For all  $s < t \in [0, T]$ , let  $\Delta_{s,t}$  denote the simplex

$$\{(u_1, u_2) \mid s \leq u_1 \leq u_2 \leq t\},$$

and we simply write  $\Delta$  for  $\Delta_{0,T}$ . In what follows, we will use  $x$  to denote an  $\mathbb{R}^d$ -valued path, and  $X$  to denote a stochastic process in  $\mathbb{R}^d$ , which is a Gaussian process in the current paper. For any integer  $N$ , let

$$T^N(\mathbb{R}^d) = \mathbb{R} \oplus \mathbb{R}^d \oplus \cdots \oplus (\mathbb{R}^d)^{\otimes N}$$

denote the truncated tensor algebra. The space of all continuous bounded variation paths  $x: [0, T] \rightarrow \mathbb{R}^d$  is denoted by  $C^{1-var}(\mathbb{R}^d)$ . For a path  $x \in C^{1-var}(\mathbb{R}^d)$ , we use the bold letter  $\mathbf{X}$  to denote its  $n$ -th level iterated tensor integral:

$$\mathbf{X}_{s,t}^n = \int_{s < u_1 < \cdots < u_n < t} dx_{u_1} \otimes \cdots \otimes dx_{u_n}.$$

The lift map  $S_N$  taking  $x$  to a  $T^N(\mathbb{R}^d)$ -valued path is defined by

$$S_N(x)_{s,t} = 1 + \sum_{n=1}^N \mathbf{X}_{s,t}^n.$$

For a path  $x$ , write  $x_{s,t} = x_t - x_s$  and we have  $x_{s,t} = \mathbf{X}_{s,t}^1$ . It is well known that  $S_N(x)$  is a multiplicative functional, that is, for any  $s < u < t$ , we have

$$S_N(x)_{s,u} \otimes S_N(x)_{u,t} = S_N(x)_{s,t},$$

where the multiplication takes place in  $T^N(\mathbb{R}^d)$ .

We equip each subspace  $(\mathbb{R}^d)^{\otimes n}$  with the projective tensor norm<sup>1</sup>  $|\cdot|$ . If  $\mathbf{X}, \mathbf{Y}$  are two multiplicative functionals in  $T^N(\mathbb{R}^d)$ , then for each  $p \geq 1$ , we define their  $p$ -variation distance by

$$\varrho_{p-var}^N(\mathbf{X}, \mathbf{Y}) := \max_{n \leq N} \sup_{\mathcal{P}} \left( \sum_i |\mathbf{X}_{t_i, t_{i+1}}^n - \mathbf{Y}_{t_i, t_{i+1}}^n|^{\frac{n}{p}} \right)^{\frac{p}{n}},$$

where the supremum is taken over all finite partitions of the interval  $[0, T]$ . If  $N = \lfloor p \rfloor$ , this defines a rough paths metric and we only write  $\varrho_{p-var}(\mathbf{X}, \mathbf{Y})$  in this case.

**Remark 2.1.** *Note that we only consider so-called inhomogeneous rough paths metrics in this paper. The reason for this is that the Itô-Lyons solution map for rough paths is locally Lipschitz with respect to these metrics (cf. [6, Chapter 10]).*

We define the subset  $G^N(\mathbb{R}^d) \subset T^N(\mathbb{R}^d)$  to be

$$G^N(\mathbb{R}^d) = \{S_N(x)_{0,1} : x \in C^{1-var}(\mathbb{R}^d)\}.$$

The multiplicativity of  $S_N$  implies that  $G^N(\mathbb{R}^d)$  is a group with multiplication  $\otimes$  and identity element 1. If  $x \in C^{1-var}(\mathbb{R}^d)$ , one can show that actually  $S_N(x)_{s,t} \in G^N(\mathbb{R}^d)$  for all  $s < t$ .

<sup>1</sup>The choice of the tensor norm is actually not important since we only work in finite dimensional spaces.

**Definition 2.2.** A function  $\omega : \Delta \rightarrow \mathbb{R}^+$  is called a control if it is continuous, vanishes on the diagonal, and is superadditive in the sense that for any  $s < u < t$ , we have

$$\omega(s, u) + \omega(u, t) \leq \omega(s, t).$$

We say a multiplicative functional  $\mathbf{X}$  in  $T^N(\mathbb{R}^d)$  has finite  $p$ -variation ( $p \geq 1$ ) controlled by  $\omega$  if for each  $n \leq N$ , there exists a constant  $C_n$  such that for all  $s < t$ , we have

$$|\mathbf{X}_{s,t}^n| \leq C_n \omega(s, t)^{\frac{n}{p}}.$$

**Definition 2.3.** Let  $p \geq 1$ . A geometric  $p$ -rough path is a continuous path in  $G^{\lfloor p \rfloor}(\mathbb{R}^d)$  which is in the  $p$ -variation closure (w.r.t the metric  $\varrho_{p\text{-var}}$ ) of the set of bounded variation paths. We use  $C^{0,p\text{-var}}([0, T], G^{\lfloor p \rfloor}(\mathbb{R}^d))$  to denote the space of geometric  $p$ -rough paths.

By Lyon’s extension theorem (cf. [11, Theorem 2.2.1]), every (geometric)  $p$ -rough path  $\mathbf{X}$  can be lifted to a  $q$ -rough path for every  $q \geq p$ . Abusing notation, we will use the same letter  $\mathbf{X}$  to denote this lift. We will write  $\mathbf{X}^n = \pi_n(\mathbf{X})$  where  $\pi_n$  denotes the projection of  $T^N(\mathbb{R}^d)$  onto the  $n$ -th tensor product,  $n \leq N$ . If  $x = \pi_1(\mathbf{X})$ , we will also use the notation

$$\mathbf{X}_{s,t}^n = \int_{\Delta_{s,t}^n} dx \otimes \cdots \otimes dx$$

(even though this integral does not have to exist as a limit of Riemann sums).

### 3 2D variation and Gaussian rough paths

If  $I = [a, b]$  is an interval, a dissection of  $I$  is a finite subset of points of the form  $\{a = t_0 < \dots < t_m = b\}$ . The family of all dissections of  $I$  is denoted by  $\mathcal{D}(I)$ .

Let  $I \subset \mathbb{R}$  be an interval and  $A = [a, b] \times [c, d] \subset I \times I$  be a rectangle. If  $f : I \times I \rightarrow V$  is a function, mapping into a normed vector space  $V$ , we define the rectangular increment  $f(A)$  by setting

$$f(A) := f \begin{pmatrix} a, b \\ c, d \end{pmatrix} := f \begin{pmatrix} b \\ d \end{pmatrix} - f \begin{pmatrix} a \\ d \end{pmatrix} - f \begin{pmatrix} b \\ c \end{pmatrix} + f \begin{pmatrix} a \\ c \end{pmatrix}.$$

**Definition 3.1.** Let  $p \geq 1$  and  $f : I \times I \rightarrow V$ . For  $[s, t] \times [u, v] \subset I \times I$ , set

$$V_p(f; [s, t] \times [u, v]) := \left( \sup_{\substack{(t_i) \in \mathcal{D}([s, t]) \\ (t'_j) \in \mathcal{D}([u, v])}} \sum_{t_i, t'_j} \left| f \begin{pmatrix} t_i, t_{i+1} \\ t'_j, t'_{j+1} \end{pmatrix} \right|^p \right)^{\frac{1}{p}}.$$

If  $V_p(f, I \times I) < \infty$ , we say that  $f$  has finite (2D)  $p$ -variation. We also define

$$V_\infty(f; [s, t] \times [u, v]) := \sup_{\substack{\sigma, \tau \in [s, t] \\ \mu, \nu \in [u, v]}} \left| f \begin{pmatrix} \sigma, \tau \\ \mu, \nu \end{pmatrix} \right|$$

**Lemma 3.2.** *Let  $f: I \times I \rightarrow V$  be a continuous map and  $1 \leq p \leq p' < \infty$ . Assume that  $f$  has finite  $p$ -variation. Then for every  $[s, t] \times [u, v] \subset I \times I$  we have*

$$V_{p'}(f; [s, t] \times [u, v]) \leq V_\infty(f; [s, t] \times [u, v])^{1-\frac{p}{p'}} V_p(f; [s, t] \times [u, v])^{\frac{p}{p'}}$$

*Proof.* Let  $(t_i) \in \mathcal{D}([s, t])$  and  $(t'_j) \in \mathcal{D}([u, v])$ . Then,

$$\sum_{t_i, t'_j} \left| f \left( \begin{matrix} t_i, t_{i+1} \\ t'_j, t'_{j+1} \end{matrix} \right) \right|^{p'} \leq V_\infty(f; [s, t] \times [u, v])^{p'-p} \sum_{t_i, t'_j} \left| f \left( \begin{matrix} t_i, t_{i+1} \\ t'_j, t'_{j+1} \end{matrix} \right) \right|^p.$$

Taking the supremum over all partitions gives the claim. □

**Lemma 3.3.** *Let  $f: I \times I \rightarrow \mathbb{R}$  be continuous with finite  $p$ -variation. Choose  $p'$  such that  $p' \geq p$  if  $p = 1$  and  $p' > p$  if  $p > 1$ . Then there is a control  $\omega$  and a constant  $C = C(p, p')$  such that*

$$V_{p'}(f; J \times J) \leq \omega(J)^{\frac{1}{p'}} \leq C V_p(f; J \times J)$$

*holds for every interval  $J \subset I$ .*

*Proof.* Follows from [7, Theorem 1]. □

Let  $X = (X^1, \dots, X^d): I \rightarrow \mathbb{R}^d$  be a centered, stochastic process. Then the covariance function  $R_X(s, t) := \text{Cov}_X(s, t) = E(X_s \otimes X_t)$  is a map  $R_X: I \times I \rightarrow \mathbb{R}^{d \times d}$  and we can ask for its  $\rho$ -variation (we will use the letter  $\rho$  instead of  $p$  in this context). Clearly,  $R_X$  has finite  $\rho$ -variation if and only if for every  $i, j \in \{1, \dots, d\}$  the map  $s, t \mapsto E(X_s^i X_t^j)$  has finite  $\rho$ -variation. In particular, if  $X^i$  and  $X^j$  are independent for  $i \neq j$ ,  $R_X$  has finite  $\rho$ -variation if and only if  $R_{X^i}$  has finite  $\rho$ -variation for every  $i = 1, \dots, d$ . In the next example, we calculate the  $\rho$ -variation for the covariances of some well-known real valued Gaussian processes. In particular, we will see that many interesting Gaussian processes have a covariance of finite 1-variation.

**Example 3.4.** (i) *Let  $X = B$  be a Brownian motion. Then  $R_B(s, t) = \min\{s, t\}$  and thus, for  $A = [s, t] \times [u, v]$ ,*

$$|R(A)| = |(s, t) \cap (u, v)| = \int_{[s, t] \times [u, v]} \delta_{x=y} dx dy.$$

*This shows that  $R_B$  has finite 1-variation on any interval  $I$ .*

(ii) *More generally, let  $f: [0, T] \rightarrow \mathbb{R}$  be a left-continuous, locally bounded function. Set*

$$X_t = \int_0^t f(r) dB_r.$$

*Then, for  $A = [s, t] \cap [u, v]$  we have by the Itô isometry,*

$$R_X(A) = E \left[ \int_{[s, t]} f dB \int_{[u, v]} f dB \right] = \int_{[s, t] \times [u, v]} \delta_{x=y} f(x) f(y) dx dy$$

*which shows that  $R_X$  has finite 1-variation.*

(iii) Let  $X$  be an Ornstein-Uhlenbeck process, i.e.  $X$  is the solution of the SDE

$$dX_t = -\theta X_t dt + \sigma dB_t \tag{3.1}$$

for some  $\theta, \sigma > 0$ . If we claim that  $X_0 = 0$ , one can show that  $X$  is centered, Gaussian and a direct calculation shows that the covariance of  $X$  has finite 1-variation on any interval  $[0, T]$ . The same is true considering the stationary solution of (3.1) instead.

(iv) If  $X$  is a continuous Gaussian martingale, it can be written as a time-changed Brownian motion. Since the  $\rho$ -variation of its covariance is invariant under time-change,  $X$  has again a covariance of finite 1-variation.

(v) If  $X: [0, T] \rightarrow \mathbb{R}$  is centered Gaussian with  $X_0 = 0$ , we can define a Gaussian bridge by

$$X_{\text{Bridge}}(t) = X_t - t \frac{X_T}{T}.$$

One can easily show that if the covariance of  $X$  has finite  $\rho$ -variation, the same is true for  $X_{\text{Bridge}}$ . In particular, the covariance of a Brownian bridge has finite 1-variation.

Next, we cite the fundamental existence result about Gaussian rough paths. For a proof, cf. [5] or [6, Chapter 15].

**Theorem 3.5** (Friz, Victoir). *Let  $X: [0, T] \rightarrow \mathbb{R}^d$  be a centered Gaussian process with continuous sample paths and independent components. Assume that there is a  $\rho \in [1, 2)$  such that  $V_\rho(R_X; [0, T]^2) < \infty$ . Then  $X$  admits a lift  $\mathbf{X}$  to a process whose sample paths are geometric  $p$ -rough paths for any  $p > 2\rho$ , i.e. with sample paths in  $C^{0, p-\text{var}}([0, T], G^{[p]}(\mathbb{R}^d))$  and  $\pi_1(\mathbf{X}_{s,t}) = X_t - X_s$  for any  $s < t$ .*

In the next proposition, we give an upper  $L^2$ -estimate for the difference of two Gaussian rough paths on the first two levels.

**Proposition 3.6.** *Let  $(X, Y) = (X^1, Y^1, \dots, X^d, Y^d): [0, T] \rightarrow \mathbb{R}^{d+d}$  be a centered Gaussian process with continuous sample paths where  $(X^i, Y^i)$  and  $(X^j, Y^j)$  are independent for  $i \neq j$ . Let  $\rho \in [1, \frac{3}{2})$  and assume that  $V_{\rho'}(R_{(X,Y)}; [0, T]^2) \leq K < +\infty$  for a constant  $K > 0$  where  $\rho' < \rho$  in the case  $\rho > 1$  and  $\rho' = 1$  in the case  $\rho = 1$ . Let  $\gamma \geq \rho$  such that  $\frac{1}{\gamma} + \frac{1}{\rho} > 1$ . Then there are constants  $C_0, C_1, C_2$  depending on  $\rho, \rho', \gamma$  and  $K$  and a control  $\omega$  such that  $\omega(0, T) \leq C_0$  and*

$$|X_{s,t} - Y_{s,t}|_{L^2} \leq C_1 \sup_{u \in [s,t]} |X_u - Y_u|_{L^2}^{1-\frac{\rho}{\gamma}} \omega(s, t)^{\frac{1}{2\gamma}}$$

and

$$\left| \int_s^t X_{s,u} \otimes dX_u - \int_s^t Y_{s,u} \otimes dY_u \right|_{L^2} \leq C_2 \sup_{u \in [s,t]} |X_u - Y_u|_{L^2}^{1-\frac{\rho}{\gamma}} \omega(s, t)^{\frac{1}{2\gamma} + \frac{1}{2\rho}}$$

hold for every  $s < t$ .

*Proof.* Note first that, by assumption on  $V_{\rho'}(R_{(X,Y)}; [0, T]^2)$ , Lemma 3.3 guarantees that there is a control  $\omega$  and a constant  $c_1 = c_1(\rho, \rho')$  such that

$$V_\rho(R_X; [s, t]^2) \vee V_\rho(R_Y; [s, t]^2) \vee V_\rho(R_{(X-Y)}; [s, t]^2) \leq \omega(s, t)^{1/\rho}$$



holds for all  $s < t$  and  $i = 1, \dots, d$  with the property that  $\omega(0, T) \leq c_1 K^\rho =: C_0$ . We will estimate both levels componentwise. We start with the first level. Let  $i \in \{1, \dots, d\}$ . Then,

$$\begin{aligned} |X_{s,t}^i - Y_{s,t}^i|_{L^2}^2 &= \left| R_{(X^i - Y^i)} \begin{pmatrix} s, t \\ s, t \end{pmatrix} \right| \\ &\leq V_\gamma(R_{(X-Y)}; [s, t]^2) \end{aligned}$$

and thus

$$|X_{s,t} - Y_{s,t}|_{L^2} \leq c_2 \sqrt{V_\gamma(R_{(X-Y)}; [s, t]^2)}.$$

For the second level, consider first the case  $i = j$ . We have, using that  $(X, Y)$  is Gaussian and that we are dealing with *geometric* rough paths,

$$\begin{aligned} \left| \int_s^t X_{s,u}^i dX_u^i - \int_s^t Y_{s,u}^i dY_u^i \right|_{L^2} &= \frac{1}{2} |(X_{s,t}^i)^2 - (Y_{s,t}^i)^2|_{L^2} \\ &= \frac{1}{2} |(X_{s,t}^i - Y_{s,t}^i)(X_{s,t}^i + Y_{s,t}^i)|_{L^2} \\ &\leq c_3 |X_{s,t}^i - Y_{s,t}^i|_{L^2} (|X_{s,t}^i|_{L^2} + |Y_{s,t}^i|_{L^2}). \end{aligned}$$

From the first part, we know that

$$|X_{s,t}^i - Y_{s,t}^i|_{L^2} \leq \sqrt{V_\gamma(R_{(X-Y)}; [s, t]^2)}.$$

Furthermore,

$$|X_{s,t}^i|_{L^2} = \sqrt{\left| R_X \begin{pmatrix} s, t \\ s, t \end{pmatrix} \right|} \leq \sqrt{V_\rho(R_X; [s, t]^2)} \leq \omega(s, t)^{\frac{1}{2\rho}}$$

and the same holds for  $|Y_{s,t}^i|_{L^2}$ . Hence

$$\left| \int_s^t X_{s,u}^i dX_u^i - \int_s^t Y_{s,u}^i dY_u^i \right|_{L^2} \leq c_4 \sqrt{V_\gamma(R_{(X-Y)}; [s, t]^2)} \omega(s, t)^{\frac{1}{2\rho}}.$$

For  $i \neq j$ ,

$$\begin{aligned} &\left| \int_s^t X_{s,u}^i dX_u^j - \int_s^t Y_{s,u}^i dY_u^j \right|_{L^2} \\ &\leq \left| \int_s^t (X^i - Y^i)_{s,u} dX_u^j \right|_{L^2} + \left| \int_s^t Y_{s,u}^i d(X^j - Y^j)_u \right|_{L^2}. \end{aligned}$$

We estimate the first term. From independence,

$$E \left[ \left( \int_s^t (X^i - Y^i)_{s,u} dX_u^j \right)^2 \right] = \int_{[s,t]^2} R_{(X^i - Y^i)} \begin{pmatrix} s, u \\ s, v \end{pmatrix} dR_{X^j}(u, v)$$

where the integral on the right is a  $2D$  Young integral.<sup>2</sup> By a  $2D$  Young estimate (cf. [16]),

$$\begin{aligned} \left| \int_{[s,t]^2} R_{(X^i - Y^i)} \begin{pmatrix} s, u \\ s, v \end{pmatrix} dR_{X^j}(u, v) \right| &\leq c_5(\rho, \gamma) V_\gamma(R_{(X^i - Y^i)}; [s, t]^2) V_\rho(R_{X^j}; [s, t]^2) \\ &\leq c_6 V_\gamma(R_{(X-Y)}; [s, t]^2) \omega(s, t)^{1/\rho}. \end{aligned}$$

<sup>2</sup>The reader might feel a bit uncomfortable at this point asking why it is allowed to put expectation inside the integral (which is not even an integral in Riemann-Stieltjes sense). However, this can be made rigorous by dealing with processes which have sample paths of bounded variation first and passing to the limit afterwards (cf. [5, 6, 4, 2]). We decided not to go too much into detail here in order not to distract the reader from the main ideas and to improve the readability.

The second term is treated exactly in the same way. Summarizing, we have shown that

$$\left| \int_s^t X_{s,u} \otimes dX_u - \int_s^t Y_{s,u} \otimes dY_u \right|_{L^2} \leq C \sqrt{V_\gamma(R_{(X-Y)}; [s, t]^2)} \omega(s, t)^{\frac{1}{2\rho}}.$$

Finally, by Lemma 3.2

$$V_\gamma(R_{(X-Y)}; [s, t]^2) \leq V_\infty(R_{(X-Y)}; [s, t]^2)^{1-\rho/\gamma} \omega(s, t)^{1/\gamma}$$

and by the Cauchy-Schwarz inequality

$$V_\infty(R_{(X-Y)}; [s, t]^2) \leq 4 \sup_{u \in [s, t]} |X_u - Y_u|_{L^2}^2$$

which gives the claim. □

**Corollary 3.7.** *Under the assumptions of Proposition 3.6, for every  $\gamma$  satisfying  $\gamma \geq \rho$  and  $\frac{1}{\gamma} + \frac{1}{\rho} > 1$ , and every  $p > 2\rho$  and  $\gamma' > \gamma$ , there is a (random) control  $\hat{\omega}$  such that*

$$|\mathbf{X}_{s,t}^n| \leq \hat{\omega}(s, t)^{n/p} \tag{3.2}$$

$$|\mathbf{Y}_{s,t}^n| \leq \hat{\omega}(s, t)^{n/p} \tag{3.3}$$

$$|\mathbf{X}_{s,t}^n - \mathbf{Y}_{s,t}^n| \leq \epsilon \hat{\omega}(s, t)^{\frac{1}{2\gamma'} + \frac{n-1}{p}} \tag{3.4}$$

holds a.s. for all  $s < t$  and  $n = 1, 2$  where  $\epsilon = \sup_{u \in [0, T]} |X_u - Y_u|_{L^2}^{1-\frac{\rho}{\gamma}}$ . Furthermore, there is a constant  $C = C(p, \rho, \gamma, \gamma', K)$  such that

$$|\hat{\omega}(0, T)|_{L^q} \leq CT(q^{p/2} + q^{\gamma'})$$

holds for all  $q \geq 1$ .

*Proof.* Let  $\omega$  be the control from Proposition 3.6. We know that

$$|\mathbf{X}_{s,t}^n - \mathbf{Y}_{s,t}^n|_{L^2} \leq c_1 \epsilon \omega(s, t)^{\frac{1}{2\gamma} + \frac{n-1}{2\rho}}$$

holds for a constant  $c_1$  for all  $s < t$  and  $n = 1, 2$ . Furthermore,  $|\mathbf{X}_{s,t}^n|_{L^2} \leq c_2 \omega(s, t)^{\frac{n}{2\rho}}$  for a constant  $c_2$  for all  $s < t$  and  $n = 1, 2$  and the same holds for  $\mathbf{Y}$  (this just follows from setting  $Y = \text{const.}$  and  $\gamma = \rho$  in Proposition 3.6). Now introduce a new process  $\tilde{\mathbf{X}} : [0, T] \rightarrow \mathbb{R}^d$  on the same sample space as  $\mathbf{X}$  such that for all sample points, we have

$$\tilde{\mathbf{X}}_{\omega(0,t)/\omega(0,T)} = \mathbf{X}_t, \quad \forall t \in [0, T],$$

and define  $\tilde{\mathbf{Y}}$  in the same way. Then  $\tilde{\mathbf{X}}, \tilde{\mathbf{Y}}$  are well defined, multiplicative, and we can replace the control  $\omega$  by  $c_3 K |t - s|$  for the two re-parametrized processes. Using that  $(X, Y)$  is Gaussian, we may conclude that  $\mathbf{X}_{s,t}^n, \mathbf{Y}_{s,t}^n$  and  $\mathbf{X}_{s,t}^n - \mathbf{Y}_{s,t}^n$  are elements in the  $n$ -th Wiener chaos for any  $s < t$  (cf. [5, Proposition 21]), hence we can pass from  $L^2$  to  $L^q$  estimates (cf. [5, Lemma 22]), thus

$$|\tilde{\mathbf{X}}_{s,t}^n|_{L^q} \leq c_4 (\sqrt{qK^{1/\rho}})^n |t - s|^{\frac{n}{2\rho}} \tag{3.5}$$

$$|\tilde{\mathbf{Y}}_{s,t}^n|_{L^q} \leq c_4 (\sqrt{qK^{1/\rho}})^n |t - s|^{\frac{n}{2\rho}} \tag{3.6}$$

$$|\tilde{\mathbf{X}}_{s,t}^n - \tilde{\mathbf{Y}}_{s,t}^n|_{L^q} \leq \tilde{\epsilon} c_4 (\sqrt{qK^{1/\rho}})^n |t - s|^{\frac{1}{2\gamma} + \frac{(n-1)}{2\rho}} \tag{3.7}$$

holds for all  $s < t$ ,  $n = 1, 2$  and  $q \geq 1$  with  $\tilde{\epsilon} = \epsilon K^{\frac{1}{2\gamma} - \frac{1}{2\rho}}$ . Using Lemma 6.1 in the appendix, we see that there is a constant  $c_5 = c_5(p, \rho, \gamma, \gamma', K)$  such that

$$\left| \sup_{s < t \in [0, T]} \frac{|\tilde{\mathbf{X}}_{s,t}^n|}{|t - s|^{n/p}} \right|_{L^q} \leq c_5 q^{n/2} \tag{3.8}$$

$$\left| \sup_{s < t \in [0, T]} \frac{|\tilde{\mathbf{Y}}_{s,t}^n|}{|t - s|^{n/p}} \right|_{L^q} \leq c_5 q^{n/2} \tag{3.9}$$

$$\left| \sup_{s < t \in [0, T]} \frac{|\tilde{\mathbf{X}}_{s,t}^n - \tilde{\mathbf{Y}}_{s,t}^n|}{|t - s|^{1/p(n)}} \right|_{L^q} \leq \epsilon c_5 q^{n/2} \tag{3.10}$$

hold for  $q$  sufficiently large and  $n = 1, 2$  where  $\frac{1}{p(n)} = \frac{1}{2\gamma'} + \frac{n-1}{p}$ . Set

$$\begin{aligned} \hat{\omega}_X^n(s, t) &:= \sup_{D \subset [s, t]} \sum_{t_i \in D} |\mathbf{X}_{t_i, t_{i+1}}^n|^{p/n} \\ \hat{\omega}_Y^n(s, t) &:= \sup_{D \subset [s, t]} \sum_{t_i \in D} |\mathbf{Y}_{t_i, t_{i+1}}^n|^{p/n} \\ \hat{\omega}_{X-Y}^n(s, t) &:= \sup_{D \subset [s, t]} \sum_{t_i \in D} |\mathbf{X}_{t_i, t_{i+1}}^n - \mathbf{Y}_{t_i, t_{i+1}}^n|^{p(n)} \end{aligned}$$

and

$$\hat{\omega}(s, t) := \sum_{n=1,2} \hat{\omega}_X^n(s, t) + \hat{\omega}_Y^n(s, t) + \epsilon^{\frac{1}{p(n)}} \hat{\omega}_{X-Y}^n(s, t).$$

for  $s < t$ . Clearly,  $\hat{\omega}$  fulfils (3.2), (3.3) and (3.4). Moreover, the notion of  $p$ -variation is invariant under reparametrization, hence

$$\hat{\omega}_X^n(0, T) = \sup_{D \subset [0, T]} \sum_{t_i \in D} |\mathbf{X}_{t_i, t_{i+1}}^n|^{p/n} = \sup_{D \subset [0, T]} \sum_{t_i \in D} |\tilde{\mathbf{X}}_{t_i, t_{i+1}}^n|^{p/n} \leq T \sup_{s < t \in [0, T]} \frac{|\tilde{\mathbf{X}}_{s,t}^n|^{p/n}}{|t - s|}$$

and a similar estimate holds for  $\hat{\omega}_Y^n(0, T)$  and  $\hat{\omega}_{X-Y}^n(0, T)$ . By the triangle inequality and the estimates (3.5), (3.6) and (3.7),

$$\begin{aligned} |\hat{\omega}(0, T)|_{L^q} &\leq \sum_{n=1,2} |\hat{\omega}_X^n(0, T)|_{L^q} + |\hat{\omega}_Y^n(0, T)|_{L^q} + \epsilon^{\frac{1}{p(n)}} |\hat{\omega}_{X-Y}^n(0, T)|_{L^q} \\ &\leq c_6 T \left( q^{p/2} + q^{\frac{p(1)}{2}} + q^{p(2)} \right) \leq c_7 T (q^{p/2} + q^{\gamma'}) \end{aligned}$$

for  $q$  large enough. We can extend the estimate to all  $q \geq 1$  by making the constant larger if necessary. □

**Corollary 3.8.** *Let  $\hat{\omega}$  be the random control defined in the previous corollary. Then, for every  $n$ , there exists a constant  $c_n$  such that*

$$|\mathbf{X}_{s,t}^n| < c_n \hat{\omega}(s, t)^{\frac{n}{p}}, \quad |\mathbf{Y}_{s,t}^n| < c_n \hat{\omega}(s, t)^{\frac{n}{p}}$$

a.s. for all  $s < t$ . The constants  $c_n$  are deterministic and can be chosen such that  $c_n \leq \frac{2^n}{(n/p)!}$ , where  $x! := \Gamma(x + 1)$ .

*Proof.* Follows from the extension theorem, cf. [11, Theorem 2.2.1] or [12, Theorem 3.7]. □

## 4 Main estimates

In what follows, we let  $p \in (2\rho, 3)$ . Let  $\gamma \geq \rho$  such that  $\frac{1}{\gamma} + \frac{1}{\rho} > 1$ . We write  $\log^+ x = \max\{x, 0\}$ , and set

$$\epsilon = \sup_{u \in [0, T]} |X_u - Y_u|_{L^2}^{1 - \frac{\rho}{\gamma}}.$$

### 4.1 Higher level estimates

We first introduce some notations. Suppose  $\mathbf{X}$  is a multiplicative functional in  $T^N(\mathbb{R}^d)$  with finite  $p$ -variation controlled by  $\omega$ ,  $N \geq [p]$ . Then, define

$$\hat{\mathbf{X}}_{s,t} = 1 + \sum_{n=1}^N \mathbf{X}_{s,t}^n \in T^{N+1}(\mathbb{R}^d).$$

Then,  $\hat{\mathbf{X}}$  is multiplicative in  $T^N$ , but in general not in  $T^{N+1}$ . For any partition  $D = \{s = u_0 < u_1 < \dots < u_L < u_{L+1} = t\}$ , define

$$\hat{\mathbf{X}}_{s,t}^D := \hat{\mathbf{X}}_{s,u_1} \otimes \dots \otimes \hat{\mathbf{X}}_{u_L,t} \in T^{N+1}(\mathbb{R}^d).$$

The following lemma gives a construction of the unique multiplicative extension of  $\mathbf{X}$  to higher degrees. It was first proved in Theorem 2.2.1 in [11].

**Lemma 4.1.** *Let  $\mathbf{X}$  be a multiplicative functional in  $T^N$ . Let  $D = \{s < u_1 < \dots < u_L < t\}$  be any partition of  $(s, t)$ , and  $D^j$  denote the partition with the point  $u_j$  removed from  $D$ . Then,*

$$\hat{\mathbf{X}}_{s,t}^D - \hat{\mathbf{X}}_{s,t}^{D^j} = \sum_{n=1}^N \mathbf{X}_{u_{j-1},u_j}^n \otimes \mathbf{X}_{u_j,u_{j+1}}^{N+1-n} \in T^{N+1}(\mathbb{R}^d). \quad (4.1)$$

*In particular, its projection onto the subspace  $T^N$  is the 0-vector. Suppose further that  $\mathbf{X}$  has finite  $p$ -variation controlled by  $\omega$ , and  $N \geq [p]$ , then the limit*

$$\lim_{|D| \rightarrow 0} \hat{\mathbf{X}}_{s,t}^D \in T^{N+1}(\mathbb{R}^d)$$

*exists. Furthermore, it is the unique multiplicative extension of  $\mathbf{X}$  to  $T^{N+1}$  with finite  $p$ -variation controlled by  $\omega$ .*

**Theorem 4.2.** *Let  $(X, Y)$  and  $\rho, \gamma$  as in Proposition 3.6. Then for every  $p \in (2\rho, 3)$  and  $\gamma' > \gamma$  there exists a constant  $C_3$  depending on  $p$  and  $\gamma'$  and a (random) control  $\hat{\omega}$  such that for all  $q \geq 1$ , we have*

$$|\hat{\omega}(0, T)|_{L^q} \leq M < +\infty,$$

*where  $M = M(p, \rho, \gamma, \gamma', K, q)$ , and the following holds a.s. for all  $[s, t]$ :*

(i) *If  $\frac{1}{2\gamma'} + \frac{2}{p} > 1$ , then*

$$|\mathbf{X}_{s,t}^3 - \mathbf{Y}_{s,t}^3| < C_3 \epsilon \hat{\omega}(s, t)^{\frac{1}{2\gamma'} + \frac{2}{p}}.$$

(ii) If  $\frac{1}{2\gamma'} + \frac{2}{p} = 1$ , then

$$|\mathbf{X}_{s,t}^3 - \mathbf{Y}_{s,t}^3| < C_3 \epsilon \cdot (1 + \log^+ [\hat{\omega}(0, T)/\epsilon^{1-\frac{p}{2\gamma'}}]) \cdot \hat{\omega}(s, t).$$

(iii) If  $\frac{1}{2\gamma'} + \frac{2}{p} < 1$ , then

$$|\mathbf{X}_{s,t}^3 - \mathbf{Y}_{s,t}^3| < C_3 \epsilon^{\frac{3-p}{1-p/2\gamma'}} \hat{\omega}(s, t),$$

*Proof.* Let  $s < t \in [0, T]$  and let  $\hat{\omega}$  be the (random) control defined in Corollary 3.7. Then, by the same corollary, for every  $q \geq 1$ ,  $|\hat{\omega}(0, T)|_{L^q} \leq M$ . Fix an enhanced sample rough path  $(\mathbf{X}, \mathbf{Y})$  up to level 2 and for simplicity, we will use  $\omega$  to denote the corresponding realisation of the (random) control  $\hat{\omega}$ . We can assume without loss of generality that

$$\epsilon < \omega(s, t)^{\frac{1}{p} - \frac{1}{2\gamma'}}, \tag{4.2}$$

otherwise the triangle inequality and Corollary 3.8 readily imply the stated estimates. Let  $D = \{s = u_0 < \dots < u_{L+1} = t\}$  be a dissection. Then (cf. [11, Lemma 2.2.1]), there exists a  $j$  such that

$$\omega(u_{j-1}, u_{j+1}) \leq \frac{2}{L} \omega(s, t), \quad L \geq 1. \tag{4.3}$$

Let  $D^j$  denote the dissection with the point  $u_j$  removed from  $D$ . Then, we have

$$\begin{aligned} |(\hat{\mathbf{X}}_{s,t}^D - \hat{\mathbf{Y}}_{s,t}^D)^3| &< |(\hat{\mathbf{X}}_{s,t}^{D^j} - \hat{\mathbf{Y}}_{s,t}^{D^j})^3| + \sum_{k=1}^2 (|\mathbf{R}_{u_{j-1}, u_j}^k \otimes \mathbf{X}_{u_j, u_{j+1}}^{3-k}| \\ &+ |\mathbf{X}_{u_{j-1}, u_j}^k \otimes \mathbf{R}_{u_j, u_{j+1}}^{3-k}| + |\mathbf{R}_{u_{j-1}, u_j}^k \otimes \mathbf{R}_{u_j, u_{j+1}}^{3-k}|), \end{aligned}$$

where  $\mathbf{R}_{s,t} = \mathbf{Y}_{s,t} - \mathbf{X}_{s,t}$ . By assumption,

$$|\mathbf{R}_{u_{j-1}, u_j}^k \otimes \mathbf{R}_{u_j, u_{j+1}}^{3-k}| < C \cdot \min \left\{ \epsilon \left( \frac{1}{L} \omega(s, t) \right)^{\frac{1}{2\gamma'} + \frac{2}{p}}, \left( \frac{1}{L} \omega(s, t) \right)^{\frac{3}{p}} \right\}, \tag{4.4}$$

and similar inequalities hold for the other two terms in the bracket. Thus, we have

$$|(\hat{\mathbf{X}}_{s,t}^D - \hat{\mathbf{Y}}_{s,t}^D)^3| < |(\hat{\mathbf{X}}_{s,t}^{D^j} - \hat{\mathbf{Y}}_{s,t}^{D^j})^3| + C_3 \min \left\{ \epsilon \left( \frac{1}{L} \omega(s, t) \right)^{\frac{1}{2\gamma'} + \frac{2}{p}}, \left( \frac{1}{L} \omega(s, t) \right)^{\frac{3}{p}} \right\}.$$

Let  $N$  be the integer for which

$$\left[ \frac{1}{N+1} \omega(s, t) \right]^{\frac{1}{p} - \frac{1}{2\gamma'}} \leq \epsilon < \left[ \frac{1}{N} \omega(s, t) \right]^{\frac{1}{p} - \frac{1}{2\gamma'}}, \tag{4.5}$$

then

$$\epsilon \left[ \frac{1}{L} \omega(s, t) \right]^{\frac{1}{2\gamma'} + \frac{2}{p}} < \left[ \frac{1}{L} \omega(s, t) \right]^{\frac{3}{p}}$$

if and only if  $L \leq N$ . By Lemma 4.1, we have

$$\mathbf{X}_{s,t}^3 = \lim_{|D| \rightarrow 0} (\hat{\mathbf{X}}_{s,t}^D)^3, \quad \mathbf{Y}_{s,t}^3 = \lim_{|D| \rightarrow 0} (\hat{\mathbf{Y}}_{s,t}^D)^3.$$

Thus, for a fixed partition  $D$ , we choose a point each time according to (4.3), and drop them successively. By letting  $|D| \rightarrow +\infty$ , we have

$$|\mathbf{X}_{s,t}^3 - \mathbf{Y}_{s,t}^3| \leq C_3 \left[ \epsilon \sum_{L=1}^N \left( \frac{1}{L} \omega(s, t) \right)^{\frac{1}{2\gamma'} + \frac{2}{p}} + \sum_{L=N+1}^{+\infty} \left( \frac{1}{L} \omega(s, t) \right)^{\frac{3}{p}} \right].$$

Approximating the sums by integrals, we have

$$|\mathbf{X}_{s,t}^3 - \mathbf{Y}_{s,t}^3| < C_3[\epsilon\omega(s,t)^{\frac{1}{2\gamma'} + \frac{2}{p}}(1 + \int_1^N x^{-(\frac{1}{2\gamma'} + \frac{2}{p})} dx) + \omega(s,t)^{\frac{3}{p}} \int_N^{+\infty} x^{-\frac{3}{p}} dx].$$

Compute the second integral, and use

$$[\frac{1}{N+1}\omega(s,t)]^{(\frac{1}{p} - \frac{1}{2\gamma'})} \leq \epsilon,$$

we obtain

$$|\mathbf{X}_{s,t}^3 - \mathbf{Y}_{s,t}^3| < C_3[\epsilon\omega(s,t)^{\frac{1}{2\gamma'} + \frac{2}{p}}(1 + \int_1^N x^{-(\frac{1}{2\gamma'} + \frac{2}{p})} dx) + \epsilon^{\frac{3-p}{1-p/2\gamma'}}\omega(s,t)]. \tag{4.6}$$

Now we apply the above estimates to the three situations respectively.

**1.**  $\frac{1}{2\gamma'} + \frac{2}{p} > 1.$

In this case, the integral

$$\int_1^N x^{-(\frac{1}{2\gamma'} + \frac{2}{p})} dx < \int_1^{+\infty} x^{-(\frac{1}{2\gamma'} + \frac{2}{p})} dx < +\infty$$

converges. On the other hand, (4.2) implies

$$\epsilon^{\frac{3-p}{1-p/2\gamma'}}\omega(s,t) < \epsilon\omega(s,t)^{\frac{1}{2\gamma'} + \frac{2}{p}},$$

thus, from (4.6), we get

$$|\mathbf{X}_{s,t}^3 - \mathbf{Y}_{s,t}^3| < C_3\epsilon\omega(s,t)^{\frac{1}{2\gamma'} + \frac{2}{p}}.$$

**2.**  $\frac{1}{2\gamma'} + \frac{2}{p} = 1.$

In this case,  $\frac{1}{p} - \frac{1}{2\gamma'} = \frac{3-p}{p}$ , and  $\frac{3-p}{1-p/2\gamma'} = 1$ . Thus, by the second inequality in (4.5), we have

$$\int_1^N x^{-1} dx = \log N < \log \omega(s,t) - \frac{p}{3-p} \log \epsilon.$$

On the other hand, (4.2) gives

$$\log \omega(s,t) - \frac{p}{3-p} \log \epsilon > 0.$$

Combining the previous two bounds with (4.6), we get

$$|\mathbf{X}_{s,t}^3 - \mathbf{Y}_{s,t}^3| < C_3\epsilon[1 + \log \omega(s,t) - \frac{p}{3-p} \log \epsilon]\omega(s,t).$$

We can simplify the above inequality to

$$|\mathbf{X}_{s,t}^3 - \mathbf{Y}_{s,t}^3| < C_3\epsilon[1 + \log^+(\omega(0,T)/\epsilon^{\frac{p}{3-p}})]\omega(s,t),$$

where we have also included the possibility of  $\epsilon \geq \omega(0,T)^{\frac{3}{p}-1}$ .

**3.**  $\frac{1}{2\gamma'} + \frac{2}{p} < 1.$

Now we have

$$1 + \int_1^N x^{-(\frac{1}{2\gamma'} + \frac{2}{p})} dx < CN^{1 - \frac{1}{2\gamma'} - \frac{2}{p}} < C \cdot \epsilon^{-(1 - \frac{1}{2\gamma'} - \frac{2}{p})/\frac{1}{p} - \frac{1}{2\gamma'}} \omega(s, t)^{1 - \frac{1}{2\gamma'} - \frac{2}{p}},$$

where the second inequality follows from (4.5). Combining the above bound with (4.6), we obtain

$$|\mathbf{X}_{s,t}^3 - \mathbf{Y}_{s,t}^3| < C_3 \epsilon^{\frac{3-p}{1-p/2\gamma'}} \omega(s, t).$$

□

The following theorem, obtained with the standard induction argument, gives estimates for all levels  $n = 1, 2, \dots$ .

**Theorem 4.3.** *Let  $(X, Y)$  and  $\rho, \gamma$  as in Proposition 3.6,  $p \in (2\rho, 3)$  and  $\gamma' > \gamma$ . Then there exists a (random) control  $\hat{\omega}$  such that for every  $q \geq 1$ , we have*

$$|\hat{\omega}(0, T)|_{L^q} \leq M$$

where  $M = M(p, \rho, \gamma, \gamma', q, K)$ , and for each  $n$  there exists a (deterministic) constant  $C_n$  depending on  $p$  and  $\gamma'$  such that a.s. for all  $[s, t]$ :

(i) *If  $\frac{1}{2\gamma'} + \frac{2}{p} > 1$ , then we have*

$$|\mathbf{X}_{s,t}^n - \mathbf{Y}_{s,t}^n| < C_n \epsilon \hat{\omega}(s, t)^{\frac{1}{2\gamma'} + \frac{n-1}{p}}$$

(ii) *If  $\frac{1}{2\gamma'} + \frac{2}{p} = 1$ , then we have*

$$|\mathbf{X}_{s,t}^n - \mathbf{Y}_{s,t}^n| < C_n \epsilon \cdot (1 + \log^+ [\hat{\omega}(0, T)/\epsilon^{1 - \frac{p}{2\gamma'}}]) \cdot \hat{\omega}(s, t)^{\frac{1}{2\gamma'} + \frac{n-1}{p}}.$$

(iii) *If  $\frac{1}{2\gamma'} + \frac{2}{p} < 1$ , then for all  $s < t$  and all small  $\epsilon$ , we have*

$$|\mathbf{X}_{s,t}^n - \mathbf{Y}_{s,t}^n| < C_n \epsilon^{\frac{3-p}{1-p/2\gamma'}} \hat{\omega}(s, t)^{\frac{n-1+\{p\}}{p}} \tag{4.7}$$

where  $\{p\}$  denotes the fractional part of  $p$ , i.e.  $\{p\} = p - \lfloor p \rfloor = p - 2$ .

*Proof.* We prove the case when  $\frac{1}{2\gamma'} + \frac{2}{p} < 1$ ; the other two situations are similar. Let  $\hat{\omega}$  be the control in the previous theorem. Fix an enhanced sample path  $(\mathbf{X}, \mathbf{Y})$ , the corresponding realisation  $\omega$  of  $\hat{\omega}$ , and  $s < t \in [0, T]$ . We may still assume (4.2) without loss of generality (otherwise we just apply Corollary 3.8). Thus, for  $n = 1, 2$ , we have

$$|\mathbf{X}_{s,t}^n - \mathbf{Y}_{s,t}^n| < \epsilon \omega(s, t)^{\frac{1}{2\gamma'} + \frac{n-1}{p}} < C_n \epsilon^{\frac{3-p}{1-p/2\gamma'}} \omega(s, t)^{\frac{n-1+\{p\}}{p}},$$

where the second inequality comes from (4.2). The above inequality also holds for  $n = 3$  by the previous theorem. Now, suppose (4.7) holds for  $n = 1, \dots, k$ , where  $k \geq 3$ , then for level  $n = k + 1$ , we have in the exponent

$$\frac{n - 1 + \{p\}}{p} > 1,$$

so that the usual induction procedure works (cf. [11], Theorem 2.2.2.). Thus, we prove (4.7) for all  $n$ . □

**4.2 Proof of Theorem 1.1**

*Proof.* We distinguish two cases:

**Case 1:** Assume that  $1 - \frac{2\rho}{\sigma} \leq 3 - 2\rho$ , i.e.  $1 \leq \frac{1}{\rho} + \frac{1}{\sigma}$ . This means that for every choice  $\gamma < \sigma/2$  we have

$$\frac{1}{\rho} + \frac{1}{2\gamma} > 1.$$

We can choose  $\gamma$  close to  $\sigma/2$  such that also

$$\beta < 1 - \frac{\rho}{\gamma}.$$

Now choose  $p$  and  $\gamma'$  such that  $2\rho < p < \sigma$ ,  $\gamma < \gamma' < \sigma/2$  and

$$\frac{2}{p} + \frac{1}{2\gamma'} > 1.$$

Theorem 4.3 implies that

$$|\mathbf{X}_{s,t}^n - \mathbf{Y}_{s,t}^n| < C_n \epsilon \hat{\omega}(s, t)^{\frac{1}{2\gamma'} + \frac{n-1}{p}}$$

holds for any  $s < t$  and  $n \in \mathbb{N}$ . Our choices on  $p$  and  $\gamma'$  imply that  $\frac{1}{2\gamma'} + \frac{n}{p} \geq \frac{n}{\sigma}$  for every  $n$ , therefore

$$\begin{aligned} \varrho_{\sigma\text{-var}}^N(\mathbf{X}, \mathbf{Y}) &\leq C_N \epsilon \max_{n=1, \dots, N} \hat{\omega}(0, T)^{\frac{1}{2\gamma'} + \frac{n-1}{p}} \\ &= C_N \sup_{u \in [0, T]} |X_u - Y_u|_{L^2}^{1 - \frac{\rho}{\gamma}} \max_{n=1, \dots, N} \hat{\omega}(0, T)^{\frac{n-1+\{p\}}{p}}. \end{aligned}$$

Taking the  $L^q$  norm on both sides gives the claim.

**Case 2:** Now let  $3 - 2\rho < 1 - \frac{2\rho}{\sigma}$ , i.e.

$$\frac{1}{\rho} + \frac{1}{\sigma} < 1. \tag{4.8}$$

Choose  $\gamma$  such that

$$\rho < \gamma < \min \left\{ \frac{\rho}{\rho-1}, \frac{\sigma}{2} \right\}.$$

We claim that we can choose  $\gamma$  such that also

$$\frac{1}{\rho} + \frac{1}{2\gamma} < 1 \tag{4.9}$$

holds. Indeed: In the case  $\gamma < \sigma/2 \leq \frac{\rho}{\rho-1}$ , we may choose  $\gamma$  close to  $\sigma/2$  using (4.8) to obtain (4.9). The case  $\gamma < \frac{\rho}{\rho-1} < \sigma/2$  may only happen if  $\rho > 1$ , i.e. we may choose  $\gamma$  close to  $\frac{\rho}{\rho-1}$  to get (4.9) which shows the claim. Now choose  $p > 2\rho$  and  $\gamma' \in (\gamma, \sigma/2)$  with the properties that  $\beta < 3 - p < 3 - 2\rho$  and  $\frac{2\rho}{\gamma} = \frac{p}{\gamma'}$  hold. With these choices,  $\frac{2}{p} + \frac{1}{2\gamma'} < 1$  and the corresponding estimate of Theorem 4.3 shows that

$$|\mathbf{X}_{s,t}^n - \mathbf{Y}_{s,t}^n| < C_n \epsilon^{\frac{3-p}{1-\rho/\gamma}} \hat{\omega}(s, t)^{\frac{n-1+\{p\}}{p}}$$

holds for every  $s < t$  and  $n \in \mathbb{N}$ . Since

$$\frac{\sigma(n-1+\{p\})}{np} = \frac{\sigma}{p} \left( 1 + \frac{p-3}{n} \right) \geq \sigma \left( 1 - \frac{2}{p} \right) \geq 2\gamma \left( 1 - \frac{1}{\rho} \right) \geq 1$$



we may conclude that

$$\begin{aligned} \varrho_{\sigma\text{-var}}^N(\mathbf{X}, \mathbf{Y}) &\leq C_N \epsilon^{\frac{3-p}{1-\rho/\gamma}} \max_{n=1, \dots, N} \hat{\omega}(0, T)^{\frac{n-1+\{p\}}{p}} \\ &= C_N \sup_{u \in [0, T]} |X_u - Y_u|_{L^2}^{3-p} \max_{n=1, \dots, N} \hat{\omega}(0, T)^{\frac{n-1+\{p\}}{p}}. \end{aligned}$$

Taking  $L^q$  norms yields the assertion. □

## 5 Applications

### 5.1 Convergence rates of rough differential equation

Consider the *rough differential equation* of the form

$$dY_t = \sum_{i=1}^d V_i(Y_t) dX_t^i =: V(Y_t) dX_t; \quad Y_0 \in \mathbb{R}^e \tag{5.1}$$

where  $X$  is a centered Gaussian process in  $\mathbb{R}^d$  with independent components and  $V = (V_i)_{i=1}^d$  a collection of bounded, smooth vector fields with bounded derivatives in  $\mathbb{R}^e$ . Rough path theory gives meaning to the pathwise solution to (5.1) in the case when the covariance  $R_X$  has finite  $\rho$ -variation for some  $\rho < 2$ . Assume that  $\rho \in [1, \frac{3}{2})$  and that there is a constant  $K$  such that

$$V_\rho(R_X; [s, t]^2) \leq K|t - s|^{\frac{1}{\rho}} \tag{5.2}$$

for all  $s < t$  (note that this condition implies that the sample paths of  $X$  are  $\alpha$ -Hölder for all  $\alpha < \frac{1}{2\rho}$ ). For simplicity, we also assume that  $[0, T] = [0, 1]$ . For every  $k \in \mathbb{N}$ , we can approximate the sample paths of  $X$  piecewise linear at the time points  $\{0 < 1/k < 2/k < \dots < (k-1)/k < 1\}$ . We will denote this process by  $X^{(k)}$ . Clearly,  $X^{(k)} \rightarrow X$  uniformly as  $k \rightarrow \infty$ . Now we substitute  $X$  by  $X^{(k)}$  in (5.1), solve the equation and obtain a solution  $Y^{(k)}$ ; we call this the Wong-Zakai approximation of  $Y$ . One can show, using rough path theory, that  $Y^{(k)} \rightarrow Y$  a.s. in uniform topology as  $k \rightarrow \infty$ . The proposition below is an immediate consequence of Theorem 1.1 and gives us rates of convergence.

**Proposition 5.1.** *The mesh size  $\frac{1}{k}$  Wong-Zakai approximation converges uniformly to the solution of (5.1) with a.s. rate at least  $k^{-(\frac{3}{2\rho}-1-\delta)}$  for any  $\delta \in (0, \frac{3}{2\rho}-1)$ . In particular, the rate is arbitrarily close to  $\frac{1}{2}$  when  $\rho = 1$ , which is the sharp rate in that case.*

*Proof.* First, one shows that (5.2) implies that

$$\sup_{t \in [0, 1]} |X_t^{(k)} - X_t|_{L^2} = O(k^{-\frac{1}{2\rho}}).$$

One can show (cf. [6, Chapter 15.2.3]) that there is a constant  $C$  such that

$$\sup_{k \in \mathbb{N}} V_\rho(R_{(X, X^{(k)})}; [s, t]^2) \leq C|t - s|^{\frac{1}{\rho}}$$

holds for all  $s < t$ . By choosing  $q$  large enough, a Borel-Cantelli type argument applied to Theorem 1.1 shows that  $\varrho_{\sigma\text{-var}}(\mathbf{X}, \mathbf{X}^{(k)}) \rightarrow 0$  a.s. for  $k \rightarrow \infty$  with rate arbitrarily close to  $k^{-\beta}$ , any  $\beta < \left\{ \frac{1}{2\rho} - \frac{1}{\sigma}, \frac{3}{2\rho} - 1 \right\}$ . Choosing  $\sigma$  large enough, we obtain the desired rate.

Using the local Lipschitz property of the Itô Lyons map (cf. [6, Theorem 10.26]), we conclude that the Wong-Zakai convergence rate is faster than

$$k^{-(\frac{3}{2\rho}-1-\delta)}$$

for any  $\delta > 0$  (but not for  $\delta = 0$ ). □

**Remark 5.2.** For  $\rho \in (1, \frac{3}{2})$ , the rate above is not optimal. In fact, the sharp rate in this case is ‘almost’  $\frac{1}{\rho} - \frac{1}{2}$ , as shown in [4]. The reason for the non-optimality of the rate is that we obtain the third level estimate merely based on the first two levels, which leads to a reduction in the exponent in the rate. On the other hand, this method does not use any Gaussian structure on the third level, and can be applied to more general processes. For the case  $\rho = 1$ , we recover the sharp rate of ‘almost’  $\frac{1}{2}$ .

## 5.2 The stochastic heat equation

In the theory of stochastic partial differential equations (SPDEs), one typically considers the SPDE as an evolution equation in a function space. When it comes to the question of time and space regularity of the solution, one discovers that they will depend on the particular choice of this space. As a rule of thumb, the smaller the space, the lower the time regularity ([9], Section 5.1). The most prominent examples of such spaces are Hilbert spaces, typically Sobolev spaces. However, in some cases, it can be useful to choose rough paths spaces instead ([10]). A natural question now is whether the known regularity results for Hilbert spaces are also true for rough paths spaces. In this section, we study the example of a modified stochastic heat equation for which we can give a positive answer.

Consider the stochastic heat equation:

$$d\psi = (\partial_{xx} - 1)\psi dt + \sigma dW \tag{5.3}$$

where  $\sigma$  is a positive constant, the spatial variable  $x$  takes values in  $[0, 2\pi]$ ,  $W$  is space-time white noise, i.e. a standard cylindrical Wiener process on  $L^2([0, 2\pi], \mathbb{R}^d)$ , and  $\psi$  denotes the stationary solution with values in  $\mathbb{R}^d$ . The solution  $\psi$  is known to be almost  $\frac{1}{4}$ -Hölder continuous in time and almost  $\frac{1}{2}$ -Hölder continuous in space (cf. [9]). In the next Theorem, we show that this is also the case if we consider the solution as an evolution equation in the appropriate rough paths space.

**Theorem 5.3.** Let  $p > 2$ . Then, for any fixed  $t \geq 0$ , the process  $x \mapsto \psi_t(x)$  is a Gaussian process (in space) which can be lifted to an enhanced Gaussian process  $\Psi_t(\cdot)$ , a process with sample paths in  $C^{0,p\text{-var}}([0, 2\pi], G^{\lfloor p \rfloor}(\mathbb{R}^d))$ . Moreover,  $t \mapsto \Psi_t(\cdot)$  has a Hölder continuous modification (which we denote by the same symbol). More precisely, for every  $\alpha \in (0, \frac{1}{4} - \frac{1}{2p})$ , there exists a (random) constant  $C$  such that

$$\varrho_{p\text{-var}}(\Psi_s, \Psi_t) \leq C|t - s|^\alpha$$

holds almost surely for all  $s < t$ . In particular, choosing  $p$  large gives a time regularity of almost  $\frac{1}{4}$ -Hölder.

*Proof.* The fact that  $x \mapsto \psi_t(x)$  can be lifted to a process with rough sample paths and that there is some Hölder-continuity in time was shown in Lemma 3.1 in [10], see also [2]. We quickly repeat the argument and show where we can use our results in order to derive the exact Hölder exponents. Using the standard Fourier basis

$$e_k(x) = \begin{cases} \frac{1}{\sqrt{\pi}} \sin(kx) & \text{if } k > 0 \\ \frac{1}{\sqrt{2\pi}} & \text{if } k = 0 \\ \frac{1}{\sqrt{\pi}} \cos(kx) & \text{if } k < 0 \end{cases}$$

the equation (5.3) can be rewritten as a system of SDEs

$$dY_t^k = -(k^2 + 1)Y_t^k dt + \sigma dW_t^k$$

where  $(W^k)_{k \in \mathbb{Z}}$  is a collection of independent standard Brownian motions and  $(Y^k)_{k \in \mathbb{Z}}$  are the stationary solutions of the SDEs, i.e. a collection of centered, independent, stationary Ornstein-Uhlenbeck processes. The solution of (5.3) is thus given by the infinite sum  $\psi_t(x) = \sum_{k \in \mathbb{Z}} Y_t^k e_k(x)$ . One can easily see that

$$E[\psi_s(x) \otimes \psi_t(y)] = \frac{\sigma^2}{4\pi} \sum_{k \in \mathbb{Z}} \frac{\cos(k(x-y))}{1+k^2} e^{-(1+k^2)|t-s|} \times I_d$$

where  $I_d$  denotes the identity matrix in  $\mathbb{R}^{d \times d}$ . In particular, for  $s = t$ ,

$$E[\psi_t(x) \otimes \psi_t(y)] = K(x-y) \times I_d$$

where  $K$  is given by

$$K(x) = \frac{\sigma^2}{4 \sinh(\pi)} \cosh(|x| - \pi)$$

for  $x \in [-\pi, \pi]$  and extended periodically for the remaining values of  $x$  (this can be derived by a Fourier expansion of the function  $x \mapsto \cosh(|x| - \pi)$ ). In particular, one can calculate that  $x \mapsto \psi_t(x)$  is a Gaussian process with covariance of finite 1-variation (see the remark at the end of the section for this fact), hence  $\psi_t$  can be lifted to process  $\Psi_t$  with sample paths in the rough paths space  $C^{0,p\text{-var}}([0, 2\pi], G^{\lfloor p \rfloor}(\mathbb{R}^d))$  for any  $p > 2$ .

Furthermore, for any  $s < t$ ,  $x \mapsto (\psi_s(x), \psi_t(x))$  is a Gaussian process which fulfils the assumptions of Theorem 1.1 and the covariance  $R_{(\psi_s, \psi_t)}$  also has finite 1-variation, uniformly bounded for all  $s < t$ , hence

$$\sup_{s < t} |R_{(\psi_s, \psi_t)}|_{1\text{-var}; [0, 2\pi]^2} =: c_1 < \infty.$$

Therefore, for any  $\beta \in (1, p/2)$  and  $q \geq 1$  there is a constant  $C = C(p, \beta, c_1, q)$  such that

$$|\varrho_{p\text{-var}}(\Psi_s, \Psi_t)|_{L^q} \leq C \sup_{x \in [0, 2\pi]} |\psi_t(x) - \psi_s(x)|_{L^2}^\beta$$

holds for all  $s < t$ . A straightforward calculation (cf. [10, Lemma 3.1]) shows that

$$|\psi_t(x) - \psi_s(x)|_{L^2} \leq c_2 |t - s|^{1/4}$$

for a constant  $c_2$ .

Since  $C^{0,p\text{-var}}$  is a Polish space, we can apply the usual Kolmogorov continuity criterion to conclude.  $\square$

**Remark 5.4.** We emphasize that here, for every fixed  $t$ , the process  $\psi_t(\cdot)$  is a Gaussian process, where the spatial variable  $x$  should now be viewed as 'time'. This idea is due to M.Hairer. Knowing that the spatial regularity is 'almost'  $1/2$  for every fixed time  $t$ , one could guess that covariance of this spatial Gaussian process has finite 1-variation. For a formal calculation, we refer to [9] or [2].

## 6 Appendix

The next Lemma is a slight modification of [6, Theorem A.13]. The proof follows the ideas of [2, Theorem 3.1].

**Lemma 6.1** (Kolmogorov for multiplicative functionals). *Let  $\mathbf{X}, \mathbf{Y}: [0, T] \times \Omega \rightarrow T^N(V)$  be random multiplicative functionals and assume that  $\mathbf{X}(\omega)$  and  $\mathbf{Y}(\omega)$  are continuous for all  $\omega \in \Omega$ . Let  $\beta, \delta \in (0, 1]$  and choose  $\beta' < \beta$  and  $\delta' < \delta$ . Assume that there is a constant  $M > 0$  such that*

$$\begin{aligned} |\mathbf{X}_{s,t}^n|_{L^{q/n}} &\leq M^n |t-s|^{n\beta} \\ |\mathbf{Y}_{s,t}^n|_{L^{q/n}} &\leq M^n |t-s|^{n\beta} \\ |\mathbf{X}_{s,t}^n - \mathbf{Y}_{s,t}^n|_{L^{q/n}} &\leq M^n \epsilon |t-s|^{\delta+(n-1)\beta} \end{aligned}$$

hold for all  $s < t \in [0, T]$  and  $n = 1, \dots, N$  where  $\epsilon$  is a positive constant and  $q \geq q_0$  where

$$q_0 := 1 + \left( \frac{1}{\beta - \beta'} \vee \frac{1}{\delta - \delta'} \right).$$

Then there is a constant  $C = C(N, \beta, \beta', \delta, \delta')$  such that

$$\left| \sup_{s < t \in [0, T]} \frac{|\mathbf{X}_{s,t}^n|}{|t-s|^{n\beta'}} \right|_{L^{\frac{q}{n}}} \leq CM^n \tag{6.1}$$

$$\left| \sup_{s < t \in [0, T]} \frac{|\mathbf{Y}_{s,t}^n|}{|t-s|^{n\beta'}} \right|_{L^{\frac{q}{n}}} \leq CM^n \tag{6.2}$$

$$\left| \sup_{s < t \in [0, T]} \frac{|\mathbf{X}_{s,t}^n - \mathbf{Y}_{s,t}^n|}{|t-s|^{\delta'+(n-1)\beta'}} \right|_{L^{\frac{q}{n}}} \leq CM^n \epsilon \tag{6.3}$$

hold for all  $n = 1, \dots, N$ .

*Proof.* W.l.o.g., we may assume  $T = 1$ . Let  $(D_k)_{k \in \mathbb{N}}$  be the sequence of dyadic partitions of the interval  $[0, 1)$ , i.e.  $D_k = \{ \frac{l}{2^k} : l = 0, \dots, 2^k - 1 \}$ . Clearly,  $|D_k| = \frac{1}{\#D_k} = 2^{-k}$ . Set

$$\begin{aligned} K_{k,X}^n &:= \max_{t_i \in D_k} |\mathbf{X}_{t_i, t_{i+1}}^n| \\ K_{k,Y}^n &:= \max_{t_i \in D_k} |\mathbf{Y}_{t_i, t_{i+1}}^n| \\ K_{k,X-Y}^n &:= \frac{1}{\epsilon} \max_{t_i \in D_k} |\mathbf{X}_{t_i, t_{i+1}}^n - \mathbf{Y}_{t_i, t_{i+1}}^n| \end{aligned}$$

for  $n = 1, \dots, N$  and  $k \in \mathbb{N}$ . By assumption, we have

$$E|K_{k,X}^n|_{\frac{q}{n}} \leq E \sum_{t_i \in D_k} |\mathbf{X}_{t_i, t_{i+1}}^n|_{\frac{q}{n}} \leq \#D_k \max_{t_i \in D_k} E|\mathbf{X}_{t_i, t_{i+1}}^n|_{\frac{q}{n}} \leq M^q |D_k|^{q\beta-1}.$$

In the same way one estimates  $K_{k,Y}^n$  and  $K_{k,X-Y}^n$ , hence

$$|K_{k,X}^n|_{L^{q/n}} \leq M^n |D_k|^{n\beta-n/q} \tag{6.4}$$

$$|K_{k,Y}^n|_{L^{q/n}} \leq M^n |D_k|^{n\beta-n/q} \tag{6.5}$$

$$|K_{k,X-Y}^n|_{L^{q/n}} \leq M^n |D_k|^{\delta+(n-1)\beta-n/q}. \tag{6.6}$$

Note the following fact: For any dyadic rationals  $s < t$ , i.e.  $s < t \in \Delta := \bigcup_{k=1}^\infty D_k$ , there is a  $m \in \mathbb{N}$  such that  $|D_{m+1}| < |t - s| \leq |D_m|$  and a partition

$$s = \tau_0 < \tau_1 < \dots < \tau_N = t \tag{6.7}$$

of the interval  $[s, t]$  with the property that for any  $i = 0, \dots, N - 1$  there is a  $k \geq m + 1$  with  $[\tau_i, \tau_{i+1}) \in D_k$ , but for fixed  $k \geq m + 1$  there are at most two such intervals contained in  $D_k$ .

**Step 1:** We claim that for every  $n = 1, \dots, N$  there is a real random variable  $K_X^n$  such that  $|K_X^n|_{L^{q/n}} \leq M^n c$  where  $c = c(\beta, \beta', \delta, \delta')$  and that for any dyadic rationals  $s < t$  and  $m, (\tau_i)_{i=0}^N$  chosen as in (6.7) we have

$$\sum_{i=0}^{N-1} \frac{|\mathbf{X}_{\tau_i, \tau_{i+1}}^n|}{|t - s|^{n\beta'}} \leq K_X^n. \tag{6.8}$$

Furthermore, the estimate (6.8) also holds for  $\mathbf{Y}^n$  and a random variable  $K_Y^n$ . Indeed: By the choice of  $m$  and  $(\tau_i)_{i=0}^N$ ,

$$\sum_{i=0}^{N-1} \frac{|\mathbf{X}_{\tau_i, \tau_{i+1}}^n|}{|t - s|^{n\beta'}} \leq \sum_{k=m+1}^\infty \frac{2K_{k,X}^n}{|D_{m+1}|^{n\beta'}} \leq 2 \sum_{k=m+1}^\infty \frac{K_{k,X}^n}{|D_k|^{n\beta'}} \leq 2 \sum_{k=1}^\infty \frac{K_{k,X}^n}{|D_k|^{n\beta'}} =: K_X^n.$$

It remains to prove that  $|K_X^n|_{L^{q/n}} \leq M^n c$ . By the triangle inequality and the estimate (6.4),

$$\left| \sum_{k=1}^\infty \frac{K_{k,X}^n}{|D_k|^{n\beta'}} \right|_{L^{q/n}} \leq M^n \sum_{k=1}^\infty |D_k|^{n(\beta-1/q-\beta')} \leq M^n \sum_{k=1}^\infty |D_k|^{(\beta-1/q_0-\beta')} < \infty$$

since  $\beta - 1/q_0 - \beta' > 0$  which shows the claim.

**Step 2:** We show that (6.1) and (6.2) hold for all  $n = 1, \dots, N$ . It is enough to consider  $\mathbf{X}$ . Note first that, due to continuity, it is enough to show the estimate for  $\sup_{s < t \in \Delta} \frac{|\mathbf{X}_{s,t}^n|}{|t-s|^{n\beta'}}$ . By induction over  $n$ : For  $n = 1$ , this just follows from the usual Kolmogorov continuity criterion. Assume that the estimate is proven up to level  $n - 1$ . Let  $s < t$  be any dyadic rationals and choose  $m$  and  $(\tau_i)_{i=0}^N$  as in (6.7). Since  $\mathbf{X}$  is a multiplicative functional,

$$|\mathbf{X}_{s,t}^n| \leq \sum_{i=0}^{N-1} |\mathbf{X}_{\tau_i, \tau_{i+1}}^n| + \sum_{l=1}^{n-1} \max_{i=1, \dots, N} |\mathbf{X}_{s, \tau_i}^{n-l}| \sum_{i=0}^{N-1} |\mathbf{X}_{\tau_i, \tau_{i+1}}^l|$$

and thus, using step 1,

$$\begin{aligned} \frac{|\mathbf{X}_{s,t}^n|}{|t-s|^{n\beta'}} &\leq \sum_{i=0}^{N-1} \frac{|\mathbf{X}_{\tau_i, \tau_{i+1}}^n|}{|t-s|^{n\beta'}} + \sum_{l=1}^{n-1} \sup_{u < v \in \Delta} \frac{|\mathbf{X}_{u,v}^{n-l}|}{|v-u|^{(n-l)\beta'}} \sum_{i=0}^{N-1} \frac{|\mathbf{X}_{\tau_i, \tau_{i+1}}^l|}{|t-s|^{l\beta'}} \\ &\leq K_X^n + \sum_{l=1}^{n-1} \sup_{u < v \in \Delta} \frac{|\mathbf{X}_{u,v}^{n-l}|}{|v-u|^{(n-l)\beta'}} K_X^l. \end{aligned}$$

We can now take the supremum over all  $s < t \in \Delta$  on the left. Taking the  $L^{q/n}$ -norm on both sides, using first the triangle, then the Hölder inequality and the estimates from step 1 together with the induction hypothesis gives the claim.

**Step 3:** As in step 1, we claim that for any  $n = 1, \dots, N$  there is a random variable  $K_{X-Y}^n \in L^{q/n}$  such that for any dyadic rationals  $s < t$  and  $m, (\tau_i)_{i=0}^N$  chosen as above we have

$$\sum_{i=0}^{N-1} \frac{|\mathbf{X}_{\tau_i, \tau_{i+1}}^n - \mathbf{Y}_{\tau_i, \tau_{i+1}}^n|}{|t-s|^{\delta'+(n-1)\beta'}} \leq K_{X-Y}^n \epsilon. \tag{6.9}$$

Furthermore, we claim that  $|K_{X-Y}^n|_{L^{q/n}} \leq M^n \tilde{c}$  where  $\tilde{c} = \tilde{c}(\beta, \beta', \delta, \delta')$ . The proof follows the lines of step 1, setting

$$\frac{1}{\epsilon} \sum_{i=0}^{N-1} \frac{|\mathbf{X}_{\tau_i, \tau_{i+1}}^n - \mathbf{Y}_{\tau_i, \tau_{i+1}}^n|}{|t-s|^{\delta'+(n-1)\beta'}} \leq 2 \sum_{k=1}^{\infty} \frac{K_{k, X-Y}^n}{|D_k|^{\delta'+(n-1)\beta'}} =: K_{X-Y}^n.$$

**Step 4:** We prove that (6.3) holds for all  $n = 1, \dots, N$ . By induction over  $n$ : The case  $n = 1$  is again just the usual Kolmogorov continuity criterion applied to  $t \mapsto \epsilon^{-1}(X_t - Y_t)$ . Assume the assertion is shown up to level  $n - 1$  and chose two dyadic rationals  $s < t$ . Using the multiplicative property, we have

$$\begin{aligned} |\mathbf{X}_{s,t}^n - \mathbf{Y}_{s,t}^n| &\leq \sum_{i=0}^{N-1} |\mathbf{X}_{\tau_i, \tau_{i+1}}^n - \mathbf{Y}_{\tau_i, \tau_{i+1}}^n| + \sum_{l=1}^{n-1} \max_{i=1, \dots, N} |\mathbf{X}_{s, \tau_i}^{n-l}| \sum_{i=0}^{N-1} |\mathbf{X}_{\tau_i, \tau_{i+1}}^l - \mathbf{Y}_{\tau_i, \tau_{i+1}}^l| \\ &\quad + \sum_{l=1}^{n-1} \max_{i=1, \dots, N} |\mathbf{X}_{s, \tau_i}^{n-l} - \mathbf{Y}_{s, \tau_i}^{n-l}| \sum_{i=0}^{N-1} |\mathbf{Y}_{\tau_i, \tau_{i+1}}^l|. \end{aligned}$$

Now we proceed as in step 2, using the estimates from step 1 to step 3 and the induction hypothesis. □

## References

- [1] L. Coutin and Z. Qian, Stochastic analysis, rough path analysis and fractional Brownian motions, *Probability Theory and Related Fields*, Volume 122, No.1, 2002, Pages 108-140. MR-1883719
- [2] P. Friz and M. Hairer, A Short Course on Rough Paths, preprint.
- [3] P. Friz and S. Riedel, Convergence rates for the full Brownian rough paths with applications to limit theorems for stochastic flows, *Bulletin des Sciences Mathematics (proceeding in memory of P. Malliavin)*, Volume 135, Issues 6-7, September-November 2011, Pages 613-628, ISSN 0007-4497. MR-2838093
- [4] P. Friz and S. Riedel, Convergence rates for the full Gaussian rough paths, to appear in *Annales de l'Institut Henri Poincaré (B) Probability and Statistics*.
- [5] P. Friz and N. Victoir, Differential Equations driven by Gaussian Signals, *Annales de l'Institut Henri Poincaré (B) Probability and Statistics*, Vol. 46, No. 2, 369-413, May 2010. MR-2667703
- [6] P. Friz and N. Victoir, Multidimensional Stochastic Processes as Rough Paths, *Cambridge Studies in Advanced Mathematics*, Cambridge University Press, 2010. MR-2604669
- [7] P. Friz and N. Victoir, A note on higher dimensional  $p$ -variation, *Electronic Journal of Probability*, Vol. 16, 2011. MR-2842090
- [8] M. Gubinelli, Controlling rough paths, *Journal of Functional Analysis*, Volume 216, No. 1, pages 86-140, 2004. MR-2091358

- [9] M. Hairer, An introduction to Stochastic PDEs, preprint arXiv:0907.4178, 2009.
- [10] M. Hairer, Rough stochastic PDEs, *Communications on Pure and Applied mathematics*, Volume 64, Issue 11, pages 1547-1585, November 2011. MR-2832168
- [11] T. Lyons, Differential equations driven by rough signals, *Rev. Mat. Iberoamericana*, Volume 14, No.2, pages 215-310, 1998. MR-1654527
- [12] T. Lyons, M. Caruana and T. Lévy, Differential Equations Driven by Rough Paths, *Lecture Notes in Mathematics* Volume 1908, Springer Berlin, 2006.
- [13] T. Lyons and Z. Qian, System Control and Rough Paths, *Oxford Mathematical Monographs* Oxford University Press, 2002. MR-2036784
- [14] T. Lyons and W. Xu, A uniform estimate for rough paths, preprint, available at <http://arxiv.org/abs/1110.5278>, 2011. MR-3116217
- [15] T. Lyons and O. Zeitouni, Conditional exponential moments for iterated Wiener integrals *Annals of Probability*, Volume 27, No.4, 1999, Pages 1728-1749. MR-1742886
- [16] N. Towghi, Multidimensional extension of L.C. Young's inequality, *JIPAM J. Inequal. Pure Appl. Math.*, 3(2): Article 22, 13 pp. (electronic), 2002. MR-1906391

**Acknowledgments.** We wish to thank our advisors, Peter Friz and Terry Lyons, for their helpful discussions and support during the project. S. Riedel is supported by the Berlin Mathematical School (BMS) and by the European Research Council under the European Union's Seventh Framework Programme ERC grant agreement nr. 258237. W. Xu is supported by the Oxford-Man Institute Scholarship. We also thank the referee for carefully reading through the manuscript and providing valuable suggestions.