Annales de l'Institut Henri Poincaré - Probabilités et Statistiques
2010, Vol. 46, No. 2, 369–413
DOI: 10.1214/09-AIHP202

ANNALES DE L'INSTITUT HENRI POINCARÉ
PROBABILITÉS ET STATISTIQUES
www.imstat.org/aihp

Differential equations driven by Gaussian signals

Peter Friz[a,1] and Nicolas Victoir

[a]*Department of Pure Mathematics and Mathematical Statistics, University of Cambridge, UK. E-mail:* P.K.Friz@statslab.cam.ac.uk

Received 25 June 2007; revised 30 July 2008; accepted 19 March 2009

Abstract. We consider multi-dimensional Gaussian processes and give a new condition on the covariance, simple and sharp, for the existence of Lévy area(s). Gaussian rough paths are constructed with a variety of weak and strong approximation results. Together with a new RKHS embedding, we obtain a powerful – yet conceptually simple – framework in which to analyze differential equations driven by Gaussian signals in the rough paths sense.

Résumé. Nous donnons une condition simple et optimale sur la covariance d'un processus gaussien pour que celui-ci puisse être associé naturellement à un rough path. Une fois ce processus construit, nous démontrons un principe de grandes déviations, un théorème du support, et plusieurs résultats d'approximations. Avec la théorie des rough paths de T. Lyons, nous obtenons ainsi un cadre puissant, bien que conceptuellement simple, dans lequel nous pouvons analyser les équations différentielles conduites par des signaux gaussiens dans le sens des rough paths.

MSC: 60G15; 60H99

Keywords: Rough paths; Gaussian processes

---

1. Introduction

Let X be a real-valued centered Gaussian process on $[0, 1]$ with continuous sample paths and (continuous) covariance $R = R(s, t) = \mathbb{E}(X_s X_t)$. From Kolmogorov's criterion, it is clear that Hölder regularity of R will imply Hölder continuity of sample paths. One can also deduce p-variation of sample paths from R. Indeed, the condition

$$\sup_{D=\{t_i\}} \sum_i \bigl|\mathbb{E}\bigl[(X_{t_{i+1}} - X_{t_i})^2\bigr]\bigr|^\rho < \infty \tag{1}$$

implies that X has sample paths of finite p-variation for $p > 2\rho$, see [22] or the survey [13]. Note that (1) can be written in terms of R and expresses some sort of "on diagonal ρ-variation" regularity of R.

The results of this paper put forward the notion of *genuine* ρ-variation regularity of R as a function on $[0, 1]^2$ as novel and, perhaps, fundamental quantity related to Gaussian processes. Similar to (1), finite ρ-variation of R, in symbols $R \in C^{\rho\text{-var}}([0, 1]^2, \mathbb{R})$, can be expressed in terms of the associated Gaussian process and amounts to say that

$$\sup_D \sum_{i,j} \bigl|\mathbb{E}\bigl[(X_{t_{i+1}} - X_{t_i})(X_{t_{j+1}} - X_{t_j})\bigr]\bigr|^\rho < \infty.$$

[1] Supported in part by a Leverhulme Research Fellowship and EPSRC Grant EP/E048609/1.

The notion of (2D) ρ-variation of the covariance leads naturally to

$$\mathcal{H} \hookrightarrow C^{\rho\text{-var}}\big([0,1],\mathbb{R}\big), \tag{2}$$

an embedding of the *Cameron–Martin* or *reproducing kernel Hilbert space (RKHS)* $\mathcal{H}$ into the space of continuous path with finite ρ-variation. Good examples to have in mind are standard Brownian Motion with $\rho = 1$ and fractional Brownian Motion with Hurst parameter $H \in (0, 1/2]$ for which $\rho = 1/(2H)$. We then consider a d-dimensional, continuous, centered Gaussian process with independent components,

$$X = \big(X^1, \ldots, X^d\big),$$

with respective covariances $R_1, \ldots, R_d \in C^{\rho\text{-var}}$ and ask under what conditions there exists an a.s. well-defined lift to a geometric rough path $\mathbf{X}$ in the sense of T. Lyons; [21,25–27]. (This amounts, first and foremost, to define Lévy's area and higher iterated integrals of X, and to establish subtle regularity properties.) The answer to this question is the sufficient (and essentially necessary) condition

$$\rho \in [1, 2)$$

under which there exists a lift of X to a *Gaussian* geometric p-rough path $\mathbf{X}$ (short: *Gaussian rough path*) for any $p > 2\rho$. For fractional Brownian Motion this requires $H > 1/4$ which is optimal [10,11] and our condition is seen to be sharp[2]. Recall that geometric p-rough paths are (limits of) paths together with their first $[p]$-iterated integrals. Assuming $\rho < 2$ one can (and should) choose $p < 4$; when X has sufficiently smooth sample paths, $\mathbf{X}_\cdot \equiv S_3(X)$ is then simply given by its coordinates in the three "tensor-levels," $\mathbb{R}^d$, $\mathbb{R}^d \otimes \mathbb{R}^d$ and $\mathbb{R}^d \otimes \mathbb{R}^d \otimes \mathbb{R}^d$, obtained by iterated integration

$$\mathbf{X}^i_\cdot = \int_0^\cdot \mathrm{d}X^i_r, \qquad \mathbf{X}^{i,j}_\cdot = \int_0^\cdot \int_0^s \mathrm{d}X^i_r \mathrm{d}X^j_s, \qquad \mathbf{X}^{i,j,k}_\cdot = \int_0^\cdot \int_0^t \int_0^s \mathrm{d}X^i_r\, \mathrm{d}X^j_s\, \mathrm{d}X^k_t,$$

with indices $i, j, k \in \{1, \ldots, d\}$. Our condition $\rho < 2$ is then easy to explain. Assuming $X_0 = 0$ and $i \neq j$, which is enough to deal with the second tensor level, we have

$$\begin{aligned}\mathbb{E}\big(\big|\mathbf{X}^{i,j}_t\big|^2\big) &= \int_{[0,t]^2} R_i(u,v) \frac{\partial^2}{\partial u\, \partial v} R_j(u,v)\, \mathrm{d}u\, \mathrm{d}v \\ &\equiv \int_{[0,t]^2} R_i(u,v)\, \mathrm{d}R_j(u,v).\end{aligned}$$

The integral which appears on the right-hand side above is a 2-dimensional (short: 2D) Young integral. It remains meaningful provided R_i, R_j have finite ρ_i resp. ρ_j-variation with $\rho_i^{-1} + \rho_j^{-1} > 1$. In particular, if R_i, R_j have both finite ρ-variation this condition reads $\rho < 2$. is required. The ρ-variation condition on the covariance encodes some decorrelation of the increments and this is the (partial) nature of the so-called (h, p)-long time memory condition that appears in [27] resp. Coutin–Qian's condition [10] which is seen to be more restrictive than our ρ-variation condition.

Let us briefly state our main continuity result for Gaussian rough paths, taken from Section 4.4.

Theorem 1. *Let $X = (X^1, \ldots, X^d)$, $Y = (Y^1, \ldots, Y^d)$ be two continuous, centered jointly Gaussian processes defined on $[0,1]$ such that (X^i, Y^i) is independent of (X^j, Y^j) when $i \neq j$. Let $\rho \in [1,2)$ and assume the covariance of (X, Y) is of finite ρ-variation,*

$$|R_{(X,Y)}|_{\rho\text{-var};[0,1]^2} \le K < \infty.$$

Let $p > 2\rho$ and $\mathbf{X}, \mathbf{Y}$ denote the natural lift of X, Y to a Gaussian rough path. Then there exist positive constants $\theta = \theta(p, \rho)$ and $C = C(p, \rho, K)$ such that for all $q \in [1, \infty)$,

$$\big|d_{p\text{-var}}(\mathbf{X}, \mathbf{Y})\big|_{L^q} \le C\sqrt{q}\,|R_{X-Y}|^{\theta}_{\infty;[0,1]^2}.$$

[2]From [22] and [30] we expect that logarithmic refinements of this condition are possible but we shall not pursue this here.

The natural lift to a Gaussian rough path is easily explained along the above estimates: take a continuous, centered d-dimensional process Z with independent components and finite $\rho \in [1,2)$-covariance and consider its piecewise linear approximations Z^n. Applying the above estimate to $\mathbf{X} = S_3(Z^n), \mathbf{Y} = S_3(Z^m)$, identifies $S_3(Z^n)$ as Cauchy sequence and we call the limit natural lift of Z. In conjunction with the *universal limit theorem* [27], i.e. the continuous dependence of solutions to (rough) differential equations of the driving signal $\mathbf{X}$ w.r.t. $d_{p\text{-var}}$, the above theorem contains a collection of powerful limit theorems which cover, for instance, piecewise linear and mollifier approximations to Stratonovich SDEs as special case. As further consequence, weak convergence results are obtained. For instance, differential equations driven by fractional Brownian Motion with Hurst parameter $H \to 1/2$ converge to the corresponding Stratonovich SDEs.

We further note that a large deviation principle holds in the present generality; thanks to the Cameron–Martin embedding (2) this follows immediately from the main result in [19]. Moreover, we shall see that in the same generality, approximations based on the L^2- or *Karhunen–Loève expansion*

$$X^i(t,\omega) = \sum_{k\in\mathbb{N}} Z_k^i(\omega) h^{i,k}(t) \tag{3}$$

converge in rough path topology to our natural lift $\mathbf{X}$. As corollary, we obtain a *support theorem* i.e. we characterize the support of $\mathbf{X}$ as closure of $S_3(\mathcal{H})$ in suitable rough path topology. The embedding (2) is absolutely crucial for these purposes: given $\rho < 2$ it tells us that elements in $\mathcal{H}$ (and in particular, Karhunen–Loève approximations which are finite sums of form (3)) admit canonicially defined second and third iterated integrals.

The lift of certain Gaussian processes including fractional Brownian Motion with Hurst parameter $H > 1/4$ is due to Coutin–Qian, [10]. A large deviation principle for the lift of fractional Brownian Motion was obtained in [28], for the Coutin–Qian class in [19]. Support statements for lifted fractional Brownian Motion for $H > 1/3$ were obtained in [17,14]; a Karhunen–Loève type approximations for fractional Brownian Motion is studied in [29].

The interest in our results goes beyond the unification and optimal extension of the above-cited articles. It identifies a general framework for differential equations driven by Gaussian signal, surprisingly well-suited for further (Gaussian) analysis: the embedding (2) combined with basic facts of Young integrals shows that, at least for $\rho < 3/2$, translations in $\mathcal{H}$-directions are well enough controlled to exploit the isoperimetric inequality for abstract Wiener spaces; applications towards regularity/integrability statements for stochastic area are discussed in [15]. Relatedly, solutions to (rough) differential equations driven by Gaussian signals are $\mathcal{H}$-differentiable which allows to establish density results using Malliavin calculus, see [7]. A Hörmander-type density result is obtained in [6] and relies on the support theorem.

1.1. *Notations*

Let (E,d) be a metric space and $x \in C([0,1],E)$. It then makes sense to speak of α-Hölder- and p-variation "norms" defined as

$$\|x\|_{\alpha\text{-Höl}} = \sup_{0\le s<t\le 1} \frac{d(x_s,x_t)}{|t-s|^\alpha}, \qquad \|x\|_{p\text{-var}} = \sup_{D=(t_i)} \left(\sum_i d(x_{t_i},x_{t_{i+1}})^p\right)^{1/p}.$$

It also makes sense to speak of a d_∞-distance of two such paths,

$$d_\infty(x,y) = \sup_{0\le t\le 1} d(x_t,y_t).$$

Given a positive integer N the truncated tensor algebra of degree N is given by the direct sum

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

With tensor product $\otimes$, vector addition and usual scalar multiplication, $T^N(\mathbb{R}^d) = (T^N(\mathbb{R}^d), \otimes, +, \cdot)$ is an algebra. Functions such as $\exp, \ln : T^N(\mathbb{R}^d) \to T^N(\mathbb{R}^d)$ are defined immediately by their power-series. Let π_i denote the canonical projection from $T^N(\mathbb{R}^d)$ onto $(\mathbb{R}^d)^{\otimes i}$. Let $p \in [1,2)$ and $x \in C^{p\text{-var}}([0,1],\mathbb{R}^d)$, the space of continuous $\mathbb{R}^d$-valued paths of bounded p-variation. We define $\mathbf{x} \equiv S_N(x) : [0,1] \to T^N(\mathbb{R}^d)$ via iterated (Young) integration,

$$\mathbf{x}_t \equiv S_N(x)_t = 1 + \sum_{i=1}^N \int_{0<s_1<\cdots<s_i<t} \mathrm{d}x_{s_1} \otimes \cdots \otimes \mathrm{d}x_{s_i}$$

noting that $\mathbf{x}_0 = 1 + 0 + \cdots + 0 = \exp(0) \equiv e$, the neutral element for $\otimes$, and that $\mathbf{x}_t$ really takes values in

$$G^N(\mathbb{R}^d) = \{g \in T^N(\mathbb{R}^d) \colon \exists x \in C^{1\text{-var}}([0,1],\mathbb{R}^d) \colon g = S_N(x)_1\},$$

a submanifold of $T^N(\mathbb{R}^d)$ and, in fact, a Lie group with product $\otimes$, called the free step-N nilpotent group with d generators. Because $\pi_1[\mathbf{x}_t] = x_t - x_0$ we say that $\mathbf{x} = S_N(x)$ is the canonical lift of x. There is a canonical notion of increments, $\mathbf{x}_{s,t} := \mathbf{x}_s^{-1} \otimes \mathbf{x}_t$. The dilation operator $\delta : \mathbb{R} \times G^N(\mathbb{R}^d) \to G^N(\mathbb{R}^d)$ is defined by

$$\pi_i(\delta_\lambda(g)) = \lambda^i \pi_i(g), \qquad i = 0, \dots, N,$$

and a continuous norm on $G^N(\mathbb{R}^d)$, homogenous with respect to δ, the *Carnot–Caratheodory norm*, is given

$$\|g\| = \inf\{\text{length}(x) \colon x \in C^{1\text{-var}}([0,1],\mathbb{R}^d), S_N(x)_1 = g\}.$$

It is symmetric, sub-additive in the sense that $\|g\| = \|g^{-1}\|$, $\|g \otimes g'\| \le \|g\| + \|g'\|$ respectively. By equivalence of continuous, homogenous norms there exists a constant K_N such that

$$\frac{1}{K_N} \max_{i=1,\dots,N} |\pi_i(g)|^{1/i} \le \|g\| \le K_N \max_{i=1,\dots,N} |\pi_i(g)|^{1/i}.$$

The norm $\|\cdot\|$ induces a (left-invariant) metric on $G^N(\mathbb{R}^d)$ known as *Carnot–Caratheodory metric*, $d(g,h) := \|g^{-1} \otimes h\|$. Now let $x, y \in C_0([0,1], G^N(\mathbb{R}^d))$, the space of continuous $G^N(\mathbb{R}^d)$-valued paths started at the neutral element $\exp(0) = e$. We define α-Hölder- and p-variation distances

$$d_{\alpha\text{-Höl}}(\mathbf{x},\mathbf{y}) = \sup_{0 \le s < t \le 1} \frac{d(\mathbf{x}_{s,t}, \mathbf{y}_{s,t})}{|t-s|^\alpha},$$

$$d_{p\text{-var}}(\mathbf{x},\mathbf{y}) = \sup_{D=(t_i)} \left(\sum_i d(\mathbf{x}_{t_i,t_{i+1}}, \mathbf{y}_{t_i,t_{i+1}})^p \right)^{1/p}$$

and also the "0-Hölder" distance, locally $1/N$-Hölder equivalent to $d_\infty(\mathbf{x},\mathbf{y})$,

$$d_0(\mathbf{x},\mathbf{y}) = d_{0\text{-Höl}}(\mathbf{x},\mathbf{y}) = \sup_{0 \le s < t \le 1} d(\mathbf{x}_{s,t}, \mathbf{y}_{s,t}).$$

Note that $d_{\alpha\text{-Höl}}(\mathbf{x},0) = \|\mathbf{x}\|_{\alpha\text{-Höl}}$, $d_{p\text{-var}}(\mathbf{x},0) = \|\mathbf{x}\|_{p\text{-var}}$ where 0 denotes the constant path $\exp(0)$, or in fact, any constant path. The following path spaces will be useful to us:

(i) $C_0^{p\text{-var}}([0,1], G^N(\mathbb{R}^d))$: the set of continuous functions $\mathbf{x}$ from $[0,1]$ into $G^N(\mathbb{R}^d)$ such that $\|\mathbf{x}\|_{p\text{-var}} < \infty$ and $\mathbf{x}_0 = \exp(0)$.

(ii) $C_0^{0,p\text{-var}}([0,1], G^N(\mathbb{R}^d))$: the $d_{p\text{-var}}$-closure of

$$\{S_N(x), x : [0,1] \to \mathbb{R}^d \text{ smooth}\}.$$

(iii) $C_0^{1/p\text{-Höl}}([0,1], G^N(\mathbb{R}^d))$: the set of continuous functions $\mathbf{x}$ from $[0,1]$ into $G^N(\mathbb{R}^d)$ such that $d_{1/p\text{-Höl}}(0,\mathbf{x}) < \infty$ and $\mathbf{x}_0 = \exp(0)$.

(iv) $C_0^{0,1/p\text{-Höl}}([0,1], G^N(\mathbb{R}^d))$: the $d_{1/p\text{-Höl}}$-closure of

$$\{S_n(x), x : [0,1] \to \mathbb{R}^d \text{ smooth}\}.$$

Recall that a geometric p-rough path is an element of $C_0^{0,p\text{-var}}([0,1],G^{[p]}(\mathbb{R}^d))$, and a weak geometric rough path is an element of $C_0^{p\text{-var}}([0,1],G^{[p]}(\mathbb{R}^d))$. For a detailed study of these spaces and their properties the reader is referred to [17].

2. 2D Young integral

2.1. *On 2D $\boldsymbol{\rho}$-variation*

For a function f from $[0,1]^2$ into a Banach space $(\mathcal{B},|\cdot|)$ we will use the notation

$$f\begin{pmatrix} s & u \\ t & v \end{pmatrix} := f(s,u)+f(t,v)-f(s,v)-f(t,u).$$

If f is the 2D distribution function of a signed measure on $[0,1]^2$ this is precisely the measure of the rectangle $(s,t]\times(u,v]$. If $f(s,t)=\mathbb{E}(X_s,X_t)\in\mathbb{R}$ for some real-valued stochastic process X, then

$$f\begin{pmatrix} s & u \\ t & v \end{pmatrix} = \mathbb{E}(X_{s,t}X_{u,v}).$$

A similar formula holds when $f(s,t)=\mathbb{E}(X_s\otimes X_t)\in\mathbb{R}^d\otimes\mathbb{R}^d$ (which we equip with its canonical Euclidean structure).

Definition 2. *Let $f:[0,1]^2\to(\mathcal{B},|\cdot|)$. We say that f has finite ρ-variation if $|f|_{\rho\text{-var},[0,1]^2}<\infty$, where*[3]

$$|f|_{\rho\text{-var},[s,t]\times[u,v]} = \sup_{\substack{D=(t_i)\ \textit{subdivision of}\ [s,t] \\ D'=(t'_j)\ \textit{subdivision of}\ [u,v]}} \left(\sum_{i,j}\left|f\begin{pmatrix} t_i & t'_j \\ t_{i+1} & t'_{j+1}\end{pmatrix}\right|^{\rho}\right)^{1/\rho}.$$

Definition 3. *A 2D control is a map ω from $(s\le t, u\le v)$ such that for all $r\le s\le t$, $u\le v$,*

$$\omega\bigl([r,s]\times[u,v]\bigr)+\omega\bigl([s,t]\times[u,v]\bigr)\le\omega\bigl([r,t]\times[u,v]\bigr),$$
$$\omega\bigl([u,v]\times[r,s]\bigr)+\omega\bigl([u,v]\times[s,t]\bigr)\le\omega\bigl([u,v]\times[r,t]\bigr),$$

and such that $\lim_{s\to t}\omega([s,t]\times[0,1])=\lim_{s\to t}\omega([0,1]\times[s,t])=0$. *Moreover, we will say that the 2D control ω is Hölder-dominated if there exists a constant C such that for all* $0\le s\le t\le 1$

$$\omega\bigl([s,t]^2\bigr)\le C|t-s|.$$

Lemma 4. *Let f be a $(\mathcal{B},|\cdot|)$-valued continuous function on* $[0,1]^2$. *Then*

(i) *If f is of finite ρ-variation for some $\rho\ge 1$,*

$$[s,t]\times[u,v]\mapsto|f|^{\rho}_{\rho\text{-var};[s,t]\times[u,v]}$$

is a 2D control.

[3]This (semi-)norm was also introduced by [36].

(ii) *f is of finite ρ-variation on $[0,1]^2$ if and only if there exists a 2D control ω such that for all $[s,t]\times[u,v]\subset[0,1]^2$,*

$$\left|f\begin{pmatrix} s & u \\ t & v\end{pmatrix}\right|^\rho \le \omega\big([s,t]\times[u,v]\big)$$

and we say that "ω controls the ρ-variation of f."

Proof. Straight-forward. □

Remark 5. *If $f:[0,T]^2\to(\mathcal{B},|\cdot|)$ is symmetric (i.e. $f(x,y)=f(y,x)$ for all x,y) and of finite ρ-variation then $[s,t]\times[u,v]\mapsto|f|^\rho_{\rho\text{-var};[s,t]\times[u,v]}$ is symmetric. In fact, one can always work with symmetric controls, it suffices to replace a given ω with $[s,t]\times[u,v]\mapsto\omega([s,t]\times[u,v])+\omega([u,v]\times[s,t])$.*

Lemma 6. *A continuous function $f:[0,1]^2\to(\mathcal{B},|\cdot|)$ is of finite ρ-variation if and only if*

$$\sup_{D=(t_i)\text{ subdivision of }[0,1]}\left(\sum_{i,j}\left|f\begin{pmatrix} t_i & t_j \\ t_{i+1} & t_{j+1}\end{pmatrix}\right|^\rho\right)^{1/\rho}<\infty.$$

Moreover, the ρ-variation of f is controlled by

$$\omega\big([s,t]\times[u,v]\big):=3^{\rho-1}\sup_{D=(t_i)\text{ subdivision of }[0,1]}\sum_{\substack{i,j\\ [t_i,t_{i+1}]\subset[s,t]\\ [t_j,t_{j+1}]\subset[u,v]}}\left|f\begin{pmatrix} t_i & t_j \\ t_{i+1} & t_{j+1}\end{pmatrix}\right|^\rho.$$

Proof. Assuming that $\omega([0,1]^2)$ is finite, it is easy to check that ω is a 2D control. Then, for any given $[s,t]$ and $[u,v]$ which do not intersect or such that $[s,t]=[u,v]$,

$$\left|f\begin{pmatrix} s & u \\ t & v\end{pmatrix}\right|^\rho \le \omega\big([s,t]\times[u,v]\big).$$

Take now $s\le u\le t\le v$, then,

$$f\begin{pmatrix} s & u \\ t & v\end{pmatrix}=f\begin{pmatrix} s & u \\ u & v\end{pmatrix}+f\begin{pmatrix} u & u \\ t & v\end{pmatrix}=f\begin{pmatrix} s & u \\ u & v\end{pmatrix}+f\begin{pmatrix} u & u \\ t & t\end{pmatrix}+f\begin{pmatrix} s & t \\ u & v\end{pmatrix}.$$

Hence,

$$\begin{aligned}\left|f\begin{pmatrix} s & u \\ t & v\end{pmatrix}\right|^\rho &\le 3^{\rho-1}\big(\omega\big([s,u]\times[u,v]\big)+\omega\big([u,t]^2\big)+\omega\big([s,u]\times[t,v]\big)\big)\\ &\le 3^{\rho-1}\omega\big([s,t]\times[u,v]\big).\end{aligned}$$

The other cases are dealt similarly, and we find at the end that for all $s\le t$, $u\le v$,

$$\left|f\begin{pmatrix} s & u \\ t & v\end{pmatrix}\right|^\rho \le 3^{\rho-1}\big(\omega[s,t]\times[u,v]\big).$$

That concludes the proof. □

Example 7. *Given two functions $g,h\in C^{\rho\text{-var}}([0,T],\mathcal{B})$ we can define*

$$(g\otimes h)(s,t):=g(s)\otimes h(t)\in\mathcal{B}\otimes\mathcal{B}$$

and $g \otimes h$ has finite 2D ρ-variation. More precisely,

$$\left|(g \otimes h)\begin{pmatrix} s & u \\ t & v \end{pmatrix}\right|^{\rho} \le |g|^{\rho}_{\rho\text{-var};[s,t]}|h|^{\rho}_{\rho\text{-var};[u,v]} =: \omega\big([s,t]\times[u,v]\big)$$

and since $\omega$is indeed a 2D control function (as product of two 1D control functions!) we see that

$$|g \otimes h|_{\rho\text{-var};[s,t]\times[u,v]} \le |g|_{\rho\text{-var};[s,t]}|h|_{\rho\text{-var};[u,v]}.$$

Remark 8. *If $\omega = \omega([s,t]\times[u,v])$ is a 2D control function, then*

$$(s,t) \mapsto \omega\big([s,t]^2\big)$$

is a 1D control function i.e. $\omega([s,t]^2) + \omega([t,u]^2) \le \omega([s,u]^2)$, and $s,t \to \omega([s,t]^2)$ is continuous and zero on the diagonal.

A function $f : [0,T]^2 \to (\mathcal{B}, |\cdot|)$ of finite q-variation can also be considered as path $t \mapsto f(t,\cdot)$ with values in the Banach space $C^{q\text{-var}}([0,T],\mathcal{B})$ with q-variation (semi-)norm. It is instructive to observe that $t \mapsto f(t,\cdot)$ has finite q-variation if and only if f has finite 2D q-variation. Let us now prove a (simplified) 2D version of a result of Musielak–Semandi [31] where they show that (in 1D) the family $C^{p\text{-var}}$ depends on p "semi-continuously from above."

Lemma 9. *For all $s < t$ and $u < v \in [0,1]$ we have $|R|_{\rho'\text{-var};[s,t]\times[u,v]} \to |R|_{\rho\text{-var};[s,t]\times[u,v]}$ as $\rho' \searrow \rho$.*

Proof. Write $Q = [s,t]\times[u,v]$ for a generic rectangle in $[0,1]^2$. Define $\omega(Q)^{1/\rho} = \liminf_{\rho'\searrow\rho} |R|_{\rho'\text{-var},Q}$. As $\rho' \mapsto |R|_{\rho'\text{-var},Q}$ is decreasing, this limit exists in $[0,\infty]$, and as $|R|_{\rho'\text{-var},[s,t]} \le |R|_{\rho\text{-var},[s,t]} < \infty$, it actually exists in $[0,\infty)$ and we have

$$\omega(Q) = \lim_{\rho'\searrow\rho} |R|^{\rho'}_{\rho'\text{-var},Q} \le |R|^{\rho}_{\rho\text{-var},Q}.$$

For all $s,t \in [0,1]$, for all ρ'

$$\big|R(Q)\big| \equiv \left|R\begin{pmatrix} s & u \\ t & v \end{pmatrix}\right| \le |R|_{\rho'\text{-var},Q}.$$

Taking the limit, we obtain $|R(Q)| \le \omega(Q)^{1/\rho}$. We now show that ω is super-additive. To this end, consider $\tilde{Q} = [r,s]\times[u,v]$, where $r < s < t$ in $[0,1]$. ($\tilde{Q}$, Q are essentially disjoint in the sense that $\tilde{Q}\cap Q$ has zero area.)Then

$$\begin{aligned}\omega(\tilde{Q}) + \omega(Q) &= \lim_{\rho'\to\rho} |R|^{\rho'}_{\rho'\text{-var},\tilde{Q}} + \lim_{\rho'\to\rho} |R|^{\rho'}_{\rho'\text{-var},Q} = \lim_{\rho'\to\rho}\big(|R|^{\rho'}_{\rho'\text{-var},\tilde{Q}} + |R|^{\rho'}_{\rho'\text{-var},Q}\big) \\ &\le \lim_{\rho'\to\rho} |R|^{\rho'}_{\rho'\text{-var},Q\cup\tilde{Q}} \le \omega(Q\cup\tilde{Q}).\end{aligned}$$

This argument shows that ω is super-addivity and so we can strengthen the estimate $|R(Q)| \le \omega(Q)^{1/\rho}$ to $|R|^{\rho}_{\rho\text{-var};Q} \le \omega(Q)$. But we also know $\omega(Q) \le |R|^{\rho}_{\rho\text{-var},Q}$ and hence have equality. □

2.2. *The integral*

Young integrals extend naturally to higher dimensions, see [36,37]. We focus on dimension 2 and $\mathcal{B} = \mathbb{R}$,which is what we need in the sequel.

Definition 10. *Let $f:[s,t]^2 \to \mathbb{R}$, $g:[u,v]^2 \to \mathbb{R}$ be continuous. Let $D=(t_i)$ be a dissection of $[s,t]$, $D'=(t'_j)$ be a dissection of $[u,v]$. If the 2D Riemann–Stieltjes sum*

$$\sum_{i,j} f(t_i,t'_j) g\begin{pmatrix} t_i & t'_j \\ t_{i+1} & t'_{j+1} \end{pmatrix}$$

converges when $\max\{\mathrm{mesh}(D),\mathrm{mesh}(D')\} \to 0$ we call the limit 2D Young-integral and write $\int_{[s,t]\times[u,v]} f\,\mathrm{d}g$ or simply $\int f\,\mathrm{d}g$ if no confusion arises.

We leave it to the reader to check that if g is of bounded variation (i.e. finite 1-variation) it induces a signed Radon measure, say λ_g, and

$$\int f\,\mathrm{d}g = \int f\,\mathrm{d}\lambda_g.$$

Example 11. *If $g(s,t)=\int_0^s \int_0^t r(x,y)\,\mathrm{d}x\,\mathrm{d}y$ (think of a 2D distribution function!) then*

$$\int_{[s,t]\times[u,v]} f\,\mathrm{d}g = \int_{[s,t]\times[u,v]} f(x,y) r(x,y)\,\mathrm{d}x\,\mathrm{d}y.$$

The following theorem was proved in [36], see also Young's original paper [37] for a (weaker) result in the same direction; we include a proof for the reader's convenience.

Theorem 12. *Let $f:[0,T]^2 \to \mathbb{R}$, $g:[0,T]^2 \to \mathbb{R}$ two continuous functions of finite q-variation (respectively of finite p-variation), with $\theta \equiv q^{-1}+p^{-1} > 1$, controlled by ω. Then the 2D Young-integral $\int_{[0,T]^2} f\,\mathrm{d}g$ exists and if $f(s,\cdot)=f(\cdot,u)=0$*

$$\left|\int_{[s,t]\times[u,v]} f\,\mathrm{d}g\right| \le C_{p,q} |f|_{q\text{-var},[s,t]\times[u,v]} \cdot |g|_{p\text{-var},[s,t]\times[u,v]}.$$

Proof. Let ω_f, ω_g be controls dominating the q-variation of f and p-variation of g, and let $\omega = \omega_f^{1/(\theta q)} \omega_g^{1/(\theta p)}$. Observe that by Hölder inequality, ω itself is a control. For a fixed $x, x' \in [s,t]$, define the functions $f_{x,x'}, g_{x,x'}$ by

$$f_{x,x'}(y) = f(x,y) - f(x',y), \quad y \in [u,v],$$
$$g_{x,x'}(y) = f(x,y) - f(x',y), \quad y \in [u,v].$$

Observe that $y \to f_{x,x'}(y)$ (resp. $y \to g_{x,x'}(y)$) is of finite q-variation (resp. p-variation) controlled by $(y,y') \to \omega_f([x,x']\times[y,y'])$ (resp. by $(y,y') \to \omega_g([x,x']\times[y,y'])$). That implies in particular by Young 1D estimates that

$$\left|\int_u^v f_{x_1,x_2}(y)\,\mathrm{d}g_{x_3,x_4}(y)\right| \le C_{q,p} \omega_f\big([x_1,x_2]\times[u,v]\big)^{1/q} \omega_g\big([x_3,x_4]\times[u,v]\big)^{1/p}.$$

For a subdivision $D=(x_i)$ of $[s,t]$, let $I_{u,v}^D = \sum_i \int_u^v f_{s,x_i}(y)\,\mathrm{d}g_{x_i,x_{i+1}}(y)$. Now, let $D\backslash\{i\}$ the subdivision D with the point x_i removed. It is easy to see that

$$\begin{aligned}
\big|I_{u,v}^D - I_{u,v}^{D\backslash\{i\}}\big| &= \left|\int_u^v f_{x_{i-1},x_i}(y)\,\mathrm{d}g_{x_i,x_{i+1}}(y)\right| \\
&\le C_{q,p} \omega_f\big([x_{i-1},x_i]\times[u,v]\big)^{1/q} \omega_g\big([x_i,x_{i+1}]\times[u,v]\big)^{1/p} \\
&\le C_{q,p} \omega_f\big([x_{i-1},x_{i+1}]\times[u,v]\big)^{1/q} \omega_g\big([x_{i-1},x_{i+1}]\times[u,v]\big)^{1/p} \\
&= C_{q,p} \omega\big([x_{i-1},x_{i+1}]\times[u,v]\big).
\end{aligned}$$

Choosing the point i such that $\omega([x_{i-1}, x_{i+1}] \times [u, v]) \le \frac{2}{r-1}\omega([s,t] \times [u,v])$, where r is the number of points in the subdivision D. Working from this point as in the proof of Young 1D estimate, we therefore obtain

$$\begin{aligned}\left|I_{u,v}^{D}\right| &\le C_{q,p}^{2}\omega\bigl([s,t]\times[u,v]\bigr)\\ &= C_{q,p}^{2}\omega_f\bigl([s,t]\times[u,v]\bigr)^{1/q}\omega_g\bigl([s,t]\times[u,v]\bigr)^{1/p}.\end{aligned}$$

We finish as in Young 1D proof, [37]. □

Remark 13. *One can take $C_{p,q}$ as a continuous function of $\theta = q^{-1} + p^{-1} \in (1,\infty)$. In the classical* 1*D case, this follows readily from the well-known expression of $C_{p,q} = 1 + \zeta(\theta)$, where ζ is Riemann's zeta function, continuous* (*and even analytic*) *for $\theta > 1$. In the* 2*D case, this follows from the constants given in* [36].

3. One dimensional Gaussian processes and the ρ-variation of their covariance

3.1. *Examples*

3.1.1. *Brownian Motion*

Standard Brownian Motion B on $[0,1]$ has covariance $R_{BM}(s,t) = \min(s,t)$. By Lemma 6, or directly from the definition, R has finite ρ-variation with $\rho = 1$, controlled by

$$\begin{aligned}\omega\bigl([s,t]\times[u,v]\bigr) &= \bigl|(s,t)\cap(u,v)\bigr|\\ &= \int_{[s,t]\times[u,v]}\delta_{x=y}(\mathrm{d}x\,\mathrm{d}y),\end{aligned}$$

where δ is the Dirac mass. Since $\omega([s,t]^2) = |t-s|$, it is Hölder dominated.

3.1.2. (*Gaussian*) *martingales*

We know that a continuous Gaussian martingale M has a deterministic bracket so that

$$M(t) \stackrel{D}{=} B_{\langle M\rangle_t}.$$

In particular,

$$R(s,t) = \min\bigl\{\langle M\rangle_s, \langle M\rangle_t\bigr\} = \langle M\rangle_{\min(s,t)}.$$

But the notion of ρ-variation is invariant under time-change and it follows that R has finite 1-variation since R_{BM} has finite 1-variation. One should notice that L^2-martingales (without assuming a Gaussian structure) have orthogonal increments i.e.

$$\mathbb{E}(X_{s,t}X_{u,v}) = 0 \quad \text{if } s < t < u < v$$

and this alone will take care of the (usually difficult to handle) off-diagonal part in the variation of the covariance $(s,t) \mapsto \mathbb{E}(X_s X_t)$.

3.1.3. *Bridges*, *Ornstein–Uhlenbeck process*

Gaussian Bridge processes are immediate generalizations of the Brownian Bridge. Given a real-valued centered Gaussian process X on $[0,1]$ with continuous covariance R of finite ρ-variation the corresponding Bridge is defined as

$$X_B(t) = X(t) - tX(1)$$

with covariance R_B. It is a simple exercise left to the reader to see that R_B has finite ρ-variation. Moreover, if R has its ρ-variation over $[s,t]^2$ dominated by a Hölder control, then R_B has also ρ-variation dominated by a Hölder control.

The usual Ornstein–Uhlenbeck (stationary or started at a fixed point) also has finite 1-variation, Hölder dominated on $[s,t]^2$. This is seen directly from the explicitly known covariance function and also left to the reader.

3.1.4. *Fractional Brownian Motion*

Finding the precise ρ-variation for the covariance of the fractional Brownian Motion is more involved. For Hurst parameter $H > 1/2$, fractional Brownian Motion has Hölder sample paths with exponent greater than $1/2$ which is, for the purpose of this paper, a trivial case.

Proposition 14. *Let B^H be fractional Brownian Motion of Hurst parameters $H \in (0, 1/2]$. Then, its covariance is of finite $1/(2H)$-variation. Moreover, its ρ-variation over $[s,t]^2$ is bounded by $C_H|t-s|$.*

Proof. (Using scaling properties of B^H the proof could be reduced to the case when $[s,t] = [0,1]$ but this does not simplify the analysis.) However, this does let $D = \{t_i\}$ be a dissection of $[s,t]$, and let us look at

$$\sum_{i,j} \left|\mathbb{E}\left(B^H_{t_i,t_{i+1}} B^H_{t_j,t_{j+1}}\right)\right|^{1/(2H)}.$$

For a fixed i and $i \neq j$, as $H \leq 1/2$, $\mathbb{E}(B^H_{t_i,t_{i+1}} B^H_{t_j,t_{j+1}})$ is negative, hence,

$$\begin{aligned}
\sum_{j} \left|\mathbb{E}\left(B^H_{t_i,t_{i+1}} B^H_{t_j,t_{j+1}}\right)\right|^{1/(2H)} &\leq \sum_{j\neq i} \left|\mathbb{E}\left(B^H_{t_i,t_{i+1}} B^H_{t_j,t_{j+1}}\right)\right|^{1/(2H)} + \mathbb{E}\left(\left|B^H_{t_i,t_{i+1}}\right|^2\right)^{1/(2H)} \\
&\leq \left|\mathbb{E}\left(\sum_{j\neq i} B^H_{t_i,t_{i+1}} B^H_{t_j,t_{j+1}}\right)\right|^{1/(2H)} + \mathbb{E}\left(\left|B^H_{t_i,t_{i+1}}\right|^2\right)^{1/(2H)} \\
&\leq \left(2^{1/(2H)-1}\left|\mathbb{E}\left(\sum_{j} B^H_{t_i,t_{i+1}} B^H_{t_j,t_{j+1}}\right)\right|^{1/(2H)} + 2^{1/(2H)-1}\mathbb{E}\left(\left|B^H_{t_i,t_{i+1}}\right|^2\right)^{1/(2H)}\right) \\
&\quad + \mathbb{E}\left(\left|B^H_{t_i,t_{i+1}}\right|^2\right)^{1/(2H)} \\
&\leq C_H\left|\mathbb{E}\left(B^H_{t_i,t_{i+1}} B^H_{s,t}\right)\right|^{1/(2H)} + C_H \mathbb{E}\left(\left|B^H_{t_i,t_{i+1}}\right|^2\right)^{1/(2H)}.
\end{aligned}$$

Hence,

$$\sum_{i,j} \left|\mathbb{E}\left(B^H_{t_i,t_{i+1}} B^H_{t_j,t_{j+1}}\right)\right|^{1/(2H)} \leq C_H \sum_i \mathbb{E}\left(\left|B^H_{t_i,t_{i+1}}\right|^2\right)^{1/(2H)} + C_H \sum_i \left|\mathbb{E}\left(B^H_{t_i,t_{i+1}} B^H_{s,t}\right)\right|^{1/(2H)}.$$

The first term is equal to $C_H|t-s|$, so we just need to prove that[4]

$$\sum_i \left|\mathbb{E}\left(B^H_{t_i,t_{i+1}} B^H_{s,t}\right)\right|^{1/(2H)} \leq C_H|t-s|. \tag{4}$$

To achieve this, it will be enough to prove that for $[u,v] \subset [s,t]$,

$$\left|\mathbb{E}\left(B^H_{u,v} B^H_{s,t}\right)\right| \leq C_H|v-u|^{2H}.$$

[4] $h(\cdot) = E(B^H_{\cdot} B^H_{s,t})$ defines a Cameron–Martin path and estimate (4) says that $|h|_{1/(2H)\text{-var};[s,t]} \leq C|t-s|^{2H}$. It is instructive to compare this with the section on Cameron–Martin spaces.

First recall that as $2H < 1$, if $0 < x < y$, then $(x+y)^{2H} - x^{2H} \le y^{2H}$. Hence, using this inequality and the triangle inequality,

$$\begin{aligned}
\left|\mathbb{E}\left(B^H_{u,v} B^H_{s,t}\right)\right| &= c_H \left|(t-v)^{2H} + (u-s)^{2H} - (v-s)^{2H} - (t-u)^{2H}\right| \\
&\le c_H\left((t-u)^{2H} - (t-v)^{2H}\right) + c_H\left((v-s)^{2H} - (u-s)^{2H}\right) \\
&\le 2c_H(v-u)^{2H}. \qquad \square
\end{aligned}$$

3.1.5. *Coutin–Qian condition on the covariance*

Coutin and Qian [10] constructed a rough paths over a class of Gaussian process. We prove here that when we look at the ρ-variation of their covariance, they are not very different than fractional Brownian Motion.[5]

Definition 15. *A real-valued Gaussian process X on $[0,1]$ satisfies the Coutin–Qian conditions if for some H*

$$\mathbb{E}\left(|X_{s,t}|^2\right) \le c_H |t-s|^{2H} \quad \text{for all } s < t, \tag{5}$$

$$\left|\mathbb{E}(X_{s,s+h} X_{t,t+h})\right| \le c_H |t-s|^{2H-2} h^2 \quad \text{for all } s, t, h \text{ with } h < t-s. \tag{6}$$

Lemma 16. *Let X be a Gaussian process on $[0,1]$ that satisfies the Coutin–Qian conditions for some $H > 0$, and let ω_H the control of the $\frac{1}{2H}$-variation of the covariance of the fractional Brownian Motion with Hurst parameter H. Then, for $s \le t$ and $u \le v$,*

$$\left|\mathbb{E}(X_{s,t} X_{u,v})\right| \le C_H \omega_H\left([s,t]\times[u,v]\right)^{2H}.$$

In particular, the covariance of X has finite $\frac{1}{2H}$-variation.

Proof. Working as in Lemma 6, at the price of a factor $3^{1/(2H-1)}$, we can restrict ourselves to the cases $s = u \le t = v$, and $s \le t \le u \le v$. The first case it given by assumption (5), so let us focus on the second one. Assume first we can write $t - s = nh$, $v - u = mh$, and that $u - t > h$. Then,

$$\mathbb{E}(X_{s,t} X_{u,v}) = \sum_{k=0}^{n-1} \sum_{l=0}^{m-1} \mathbb{E}(X_{s+kh,s+(k+1)h} X_{t+lh,t+(l+1)h}).$$

Using the triangle inequality and our assumption,

$$\begin{aligned}
\left|\mathbb{E}(X_{s,t} X_{u,v})\right| &= \sum_{k=0}^{n-1} \sum_{l=0}^{m-1} \left|\mathbb{E}(X_{s+kh,s+(k+1)h} X_{u+lh,u+(l+1)h})\right| \\
&\le C_H \sum_{k=0}^{n-1} \sum_{l=0}^{m-1} \left|(u+lh) - (s+kh)\right|^{2H-2} h^2 \\
&\le C_H \sum_{k=0}^{n-1} \sum_{l=0}^{m-1} \int_{u+(l-1)h}^{u+lh} \int_{s+kh}^{s+(k+1)h} |y-x|^{2H-2}\, \mathrm{d}x\, \mathrm{d}y \\
&\le C_H \int_{u-h}^{v-h} \int_s^t |y-x|^{2H-2}\, \mathrm{d}x\, \mathrm{d}y \\
&\le C_H \left|\mathbb{E}\left(B^H_{u-h,v-h} B^H_{s,t}\right)\right|.
\end{aligned}$$

[5] As remarked in more detail in the Introduction, a slight generalization of this condition appears in [10] and is applicable to certain non-Gaussian processes.

Letting h tends to 0, by continuity, we easily see that

$$\left|\mathbb{E}(X_{s,t}X_{u,v})\right| \le C_H\left|\mathbb{E}\left(B^H_{u,v}B^H_{s,t}\right)\right|,$$

which implies our statement for $s \le t \le u \le v$. That concludes the proof. □

3.2. *Cameron–Martin space*

We consider a real-valued centered Gaussian process X on $[0, 1]$ with continuous sample paths and covariance R. The associated Cameron–Martin space (as known as Reproducing Kernel Hilbert space) $\mathcal{H} \subset C([0, 1])$ consists of paths

$$t \mapsto h_t = \mathbb{E}(ZX_t),$$

where Z is a $\{\sigma(X_t), t \in [0, 1]\}$-measurable, Gaussian random variable. If $h' = \mathbb{E}(Z'X_\cdot)$ denotes another element in $\mathcal{H}$, the inner product on $\mathcal{H}$ is defined as

$$\langle h, h'\rangle_{\mathcal{H}} = \mathbb{E}\left(ZZ'\right).$$

Regularity of Cameron–Martin paths is not only a natural question in its own right but will prove crucial in our later sections on support theorem and large deviations.

Proposition 17. *If R is of finite ρ-variation, then $\mathcal{H} \subset C^{\rho\text{-var}}$. More, precisely, for all $h \in \mathcal{H}$*

$$|h|_{\rho\text{-var};[s,t]} \le \sqrt{\langle h, h\rangle_{\mathcal{H}}}\sqrt{R_{\rho\text{-var};[s,t]^2}}.$$

Proof. Let $h = \mathbb{E}(ZX_\cdot)$, and (t_j) a subdivision of $[s, t]$. We write $|x|_{l^r} = (\sum_i x_i^r)^{1/r}$ for $r \ge 1$. Let ρ' be the conjugate of ρ:

$$\begin{aligned}
\left(\sum_j |h_{t_j,t_{j+1}}|^\rho\right)^{1/\rho} &= \sup_{\beta,|\beta|_{l^{\rho'}}\le 1}\sum_j \beta_j h_{t_j,t_{j+1}} = \sup_{\beta,|\beta|_{l^{\rho'}}\le 1}\mathbb{E}\left(Z\sum_j \beta_j X_{t_j,t_{j+1}}\right)\\
&\le \sqrt{\mathbb{E}\left(Z^2\right)}\sup_{\beta,|\beta|_{l^{\rho'}}\le 1}\sqrt{\sum_{j,k}\beta_j\beta_k\mathbb{E}(X_{t_j,t_{j+1}}X_{t_k,t_{k+1}})}\\
&\le \sqrt{\langle h, h\rangle_{\mathcal{H}}}\sup_{\beta,|\beta|_{l^{\rho'}}\le 1}\sqrt{\left(\sum_{j,k}|\beta_j|^{\rho'}|\beta_k|^{\rho'}\right)^{1/(\rho')}\left(\sum_{j,k}\left|\mathbb{E}(X_{t_j,t_{j+1}}X_{t_k,t_{k+1}})\right|^\rho\right)^{1/\rho}}\\
&\le \sqrt{\langle h, h\rangle_{\mathcal{H}}}\left(\sum_{j,k}\left|\mathbb{E}(X_{t_j,t_{j+1}}X_{t_k,t_{k+1}})\right|^\rho\right)^{1/(2\rho)}\\
&\le \sqrt{\langle h, h\rangle_{\mathcal{H}}}\sqrt{R_{\rho\text{-var},[s,t]^2}}.
\end{aligned}$$

Optimizing over all subdivision (t_j) of $[s, t]$, we obtain our result. □

Remark 18. *Observe that for Brownian Motion ($\rho = 1$), this is a sharp result.*

3.3. *Piecewise-linear approximations*

Let X be centered real-valued continuous Gaussian process on $[0, 1]$ with covariance R assumed to be of finite ρ-variation, dominated by some 2D control function ω. Let $D = \{\tau_i\}$ be a dissection of $[0, 1]$ and let X^D denote the piecewise-linear approximation to X i.e. $X_t^D = X_t$ for $t \in D$ and X^D is linear between two successive points of D. If $(s,t)\times(u,v)\subset(\tau_i,\tau_{i+1})\times(\tau_j,\tau_{j+1})$ then the covariance of X^D, denoted by R^D, is given by

$$R^D\begin{pmatrix} s & u \\ t & v\end{pmatrix} = \mathbb{E}\left(\int_s^t \dot{X}_r^D\,\mathrm{d}r \int_u^v \dot{X}_r^D\,\mathrm{d}r\right) = \frac{t-s}{\tau_{i+1}-\tau_i}\times\frac{v-u}{\tau_{j+1}-\tau_j} R\begin{pmatrix} \tau_i & \tau_j \\ \tau_{i+1} & \tau_{j+1}\end{pmatrix}. \tag{7}$$

The aim of this section is to show that the ρ-variation of R^D is fully comparable with the ρ-variation of R. As usual, given $s\in[0,1]$, we write s_D the greatest element of D such that $s_D \le s$, and s^D the smallest element of D such that $s < s^D$.

Lemma 19.

(i) *For all* $u_1, v_1, u_2, v_2 \in D$,

$$\left|R^D\right|_{\rho\text{-var},[u_1,v_1]\times[u_2,v_2]} \le 9^{1-1/\rho}|R|_{\rho\text{-var},[u_1,v_1]\times[u_2,v_2]}.$$

(ii) *For all* $s,t\in[0,1]$, *with* $s_D \le s, t\le s^D$, *for all* $u,v\in D$,

$$\left|R^D\right|_{\rho\text{-var},[s,t]\times[u,v]} \le 9^{1-1/\rho}\left|\frac{t-s}{s^D-s_D}\right|\mathbb{E}\left(|X_{s_D,s^D}|^2\right)^{1/2}|R|^{1/2}_{\rho\text{-var},[u,v]^2}.$$

(iii) *For all* $s_1,t_1,s_2,t_2\in[0,1]$, *with* $s_{1,D}\le s_1, t_1\le s_1^D$, $s_{2,D}\le s_2, t_2 \le s_2^D$,

$$\left|R^D\right|_{\rho\text{-var},[s_1,t_1]\times[s_2,t_2]} \le \left|\frac{t_1-s_1}{s_1^D-s_{1,D}}\right|\left|\frac{t_2-s_2}{s_2^D-s_{2,D}}\right|\left|\mathbb{E}(X_{s_{1,D},s^{1,D}}X_{s_{2,D},s^{2,D}})\right|.$$

Proof. (i) Without loss of generality $[u_1,v_1]\times[u_2,v_2]=[0,1]^2$. Remark that R^D arises from the 2D function $R=R(s,t)$ simply by piecewise linear approximation of the partial functions $R(\cdot,\tau_i)$, $R(\tau_j,\cdot)$ for $\tau_i,\tau_j\in D$; in conjunction with (7) which specifies the rectangular increments of R^D, over small rectangles of form

$$[s,t]\times[u,v]\subset\underbrace{[\tau_i,\tau_{i+1}]}_{\equiv I}\times\underbrace{[\tau_j,\tau_{j+1}]}_{\equiv J},$$

directly in terms of rectangular increments of R. For the proof we introduce a 2D control function ω_D on $[0,1]^2$ as follows: for small rectangles $[s,t]\times[u,v]\subset I\times J$

$$\omega_D\left([s,t]\times[u,v]\right) := \frac{(t-s)(v-u)}{|I\times J|}|R|^\rho_{\rho\text{-var};I\times J}\quad\text{for } s,t\in I;\, u,v\in J;$$

then, for vertical "strips" of form $[s,t]\times(J_1\cup\cdots\cup J_n)$ with $s,t\in I\equiv[\tau_i,\tau_{i+1}]$ and $J_l=[\tau_{j+l-1},\tau_{j+l}]$

$$\omega_D\left([s,t]\times(J_1\cup\cdots\cup J_n)\right) := \frac{(t-s)}{|I|}|R|^\rho_{\rho\text{-var};I\times(J_1\cup\cdots\cup J_n)}\quad\text{for } s,t\in I;\, u,v\in J;$$

(similarly for horizontal strips); at last, for (possibly) large rectangle with endpoints in D we set

$$\omega_D\left((I_1\cup\cdots\cup I_m)\times(J_1\cup\cdots\cup J_n)\right) := |R|^\rho_{\rho\text{-var};(I_1\cup\cdots\cup I_m)\times(J_1\cup\cdots\cup J_n)}.$$

Now, an arbitrary rectangle $A=[a,b]\times[c,d]\subset[0,1]^2$ decomposes uniquely into (at most) 9 rectangles $A_1,\dots,A_9$ of the above type (4 small rectangles in the corners, 2 vertical and 2 horizontal strips and 1 rectangle with endpoints in D) and we define $\omega_D(A)=\sum_{i=1}^9\omega_D(A_i)$. On the other hand, it is clear from the definition of R^D that $|R^D(A_i)|^\rho \le$

$\omega_D(A_i)$ for $i = 1, \dots, 9$ and so[6]

$$\left|R^D(A)\right|^\rho = \left|\sum_{i=1}^{9} R^D(A_i)\right|^\rho \le 9^{\rho-1} \sum_{i=1}^{9} \left|R^D(A_i)\right|^\rho = 9^{\rho-1}\omega_D(A).$$

Since ω_D is (easily seen) to be a 2D control function the proof is finished with the remark that $\omega_D([0,1]) = |R|^\rho_{\rho\text{-var};[0,1]^2}$.

(ii) The second estimate is a bit more subtle. Take $s,t \in [0,1]$, with $s_D \le s,t \le s^D$, $u,v \in D$, (s_i) and (t_j) subdivisions of $[s,t]$ and $[u,v]$. Then, if $h_t^{i,D} = \mathbb{E}(X^D_{s_i,s_{i+1}} X^D_t)$, we know from Proposition 17 that

$$\begin{aligned}\left|h^{i,D}\right|_{\rho\text{-var},[u,v]} &\le \left|R^D\right|^{1/2}_{\rho\text{-var},[u,v]^2} \mathbb{E}\left(\left|X^D_{s_i,s_{i+1}}\right|^2\right)^{1/2} \\ &\le 9^{\rho-1}\frac{s_{i+1}-s_i}{s^D - s_D}|R|^{1/2}_{\rho\text{-var},[u,v]^2}\mathbb{E}\left(|X_{s_D,s^D}|^2\right)^{1/2}.\end{aligned}$$

Hence, for a fixed i,

$$\sum_j \left|\mathbb{E}\left(X^D_{s_i s_{i+1}} X^D_{t_j,t_{j+1}}\right)\right|^\rho \le \left|h^{i,D}\right|^\rho_{\rho\text{-var};[u,v]} \le \left(9^{\rho-1}\frac{s_{i+1}-s_i}{s^D-s_D}|R|^{1/2}_{\rho\text{-var}[u,v]^2}\mathbb{E}\left(|X_{s_D,s^D}|^2\right)^{1/2}\right)^\rho.$$

Summing over i and taking the supremum over all subdivision ends the proof of the second estimate. We leave the easy proof of the third estimate to the reader. □

Corollary 20. *Let X be continuous centered real-valued continuous Gaussian process on $[0,1]$ with covariance R assumed to be of finite ρ-variation. Then*:

(i) *for $s,t \in D$, the ρ-variation of R^D, the covariance of X^D, is bounded by $9^{\rho-1}|R|_{\rho\text{-var};[s,t]^2}$*,

(ii) *for all $s,t,u,v \in [0,1]$ and $\rho' > \rho$ the ρ'-variation of R^D over $[s,t]\times[u,v]$ converges to $|R|_{\rho'\text{-var};[s,t]\times[u,v]}$ when $|D| \to 0$*,

(iii) *if $|R|_{\rho\text{-var};[s,t]^2} \le C^R_{20}|t-s|^{1/\rho}$, then, $|R^D|_{\rho\text{-var};[s,t]^2} \le 9C^R_{20}|t-s|^{1/\rho}$.*

The same estimates apply to the covariance of (X, X^D).

Proof. The first statement is an easy corollary of the previous lemma. For the second we note that, by interpolation, $R^D \to R$ in ρ'-variation for any $\rho' > \rho$ so that

$$\left|R^D\right|_{\rho'\text{-var};[s,t]\times[u,v]} \to |R|_{\rho'\text{-var};[s,t]\times[u,v]} \quad \text{with } |D| \to 0.$$

(Note that we do not have $R^D \to R$ in ρ-variation in general but see remark below.) For the third one, without loss of generalities, we assume that $C^R_{20} = 1$. Then, by subadditivity of the ρ-variation at the power ρ,

$$\begin{aligned}\left|R^D\right|^\rho_{\rho\text{-var},[s,t]^2} \le{}& \left|R^D\right|^\rho_{\rho\text{-var},[s,s^D]^2} + \left|R^D\right|^\rho_{\rho\text{-var},[s,s^D]\times[s^D,t_D]} \\ &+ \left|R^D\right|^\rho_{\rho\text{-var},[s,s^D]\times[t_D,t]} + \left|R^D\right|^\rho_{\rho\text{-var},[s^D,t_D]\times[s,s^D]} \\ &+ \left|R^D\right|^\rho_{\rho\text{-var},[s^D,t_D]\times[s^D,t_D]} + \left|R^D\right|^\rho_{\rho\text{-var},[s^D,t_D]\times[t_D,t]} \\ &+ \left|R^D\right|^\rho_{\rho\text{-var},[t_D,t]\times[s,s^D]} + \left|R^D\right|^\rho_{\rho\text{-var},[t_D,t]\times[s^D,t_D]} \\ &+ \left|R^D\right|^\rho_{\rho\text{-var},[t_D,t]^2}.\end{aligned}$$

[6] $R^D(A)$ is the rectangular increment $R^D\binom{a\ c}{b\ d}$.

We bound each term using the previous lemma, and using on top estimates of the type

$$\begin{aligned}\left|\frac{t-s}{s^D-s_D}\right|\mathbb{E}\big(|X_{s_D,s^D}|^2\big)^{1/2} &\le \left|\frac{t-s}{s^D-s_D}\right|\left|s^D-s_D\right|^{1/(2\rho)}\\ &= \left|\frac{t-s}{s^D-s_D}\right|^{1-1/\rho}|t-s|^{1/(2\rho)}\\ &\le |t-s|^{1/\rho}.\end{aligned}$$

We leave the extension to the estimates on the covariation of (X, X^D) to the reader. □

4. Multidimensional Gaussian processes

As remarked in the Introduction, any $\mathbb{R}^d$-valued centered Gaussian process $X=(X^1,\dots,X^d)$ with continuous sample paths gives rise to an abstract Wiener space $(E,\mathcal{H},\mathbb{P})$ with $E=C([0,1],\mathbb{R}^d)$ and $\mathcal{H}\subset C([0,1],\mathbb{R}^d)$. If $\mathcal{H}^i$ denotes the Cameron–Martin space associated to the one dimensional Gaussian process X^i and all $\{X^i\colon i=1,\dots,d\}$ are independent then $\mathcal{H}\cong\bigoplus_{i=1}^d\mathcal{H}^i$.

4.1. *Wiener chaos*

Given an abstract Wiener space, there is a decomposition of $L^2(\mathbb{P})$ known as Wiener–Itô chaos decomposition, see [23,32,33] for the case of Wiener measure. Our interest in Wiener chaos comes from the following simple fact.

Proposition 21. *Assume the $\mathbb{R}^d$-valued continuous centered Gaussian process $X=(X^1,\dots,X^d)$ has sample paths of finite variation and let $S_N(X)\equiv\mathbf{X}$ denote its natural lift to a process with values in $G^N(\mathbb{R}^d)\subset T^N(\mathbb{R}^d)$. Then, for $n=1,\dots,N$ and any $s,t\in[0,1]$ the random variable $\pi_n(\mathbf{X}_{s,t})$ is an element in the nth (in general, not homogenous) Wiener chaos.*[7]

Proof. $\pi_n(\mathbf{X})$ is given by n iterated integrals which can be written out in terms of (a.s. convergent) Riemann–Stieltjes sums. Each such Riemann–Stieljes sum is a polynomial of degree at most n and of variables of form $X_{s,t}$. It now suffices to remark that the nth Wiener chaos contains all such polynomials and is closed under convergence in probability. □

As a consequence of the hypercontractivity property of the Ornstein–Uhlenbeck semigroup, L^p- and L^q-norms are equivalent on the nth Wiener chaos. Usually this is stated for the *homogenous* chaos, [23,32], but the extension to the nth (non-homogenous) chaos is not difficult, at least if we do not worry too much about optimal constants.

Lemma 22. *Let $n\in\mathbb{N}$ and Z be a random variable in the nth Wiener chaos. Assume $1<p<q<\infty$. Then*

$$|Z|_{L^p}\le|Z|_{L^q}\le|Z|_{L^p}(n+1)(q-1)^{n/2}\max\big(1,(p-1)^{-n}\big).$$

In particular, for $q>2$,

$$|Z|_{L^2}\le|Z|_{L^q}\le|Z|_{L^2}(n+1)(q-1)^{n/2}.$$

Proof. Only the second inequality requires proof. We need two well-known facts, both found in [32], for instance. First, if Z_k is a random variable in the kth homogeneous Wiener chaos then

$$|Z_k|_{L^q}\le\left(\frac{q-1}{p-1}\right)^{k/2}|Z_k|_{L^p}. \tag{8}$$

[7] Strictly speaking, the $(\mathbb{R}^d)^{\otimes n}$-valued chaos.

Secondly, the L^2-projection on the kth homogeneous chaos, denoted by J_k, is a bounded operator from $L^p \to L^p$ for any $1 < p < \infty$; more precisely[8]

$$|J_k Z|_{L^p} \le \begin{cases} (p-1)^{k/2}|Z|_{L^p} & \text{if } p \ge 2, \\ (p-1)^{-k/2}|Z|_{L^p} & \text{if } p < 2. \end{cases}$$

From $Z = \sum_{k=0}^n J_k Z$, we have $|Z|_{L^q} \le \sum_{k=0}^n |J_k Z|_{L^q}$ and hence

$$\begin{aligned} |Z|_{L^q} &\le \sum_{k=0}^n \left(\frac{q-1}{p-1}\right)^{k/2} |J_k Z|_{L^p} \\ &\le |Z|_{L^p} \sum_{k=0}^n \begin{cases} (q-1)^{k/2} & \text{if } p \ge 2, \\ (q-1)^{k/2}(p-1)^{-k} & \text{if } p < 2 \end{cases} \\ &\le |Z|_{L^p}(n+1)(q-1)^{n/2} \max\bigl(1, (p-1)^{-n}\bigr). \end{aligned}$$

□

Here is a immediate, yet useful, application. Assume Z, W are in the nth Wiener chaos. Then there exists $C = C(n)$

$$|WZ|_{L^2} \le C|W|_{L^2}|Z|_{L^2}. \tag{9}$$

(There is nothing special about L^2 here, but this is how we usually use it.) We now discuss more involved corollaries.

Corollary 23. *Let g be a random element of $G^N(\mathbb{R}^d)$ such that for all $1 \le n \le N$ the projection $\pi_n(g)$ is an element of the nth Wiener chaos. Let δ be a positive real. Then, the following statements* 1–6 *are equivalent*:

(i) *There exists a constant $C_1 > 0$ such that for all $n = 1, \dots, N$ there exists $q = q(n) \in (1, \infty)$: $|\pi_n(g)|_{L^q} \le C_1 \delta^n$*;
(ii) *There exists a constant $C_2 > 0$ such that for all $n = 1, \dots, N$ and for all $q \in [1, \infty)$: $|\pi_n(g)|_{L^q} \le C_2 q^{n/2} \delta^n$*;
(iii) *There exists a constant $C_3 > 0$ such that for all $n = 1, \dots, N$ there exists $q = q(n) \in (1, \infty)$: $|\pi_n(\log(g))|_{L^q} \le C_3 \delta^n$*;
(iv) *There exists a constant $C_4 > 0$ such that for all $n = 1, \dots, N$ and for all $q \in [1, \infty)$: $|\pi_n(\log(g))|_{L^q} \le C_4 q^{n/2} \delta^n$*;
(v) *There exists a constant $C_5 > 0$ and there exists $q \in (N, \infty)$: $\mathbb{E}(\|g\|^q)^{1/q} \le C_5 \delta$*;
(vi) *There exists a constant $C_6 > 0$ such that for all $q \in [1, \infty)$: $\mathbb{E}(\|g\|^q)^{1/q} \le C_6 q^{1/2} \delta$.*
When switching from ith to the jth statement, the constant C_j depends only on C_i, N and d.

Remark 24. *The restrictions on q in statements* 1, 3, 5 *comes from Lemma* 22 *where equivalence of L^p- and L^q-norms* (*on the nth Wiener chaos*) *is shown only for $p, q > 1$. In fact, this equivalence holds true for all $0 < p < q < \infty$* (*and hence statements* 1, 3, 5 *can be formulated with $q \in (0, \infty)$*). *This follows from the work of C. Borell* [3–5] *and is easy to see if one accepts a results of Schreiber* [34] *that convergence in probability and in L^q are equivalent on the nth Wiener chaos. Indeed, first note that for any $p > 0$, L^p-convergence implies convergence in probability and hence in L^q so that the identity map from $L^p \to L^q$ is continuous. Assume it is not bounded. Then there exists a sequence of random variables (Z^n) such that $|Z^n|_{L^q} > n|Z^n|_{L^p}$. But $W^n := Z^n/|Z^n|_{L^q}$ satisfies $1/n > |W^n|_{L^p}$ and hence converges to* 0 *in L^p which contradicts $|W^n|_{L^q} \equiv 1$.*

Proof. Clearly, (vi)$\Longrightarrow$(v), (iv)$\Longrightarrow$(iii), (ii)$\Longrightarrow$(i), and Lemma 22 shows (iii)$\Longrightarrow$(iv), and (i)$\Longrightarrow$(ii). It is therefore enough to prove (ii)$\Longrightarrow$(vi), (v)$\Longrightarrow$(i), and (ii)$\Leftrightarrow$(iv).

[8] In fact, this is a simple consequence of (8) when $p > 2$ and combined with a duality argument for $1 < p < 2$.

(ii)$\Longrightarrow$(vi) By equivalence of homogeneous norm, there exists a constant $C > 0$ such that,

$$\|g\| \le C \max_{n=1,\dots,N} \left|\pi_n(g)\right|^{1/n},$$

so that,

$$\begin{aligned}
\mathbb{E}\left(\|g\|^q\right)^{1/q} &\le C\mathbb{E}\Big(\max_{n=1,\dots,N} \left|\pi_n(g)\right|^{q/n}\Big)^{1/q} \le C\left(\sum_{n=1}^{N} \mathbb{E}\left(\left|\pi_n(g)\right|^{q/n}\right)\right)^{1/q} \\
&\le C\left(\sum_{n=1}^{N} \mathbb{E}\left(\left|\pi_n(g)\right|^{q/n}\right)\right)^{1/q} \le C\left(\sum_{n=1}^{N} C_2^{q/n}\left(q^{n/2}\delta^n\right)^{q/n}\right)^{1/q} \\
&\le C_6 q^{1/2}\delta.
\end{aligned}$$

(v)$\Longrightarrow$(i) By equivalence of homogeneous norm, there exists a constant $c > 0$ such that,

$$\left|\pi_n(g)\right|^{1/n} \le c\|g\|.$$

Hence,

$$\mathbb{E}\left(\left|\pi_n(g)\right|^{q_0/n}\right)^{n/q_0} \le c^n \mathbb{E}\left(\|g\|^{q_0}\right)^{n/q_0} \le |cC|^N \delta^n.$$

(ii)$\Leftrightarrow$(iv) An easy consequence of (9) and the formulas

$$\pi_n(g) = \sum_{\substack{k_1,\dots,k_l \\ \sum_i k_i = n}} a_{k_1,\dots,k_l} \bigotimes_i \pi_i(\ln g),$$

$$\pi_n(\ln g) = \sum_{\substack{k_1,\dots,k_l \\ \sum_i k_i = n}} b_{k_1,\dots,k_l} \bigotimes_i \pi_i(g),$$

where the real coefficients $a_{k_1,\dots,k_l}$ and $b_{k_1,\dots,k_l}$ can be explicitly computed from the power series definition of ln and exp. □

Proposition 25. *Let* $\mathbf{X}$ *be a continuous* $G^N(\mathbb{R}^d)$*-valued stochastic process. Assume that for all* $s < t$ *in* $[0, 1]$ *and* $n = 1, \dots, N$, *the projection* $\pi_n(\mathbf{X}_{s,t})$ *is an element in the* n*th Wiener chaos and that, for some constant* C *and* 1D *control function* ω,

$$\left|\pi_n(\ln \mathbf{X}_{s,t})\right|_{L^2} \le C\omega(s,t)^{n/(2\rho)}. \tag{10}$$

Then exists a constant $C' = C'(\rho, N)$ *such that for all* $q \in [1, \infty)$

$$\left|d(\mathbf{X}_s, \mathbf{X}_t)\right|_{L^q} \le C'\sqrt{q}\,\omega(s,t)^{1/(2\rho)}. \tag{11}$$

(i) *If* $p > 2\rho$ *then* $\|\mathbf{X}\|_{p\text{-var};[0,1]}$ *has a Gauss tail i.e. there exists* $\eta = \eta(p, \rho, N, K) > 0$, *with* $\omega([0, 1]) \le K$, *such that*

$$\mathbb{E}\left(\mathrm{e}^{\eta\|\mathbf{X}\|^2_{p\text{-var};[0,1]}}\right) < \infty.$$

In particular, $\mathbf{X}$ *has a.s. sample paths of finite* p*-variation.*

(ii) *If* $\omega(s,t) \le K|t-s|$, *then* $\|\mathbf{X}\|_{p\text{-var}}$ *above may be replaced by* $\|\mathbf{X}\|_{1/p\text{-Höl}}$ *and* $\mathbf{X}$ *has a.s. sample paths of* $1/p$-*Hölder regularity.*

Proof. Equation (11) is a clear consequence of the Corollary 67. The rest follows from the results in Appendix B. Indeed, after setting $M = C'\sqrt{q}$ and $r = 2\rho$ so that (11) reads

$$\left|d(\mathbf{X}_s, \mathbf{X}_t)\right|_{L^q} \le M\omega(s,t)^{1/r}$$

we can appeal to Corollary 66 to obtain

$$\left|\|\mathbf{X}\|_{p\text{-var};[0,1]}\right|_{L^q} \le c_1 M\omega(0,1)^{1/r} = c_2\sqrt{q}$$

(where c_1 depends on ρ, p and c_2 depends on ρ, p, N, K), valid for all q large enough, $q \ge q_0(\rho, p)$. At last, as is well known, $\mathrm{O}(\sqrt{q})$-growth of the qth moment implies a Gauss tail. More quantitatively, a Taylor expansion of $x \mapsto \mathrm{e}^{\eta x^2}$ shows that $\mathbb{E}(\mathrm{e}^{\eta\|\mathbf{X}\|^2_{p\text{-var};[0,1]}}) < \infty$ provided $\eta = \eta(c_2)$ small enough. □

The same argument, but using Corollary 68 in Appendix B leads to:

Proposition 26. *Let* $\mathbf{X}, \mathbf{Y}$ *be two continuous* $G^N(\mathbb{R}^d)$-*valued stochastic processes. Assume that for all* $s < t$ *in* $[0,1]$ *and* $n = 1, \dots, N$ *the projection* $\pi_n(\mathbf{X}_{s,t}^{-1} \otimes \mathbf{Y}_{s,t})$ *is an element in the nth Wiener chaos and that, for some* $C > 0$, $\varepsilon \in [0,1)$ *and* $1D$ *control function* ω,

$$\left|\pi_n(\ln \mathbf{X}_{s,t})\right|_{L^2}, \left|\pi_n(\ln \mathbf{Y}_{s,t})\right|_{L^2} \le C\omega(s,t)^{n/(2\rho)}, \tag{12}$$

$$\left|\pi_n\left(\ln\left(\mathbf{X}_{s,t}^{-1} \otimes \mathbf{Y}_{s,t}\right)\right)\right|_{L^2} \le C\varepsilon\omega(s,t)^{n/(2\rho)}. \tag{13}$$

Then for all $q \in [1,\infty)$ *there exists a constant* $C' = C'(\rho, N, C) > 0$ *such that*

$$\left\|d(\mathbf{X}_{s,t}, \mathbf{Y}_{s,t})\right\|_{L^q} \le C'\varepsilon^{1/N}\sqrt{q}\,\omega(s,t)^{1/(2\rho)}. \tag{14}$$

(i) *If* $p > 2\rho$ *and then there exist positive constants* $\theta = \theta(p, \rho, N)$ *and* $C'' = C''(p, \rho, N, C, K)$ *with* $\omega([0,1]^2) \le K$ *such that*

$$\left|d_{p\text{-var};[0,1]}(\mathbf{X}, \mathbf{Y})\right|_{L^q} \le C''\varepsilon^\theta\sqrt{q}.$$

(ii) *If* $\omega(s,t) \le K|t-s|$, *then* $d_{p\text{-var};[0,1]}(\mathbf{X}, \mathbf{Y})$ *above may be replaced by* $d_{1/p\text{-Höl}}(\mathbf{X}, \mathbf{Y})$.

4.2. *Uniform estimates for lifts of piecewise linear Gaussian processes*

We recall that all Gaussian processes under consideration are defined on $[0,1]$, centered and with continuous sample paths. The aim of this section is to construct the lift of $X = (X^1, \dots, X^d)$ for $X^1, \dots, X^d$ independent, provided that the covariance function for each X^i has finite ρ-variation for some $\rho \in [1,2)$.

The proof of the following lemma is left to the reader.

Lemma 27. *Let* $(X_1, \dots, X_d)$ *be a* d-*dimensional Gaussian process, with covariance* R *of finite* ρ-*variation controlled by* ω. *Then, for every fixed* $\alpha = (\alpha_1, \dots, \alpha_d) \in \mathbb{R}^d$, *the covariance of*

$$\alpha_1 X_1 + \cdots + \alpha_d X_d$$

has finite ρ-*variation controlled by* ω *times a constant depending on* α.

Proposition 28. *Let* (X, Y) *be a* 2-*dimensional centered Gaussian process with covariance* R *of finite* ρ-*variation controlled by* ω. *Then, for fixed* $s < t$ *in* $[0,1]$, *the function*

$$(u,v) \in [s,t]^2 \mapsto f(u,v) := \mathbb{E}(X_{s,u}Y_{s,u}X_{s,v}Y_{s,v})$$

satisfies $f(s,\cdot)=f(\cdot,s)=0$ *and has finite* ρ*-variation. More precisely, there exists a constant* $C=C(\rho)$ *such that*

$$|f|^{\rho}_{\rho\text{-var};[s,t]^2}\le C\omega\big([s,t]^2\big)^2.$$

Proof. We fix $u<u'$, $v<v'$, all in $[s,t]$. Using

$$X_{s,u'}Y_{s,u'}-X_{s,u}Y_{s,u}=X_{u,u'}Y_{s,u'}+X_{s,u}Y_{u,u'},$$

we bound $|\mathbb{E}((X_{s,u'}Y_{s,u'}-X_{s,u}Y_{s,u})(X_{s,v'}Y_{s,v'}-X_{s,v}Y_{s,v}))|$ by

$$\big|\mathbb{E}(X_{u,u'}Y_{s,u'}X_{v,v'}Y_{s,v'})\big|+\big|\mathbb{E}(X_{s,u}Y_{u,u'}X_{v,v'}Y_{s,v'})\big|+\big|\mathbb{E}(X_{u,u'}Y_{s,u'}X_{s,v}Y_{v,v'})\big|+\big|\mathbb{E}(X_{s,u}Y_{u,u'}X_{s,v}Y_{v,v'})\big|.$$

To bound the second expression for example, we use a well-known identity for the product of Gaussian random variables,

$$\begin{aligned}\mathbb{E}(X_{s,u}Y_{u,u'}X_{v,v'}Y_{s,v'})&=\mathbb{E}(X_{s,u}Y_{u,u'})\mathbb{E}(X_{v,v'}Y_{s,v'})\\&\quad+\mathbb{E}(X_{s,u}X_{v,v'})\mathbb{E}(Y_{u,u'}Y_{s,v'})\\&\quad+\mathbb{E}(X_{s,u}Y_{s,v'})\mathbb{E}(X_{v,v'}Y_{u,u'}),\end{aligned}$$

to obtain

$$\begin{aligned}\frac{1}{C_\rho}\big|\mathbb{E}(X_{s,u}Y_{u,u'}X_{v,v'}Y_{s,v'})\big|^{\rho}&\le\omega\big([s,u]\times[u,u']\big)\omega\big([v,v']\times[s,v']\big)\\&\quad+\omega\big([s,u]\times[v,v']\big)\omega\big([u,u']\times[s,v']\big)\\&\quad+\omega\big([s,u]\times[s,v']\big)\omega\big([u,u']\times[v,v']\big)\\&\le\omega\big([s,t]\times[u,u']\big)\omega\big([v,v']\times[s,t]\big)\\&\quad+\omega\big([s,t]\times[v,v']\big)\omega\big([u,u']\times[s,t]\big)\\&\quad+\omega\big([s,t]\times[s,t]\big)\omega\big([u,u']\times[v,v']\big).\end{aligned}$$

Working similarly with all terms, we obtain that this last expression controls the ρ-variation of $(u,v)\in[s,t]^2\to\mathbb{E}(X_{s,u}Y_{s,u}X_{s,v}Y_{s,v})$, and the bound on the ρ-variation on $[s,t]^2$. $\square$

Proposition 29. *Assume* $X=(X^1,\dots,X^d)$ *is a centered continuous Gaussian process with independent components with piecewise linear sample paths. Let* $\rho\in[1,2)$ *and assume that the covariance of* X *is of finite* ρ*-variation dominated by a* 2D *control* ω. *Let* $\mathbf{X}=S_3(X)$ *denote the natural lift of* X *to a* $G^3(\mathbb{R}^d)$*-valued process. There exists* $C=C(\rho)$ *such that for all* $s<t$ *in* $[0,1]$ *and indices* $i,j,k\in\{1,\dots,d\}$,

(i) $\mathbb{E}\big(\big|X^i_{s,t}\big|^2\big)\le\omega\big([s,t]^2\big)^{1/\rho}$ *for all* i;

(ii) $\mathbb{E}\big(\big|\mathbf{X}^{i,j}_{s,t}\big|^2\big)\le C\omega\big([s,t]^2\big)^{2/\rho}$ *for* i,j *distinct*;

(iii.1) $\mathbb{E}\big(\big|\mathbf{X}^{i,i,j}_{s,t}\big|^2\big)\le C\omega\big([s,t]^2\big)^{3/\rho}$ *for* i,j *distinct*;

(iii.2) $\mathbb{E}\big(\big|\mathbf{X}^{i,j,k}_{s,t}\big|^2\big)\le C\omega\big([s,t]^2\big)^{3/\rho}$ *for* i,j,k *distinct.*

Proof. (i) is obvious. For (ii) fix $i\ne j$ and $s<t,s'<t'$. Then, using independence of X^i and X^j,

$$\begin{aligned}\mathbb{E}\big(\mathbf{X}^{i,j}_{s,t}\mathbf{X}^{i,j}_{s',t'}\big)&=\mathbb{E}\bigg(\int_s^t\int_{s'}^{t'}X^i_{s,u}X^i_{s',v}\,\mathrm{d}X^j_u\,\mathrm{d}X^j_v\bigg)\\&=\int_s^t\int_{s'}^{t'}\mathbb{E}\big(X^i_{s,u}X^i_{s',v}\big)\,\mathrm{d}\mathbb{E}\big(X^j_uX^j_v\big)\end{aligned}$$

$$= \int_s^t \int_{s'}^t \left[R_i(u,v) - R_i(s,v) - R_i\left(u,s'\right) + R_i\left(s,s'\right)\right] \mathrm{d}R_j(u,v)$$

$$\le C\omega\left([s,t] \times \left[s',t'\right]\right)^{2/\rho} \quad \text{by Young 2D estimate.}$$

(ii) follows trivially from setting $s = s'$, $t = t'$ (the general result will be used in the level (iii) estimates, see step 2 below). We break up the level (iii) estimates in a few steps, assuming $i \neq j$ throughout.

Step 1. For fixed $s < t$, $s' < t'$, $t' < u'$ we claim that

$$\mathbb{E}\left(\mathbf{X}_{s,t}^{i,j} X_{s',t'}^{i} X_{t',u'}^{j}\right) \le C\omega\left([s,t] \times \left[s',t'\right]\right)^{1/\rho} \omega\left([s,t] \times \left[t',u'\right]\right)^{1/\rho}.$$

Indeed, with $\mathrm{d}\mathbb{E}(X_{t',u'}^{j} X_u^j) \equiv \mathbb{E}(X_{t',u'}^{j} \dot{X}_u^j)\,\mathrm{d}u$ we have

$$\mathbb{E}\left(\mathbf{X}_{s,t}^{i,j} X_{s',t'}^{i} X_{t',u'}^{j}\right) = \mathbb{E}\left(\int_s^t X_{s,u}^i X_{s',t'}^i X_{t',u'}^j \,\mathrm{d}X_u^j\right) = \int_{u=s}^t \mathbb{E}\left(X_{s,u}^i X_{s',t'}^i\right) \mathrm{d}\mathbb{E}\left(X_{t',u'}^j X_u^j\right).$$

Since the 1D ρ-variation of $u \mapsto \mathbb{E}(X_{s,u}^i X_{s',t'}^i)$ is controlled by $(u,v) \mapsto \omega([u,v] \times [s',t'])$, and similarly for $u \mapsto \mathbb{E}(X_{t',u'}^j X_u^j)$, the (classical 1D) Young estimate gives

$$\left|\int_{u=s}^t \mathbb{E}\left(X_{s,u}^i X_{s',t'}^i\right) \mathrm{d}\mathbb{E}\left(X_{t',u'}^j X_u^j\right)\right| \le C\omega\left([s,t] \times \left[s',t'\right]\right)^{1/\rho} \omega\left([s,t] \times \left[t',u'\right]\right)^{1/\rho}.$$

Step 2. For fixed $s < t$, we claim that the 2D map $(u,v) \in [s,t]^2 \mapsto \mathbb{E}(\mathbf{X}_{s,u}^{i,j} \mathbf{X}_{s,v}^{i,j})$ has finite ρ-variation controlled by

$$[u_1,u_2] \times [v_1,v_2] \mapsto C\omega\left([s,t]^2\right)\omega\left([u_1,u_2] \times [v_1,v_2]\right).$$

Then, using the level (ii) estimate and step 1, for $u_1 < u_2$, $v_1 < v_2$ all in $[s,t]$,

$$\begin{aligned}
\mathbb{E}\left(\left(\mathbf{X}_{s,u_2}^{i,j} - \mathbf{X}_{s,u_1}^{i,j}\right)\left(\mathbf{X}_{s,v_2}^{i,j} - \mathbf{X}_{s,v_1}^{i,j}\right)\right) &= \mathbb{E}\left(\left(\mathbf{X}_{u_1,u_2}^{i,j} + X_{s,u_1}^i X_{u_1,u_2}^j\right)\left(\mathbf{X}_{v_1,v_2}^{i,j} + X_{s,v_1}^i X_{v_1,v_2}^j\right)\right) \\
&= \mathbb{E}\left(\mathbf{X}_{u_1,u_2}^{i,j} \mathbf{X}_{v_1,v_2}^{i,j}\right) \\
&\quad + \mathbb{E}\left(\mathbf{X}_{u_1,u_2}^{i,j} X_{s,v_1}^i X_{v_1,v_2}^j\right) \\
&\quad + \mathbb{E}(X_{s,u_1}^i X_{u_1,u_2}^j \mathbf{X}_{v_1,v_2}^{i,j}) \\
&\quad + \mathbb{E}\left(X_{s,u_1}^i X_{s,v_1}^i\right)\mathbb{E}\left(X_{u_1,u_2}^j X_{v_1,v_2}^j\right) \\
&\le \omega\left([u_1,u_2] \times [v_1,v_2]\right)^{2/\rho} \\
&\quad + \omega\left([u_1,u_2] \times [s,v_1]\right)^{1/\rho} \omega\left([u_1,u_2] \times [v_1,v_2]\right)^{1/\rho} \\
&\quad + \omega\left([s,u_1] \times [v_1,v_2]\right)^{1/\rho} \omega\left([u_1,u_2] \times [v_1,v_2]\right)^{1/\rho} \\
&\quad + \omega\left([s,u_1] \times [s,v_1]\right)^{1/\rho} \omega\left([u_1,u_2] \times [v_1,v_2]\right)^{1/\rho} \\
&\le 4\left\{\omega\left([s,t]^2\right)\omega\left([u_1,u_2] \times [v_1,v_2]\right)\right\}^{1/\rho}.
\end{aligned}$$

(Here we used that ω can be taken symmetric.)

Step 3. We now prove the level (iii) estimates and start with (iii.2). For i, j, k distinct, we have

$$\mathbb{E}\left(\left|\int_s^t \mathbf{X}_{s,u}^{i,j} \,\mathrm{d}X_u^k\right|^2\right) = \int\int_{[s,t]^2} \mathbb{E}\left(\mathbf{X}_{s,u}^{i,j} \mathbf{X}_{s,v}^{i,j}\right) \mathrm{d}R_k(u,v).$$

By Young's 2D estimate, combined with ρ-variation regularity of the integrand established in step 2, we obtain

$$\mathbb{E}\left(\left|\int_s^t \mathbf{X}_{s,u}^{i,j}\,\mathrm{d}X_u^k\right|^2\right) \le C\omega\big([s,t]^2\big)^{3/\rho},$$

as desired. The estimate (iii.1) follows from

$$\mathbb{E}\left(\left|\int_s^t \big(X_{s,u}^i\big)^2\,\mathrm{d}X_u^k\right|^2\right) = \int\int_{[s,t]^2} \mathbb{E}\big(\big(X_{s,u}^i\big)^2\big(X_{s,v}^i\big)^2\big)\,\mathrm{d}R_k(u,v)$$

and Young's 2D estimate, combined with ρ-variation regularity of the integrand which follows as a special case of Proposition 28 (the full generality will be used in the next section). □

Corollary 30. *With* $\mathbf{X}, \rho, \omega$ *as in the last proposition, there exists* $C = C(\rho, d)$ *such that for all* $s < t$ *in* $[0,1]$ *and* $n = 1, 2, 3$,

$$\mathbb{E}\big(\big|\pi_n(\ln \mathbf{X}_{s,t})\big|^2\big) \le C\omega\big([s,t]^2\big)^{n/\rho}.$$

Proof. For $n = 1, 2$ this is an immediate consequence of (i), (ii) of the preceding proposition. From Appendix C, $\pi_3(\ln \mathbf{X}_{s,t})$ expands with respect to the basis elements $[e_i, [e_j, e_k]]$ with coefficients of only four possible types

$$\mathbf{X}_{s,t}^{i,j,k}, \qquad \mathbf{X}_{s,t}^{i,i,j}, \qquad \big|\mathbf{X}_{s,t}^i\big|^2\mathbf{X}_{s,t}^j, \qquad \mathbf{X}_{s,t}^i\mathbf{X}_{s,t}^{i,j} \quad (i, j, k \text{ distinct}).$$

The first two are directly handled with (iii.1) and (iii.2). For the last two we use the estimate (9) together with (i), (ii). □

Corollary 31. *With* $\mathbf{X}, \rho, \omega$ *as in the last proposition,*[9] *there exists* $C = C(\rho, d)$ *such that for all* $q \in [1, \infty)$

$$\big|d(\mathbf{X}_s, \mathbf{X}_t)\big|_{L^q(\mathbb{P})} \le C\sqrt{q}\,\omega\big([s,t]^2\big)^{1/(2\rho)},$$

where C *can be chosen continuous in* ρ. *If* $p > 2\rho$ *then there exists* $\eta = \eta(p, \rho, K) > 0$, *with* $\omega([0,1]^2) \le K$, *such that*

$$\mathbb{E}\big(\exp\big(\eta\|\mathbf{X}\|_{p\text{-var};[0,1]}^2\big)\big) < \infty. \tag{15}$$

If $\omega(s,t) \le K|t-s|$, *then* $\|\mathbf{X}\|_{p\text{-var}}$ *above may be replaced by* $\|\mathbf{X}\|_{1/p\text{-Höl}}$.

Proof. This is just an application of Proposition 25 with 1D control $(s,t) \mapsto \omega([s,t]^2)$. To see the continuity of C with respect to $\rho \in [1,2)$, it suffices to trace it back to the first two proposition of this section: the dependence of all constants with respect to ρ arises from (classical 1D or 2D) Young inequalities and, as pointed out in Remark 13, we can choose these constants continuous in ρ. □

4.3. *Continuity estimates for lifts of piecewise linear Gaussian processes*

Proposition 32. *Let* $(X, Y) = (X^1, Y^1, \ldots, X^d, Y^d)$ *be a centered continuous Gaussian process with piecewise linear sample paths such that* (X^i, Y^i) *is independent of* (X^j, Y^j) *when* $i \ne j$. *Let* $\rho \in [1,2)$ *and* ω *a 2D control that dominates the* ρ*-variation of the covariance of* (X, Y). *Assume* $\rho' \in (\rho, 2)$ *and* $\omega([0,1]^2) \le K$. *Then there exists* $C_{32} = C_{32}(\rho, \rho', K)$ *such that for all* $s < t$ *in* $[0,1]$ *and indices* $i, j, k \in \{1, \ldots, d\}$,

(i) $\mathbb{E}\big(\big|\mathbf{X}_{s,t}^i - \mathbf{Y}_{s,t}^i\big|^2\big) \le |R_{X-Y}|_\infty^{1-\rho/\rho'}\omega\big([s,t]^2\big)^{1/\rho'}$ *for all* i;

(ii) $\mathbb{E}\big(\big|\mathbf{X}_{s,t}^{i,j} - \mathbf{Y}_{s,t}^{i,j}\big|^2\big) \le C_{32}|R_{X-Y}|_\infty^{1-\rho/\rho'}\omega\big([s,t]^2\big)^{2/\rho'}$ *for* i, j *distinct*;

[9]The optimal choice for ω is the ρ-variation of R raised to power ρ.

(iii.1) $\mathbb{E}(|\mathbf{X}^{i,i,j}_{s,t} - \mathbf{Y}^{i,i,j}_{s,t}|^2) \le C_{32}|R_{X-Y}|_\infty^{1-\rho/\rho'}\omega([s,t]^2)^{3/\rho'}$ *for* i, j *distinct*;

(iii.2) $\mathbb{E}(|\mathbf{X}^{i,j,k}_{s,t} - \mathbf{Y}^{i,j,k}_{s,t}|^2) \le C_{32}|R_{X-Y}|_\infty^{1-\rho/\rho'}\omega([s,t]^2)^{3/\rho'}$ *for* i, j, k *distinct.*

Proof. We first remark that interpolation inequalities work for 2D variation just as for 1D variation, more specifically,

$$\begin{aligned}|R_{X-Y}|_{\rho'\text{-var};[s,t]^2} &\le |R_{X-Y}|_\infty^{1-\rho/\rho'}|R_{X-Y}|^{\rho/\rho'}_{\rho\text{-var};[s,t]^2}\\ &\le |R_{X-Y}|_\infty^{1-\rho/\rho'}\omega([s,t]^2)^{1/\rho'},\end{aligned}$$

and that the ρ'-variation of the covariance of (X,Y) is also controlled by ω. The level (i) estimate is then simply

$$\mathbb{E}(|\mathbf{X}^i_{s,t} - \mathbf{Y}^i_{s,t}|^2) \le |R_{X_i-Y_i}|_{\rho'\text{-var};[s,t]^2} \le |R_{X-Y}|_\infty^{1-\rho/\rho'}\omega([s,t]^2)^{1/\rho'}.$$

For the level (ii) estimate fix $i \neq j$. By the triangle inequality,

$$\begin{aligned}|\mathbf{X}^{i,j}_{s,t} - \mathbf{Y}^{i,j}_{s,t}|_{L^2} &\le \left|\mathbf{X}^{i,j}_{s,t} - \int_s^t X^i_{s,u}\,\mathrm{d}Y^j_u\right|_{L^2} + \left|\int_s^t X^i_{s,u}\,\mathrm{d}Y^j_u - \mathbf{Y}^{i,j}_{s,t}\right|_{L^2}\\ &\le \left|\int_s^t X^i_{s,u}\,\mathrm{d}(X^j_u - Y^j_u)\right|_{L^2} + \left|\int_s^t (X^i_{s,u} - Y^i_{s,u})\,\mathrm{d}Y^j_u\right|_{L^2}.\end{aligned}$$

Using independence of $\sigma(X^i, Y^i)$ and $\sigma(X^j, Y^j)$, the variances of the Riemann–Stieltjes integrals which appear in the line above, are expressed as 2D Young integrals involving the respective covariances. Using 2D Young estimates with $1/\rho' + 1/\rho' > 1$, we see that, with changing constants c,

$$\begin{aligned}|\mathbf{X}^{i,j}_{s,t} - \mathbf{Y}^{i,j}_{s,t}|^2_{L^2} &\le c|R_{X-Y}|_{\rho'\text{-var};[s,t]^2}\omega([s,t]^2)^{1/\rho'}\\ &\le c|R_{X-Y}|_\infty^{1-\rho/\rho'}\omega([s,t]^2)^{2/\rho'}.\end{aligned}$$

We now turn to level (iii) estimates and keep $i \neq j$ fixed throughout. We have

$$\begin{aligned}|\mathbf{X}^{i,i,j}_{s,t} - \mathbf{Y}^{i,i,j}_{s,t}|^2_{L^2} &\le 2\left|\int_s^t (X^i_{s,u})^2\,\mathrm{d}(X^j_u - Y^j_u)\right|^2_{L^2}\\ &\quad + 2\left|\int_s^t \{(X^i_{s,u})^2 - (Y^i_{s,u})^2\}\,\mathrm{d}Y^j_u\right|^2_{L^2}.\end{aligned}$$

The variance of $\int_s^t (X^i_{s,u})^2\,\mathrm{d}(X^j_u - Y^j_u)$ can be written as 2D Young integral and by Proposition 28 and 2D Young estimates we obtain the bound

$$\left|\int_s^t (X^i_{s,u})^2\,\mathrm{d}(X^j_u - Y^j_u)\right|^2_{L^2} \le c|R_{X-Y}|_\infty^{1-\rho/\rho'}\omega(s,t)^{3/\rho'}.$$

To deal with the other term, we first note that, from Proposition 28, the ρ-variation of

$$(u,v) \mapsto g(u,v) \equiv \mathbb{E}[\{(X^i_{s,u})^2 - (Y^i_{s,u})^2\}\{(X^i_{s,v})^2 - (Y^i_{s,v})^2\}]$$

over $[s,t]^2$ is controlled by a constant times $\omega([s,t]^2)^2$ while its supremum norm on $[s,t]^2$ is bounded by a constant times

$$|R_{X-Y}|_\infty\omega([s,t]^2)^{1/\rho}.$$

To see the latter, it suffices to write $g(u,v)$ as expectation of the product of the four factors $X^i_{s,u} \pm Y^i_{s,u}$, $X^i_{s,v} \pm Y^i_{s,v}$, bounded by the product of the respective L^4-norms which are (everything is Gaussian) equivalent to the respective L^2-norms. This leads to

$$\begin{aligned}
&\left| \int_s^t \left\{ \left(X^i_{s,u}\right)^2 - \left(Y^i_{s,u}\right)^2 \right\} \mathrm{d}Y^j_u \right|^2_{L^2} \\
&\quad = \int_{[s,t]^2} g(u,v)\, \mathrm{d}R_{Y^j}(u,v) \\
&\quad \le c|g|_{\rho'\text{-var};[s,t]} |R_{Y^j}|_{\rho'\text{-var};[s,t]} \\
&\quad \le c|g|_\infty^{1-\rho/\rho'} |g|^{\rho/\rho'}_{\rho\text{-var};[s,t]} \omega\left([s,t]^2\right)^{1/\rho'} \\
&\quad \le c\left(|R_{X-Y}|_\infty \omega\left([s,t]^2\right)^{1/\rho}\right)^{1-\rho/\rho'} \left(\omega\left([s,t]^2\right)^2\right)^{1/\rho'} \omega\left([s,t]^2\right)^{1/\rho'} \\
&\quad = c|R_{X-Y}|_\infty^{1-\rho/\rho'} \omega\left([s,t]^2\right)^{1/\rho+2/\rho'}
\end{aligned}$$

and it follows that

$$\left|\mathbf{X}^{i,i,j}_{s,t} - \mathbf{Y}^{i,i,j}_{s,t}\right|^2_{L^2} \le c|R_{X-Y}|_\infty^{1-\rho/\rho'} \omega\left([s,t]^2\right)^{1/\rho+2/\rho'}.$$

It remains to prove (iii.2) and we fix distinct indices i, j, k. To see that

$$\mathbb{E}\left(\left|\mathbf{X}^{i,j,k}_{s,t} - \mathbf{Y}^{i,j,k}_{s,t}\right|^2\right) \le c|R_{X-Y}|_\infty^{(\rho'-\rho)/\rho'} \omega\left([s,t]^2\right)^{3/\rho'}$$

we proceed as in the proof of (ii) and start by subtract/adding

$$\int_{[s,t]} \mathbf{X}^{i,j}_{s,\cdot}\, \mathrm{d}Y^k.$$

After using the triangle inequality we are left with two terms. The first is the variance of the 2D Young integral $\int \mathbf{X}^{i,j}_{s,\cdot}\, \mathrm{d}(X-Y)^k$ which is handled via Proposition 28 and 2D Young estimates, exactly as earlier. The second term is of form $\int (\mathbf{X}^{i,j}_{s,u} - \mathbf{Y}^{i,j}_{s,u})\, \mathrm{d}Y^k_u$ and is handled by the split-up,

$$\mathbf{X}^{i,j}_{s,u} - \mathbf{Y}^{i,j}_{s,u} = \int_s^u X^i_{s,\cdot}\, \mathrm{d}\left(X^j - Y^j\right) + \int_s^u \left(X^i_{s,\cdot} - Y^i_{s,\cdot}\right) \mathrm{d}Y^j.$$

We leave the remaining details to the reader. □

Corollary 33. *With $\mathbf{X}, \mathbf{Y}, \rho, \rho', \omega, K$ as in the last proposition there exists $C_{33} = C_{33}(\rho, \rho', K)$ and $\theta = \theta(\rho, \rho') > 0$ such that for all $s < t$ in* $[0,1]$ *and $n = 1, 2, 3$,*

$$\mathbb{E}\left(\left|\pi_n\left(\ln\left(\mathbf{X}^{-1}_{s,t} \otimes \mathbf{Y}_{s,t}\right)\right)\right|^2\right) \le C_{33}|R_{X-Y}|^\theta_\infty \omega\left([s,t]^2\right)^{n/\rho'}.$$

Proof. For $n = 1$ this is a trivial consequence of (i) of the preceding proposition. From Appendix C,

$$\left|\pi_2\left(\ln\left(\mathbf{X}^{-1}_{s,t} \otimes \mathbf{Y}_{s,t}\right)\right)\right| \le \left|\pi_2(\ln \mathbf{Y}_{s,t}) - \pi_2(\ln \mathbf{X}_{s,t})\right| + \frac{1}{2}|Y_{s,t} - X_{s,t}| \cdot |Y_{s,t}|$$

which is readily handled by (i) and (ii) of the preceding proposition, noting that thanks to Wiener–Itô chaos integrability we can split up the L^2-norm of products, cf. Eq. (9),

$$\mathbb{E}\left[|Y_{s,t} - X_{s,t}|^2 |Y_{s,t}|^2\right] \le C\mathbb{E}\left(|Y_{s,t} - X_{s,t}|^2\right)\mathbb{E}\left(|Y_{s,t}|^2\right).$$

From Proposition 71 (Appendix C)

$$\left|\pi_3\left(\ln\left(\mathbf{X}_{s,t}^{-1}\otimes\mathbf{Y}_{s,t}\right)\right)\right| \le \left|\pi_3(\ln\mathbf{Y}_{s,t})-\pi_3(\ln\mathbf{X}_{s,t})\right| + \frac{1}{2}\left|\pi_2(\ln\mathbf{Y}_{s,t})-\pi_2(\ln\mathbf{X}_{s,t})\right||Y_{s,t}| + \frac{1}{12}|Y_{s,t}-X_{s,t}|\left(|X_{s,t}|^2+|Y_{s,t}|^2+6\left|\pi_2(\ln\mathbf{X}_{s,t})\right|\right).$$

The terms which appear in the last two lines are handled by split up of L^2-norm as above, the term

$$\left|\pi_3(\ln\mathbf{Y}_{s,t})-\pi_3(\ln\mathbf{X}_{s,t})\right|$$

expands with respect to the basis elements $[e_i,[e_j,e_k]]$ with coefficients of only four possible types,

$$\mathbf{Y}_{s,t}^{i,j,k}-\mathbf{X}_{s,t}^{i,j,k},\qquad \mathbf{Y}_{s,t}^{i,i,j}-\mathbf{X}_{s,t}^{i,i,j},\qquad \left|Y_{s,t}^i\right|^2Y_{s,t}^j-\left|X_{s,t}^i\right|^2X_{s,t}^j,\qquad Y_{s,t}^i\mathbf{Y}_{s,t}^{i,j}-X_{s,t}^i\mathbf{X}_{s,t}^{i,j}$$

with i,j,k distinct. The first two difference terms are handled precisely with (iii.1) and (iii.2), the remaining terms are estimated by the split up of L^2-norms combined with the elementary estimates of type

$$\left|bb'-aa'\right| \le \left|b\left(b'-a'\right)+(b-a)a'\right| \le |b|\left|b'-a'\right|+\left|a'\right||b-a|$$

and the estimates (i), (ii). □

The above estimates and Proposition 25 lead to the following important corollary. (Note that ω can be taken as $|R_{(X,Y)}|^{\rho}_{\rho\text{-var};[\cdot,\cdot],[\cdot,\cdot]}$.)

Corollary 34. *Under the above hypothesis, $|R_{X-Y}|_\infty \le 1$, $p>2\rho$ and $\omega([0,1]^2)$ bounded by K, there exists positive constants $\theta=\theta(p,\rho)>0$ and $C_{34}=C_{34}(p,\rho,K)$ such that*

$$\left|d_{p\text{-var}}\left(S_3(X),S_3(Y)\right)\right|_{L^q} \le C_{34}|R_{X-Y}|_\infty^\theta\sqrt{q}.$$

If $\omega(s,t)\le K|t-s|$, then $d_{p\text{-var}}$ above may be replaced by $d_{1/p\text{-Höl}}$.

4.4. *Natural lift of a Gaussian process*

We are now able to prove the main theorems of this chapter.

Theorem 35 (Construction of lifted Gaussian processes). *Assume $X=(X^1,\dots,X^d)$ is a centered continuous Gaussian process with independent components. Let $\rho\in[1,2)$ and assume the covariance of X if of finite ρ-variation dominated by a 2D control ω.*

(i): *(Existence) There exists a continuous $G^3(\mathbb{R}^d)$-valued process $\mathbf{X}$, such that a.e. realization is in $C_0^{0,p\text{-var}}([0,1],G^3(\mathbb{R}^d))$ for $p>2\rho$, and hence a geometric p-rough path for $p\in(2\rho,4)$, and which lifts the Gaussian process X in the sense $\pi_1(\mathbf{X}_t)=X_t-X_0$. If ω is Hölder dominated a.e. realization is in $C_0^{0,1/p\text{-Hölder}}([0,1],G^3(\mathbb{R}^d))$. Finally, there exists $C_{35}=C_{35}(\rho)$ such that for all $s<t$ in $[0,1]$ and $q\in[1,\infty)$,*

$$\left|d(\mathbf{X}_s,\mathbf{X}_t)\right|_{L^q} \le C_{35}\sqrt{q}\,\omega\left([s,t]^2\right)^{1/(2\rho)}; \tag{16}$$

and the random variables $\pi_n(\mathbf{X}_{s,t}),\pi_n(\ln\mathbf{X}_{s,t})$, $n=1,2,3$, are in the nth (not necessarily homogenous) Wiener–Itô chaos.

(ii): *(Fernique-estimates) Let $p>2\rho$ and $\omega([0,1]^2)\le K$. Then there exists $\eta=\eta(p,\rho,K)>0$, such that*

$$\mathbb{E}\left(\exp\left(\eta\|\mathbf{X}\|^2_{p\text{-var},[0,1]}\right)\right)<\infty.$$

If $\omega([s,t]^2)\le K|t-s|$ for all $s<t$ in $[0,1]$, then we can replace $\|\mathbf{X}\|_{p\text{-var},[0,1]}$ by $\|\mathbf{X}\|_{1/p\text{-Höl};[0,1]}$ above.

(iii): (*Uniqueness*) *The lift* $\mathbf{X}$ *is unique in the sense that it is the* $d_{p\text{-var}}$*-limit in* $L^q(\mathbb{P})$, *for any* $q \in [1, \infty)$, *of any sequence* $S_3(X^D)$ *with* $|D| \to 0$. (*As usual*, X^D *denotes the piecewise linear approximation of* X *based on a dissection* D *of* $[0, 1]$.)

(iv): (*Consistency*) *If* X *has a.s. sample paths of finite* $[1, 2)$*-variation*, $\mathbf{X}$ *coincides with the canonical lift obtained by iterated Young-integration of* X. *If* $\tilde{\mathbf{X}} = (1, \pi_1(\mathbf{X}), \pi_2(\mathbf{X})) \in C_0^{0,p\text{-var}}([0, 1], G^2(\mathbb{R}^d))$ *a.s. for* $p < 3$ *then* $\tilde{\mathbf{X}}$ *is a geometric* p*-rough path and* $\mathbf{X}$ *coincides with the Young–Lyons lift of* $\tilde{\mathbf{X}}$.

Definition 36. *We call* $\mathbf{X}$ *natural lift* (*of the Gaussian process*) X. *A typical realizations of* $\mathbf{X}$ *is called a Gaussian rough path*.

Proof (Existence, Uniqueness). Let (D_n) be a sequence of dissections with mesh $|D_n| \to 0$. Clearly, $|R_{X^{D_n} - X^{D_m}}|_\infty \to 0$ and from Corollary 34 for every $p > 2\rho$,

$$\left|d_{p\text{-var}}\big(S_3\big(X^{D_n}\big), S_3\big(X^{D_m}\big)\big)\right|_{L^q} \to 0.$$

In particular, we see that $(S_3(X^{D_n}))$ is Cauchy in probability as sequence of $C_0^{0,p\text{-var}}$-valued random variables[10] and so there exists $\mathbf{X} \in C_0^{0,p\text{-var}}([0, 1], G^3(\mathbb{R}^d))$ so that $d_{p\text{-var}}(S_3(X^{D_n}), \mathbf{X}) \to 0$ in probability and from the uniform estimates from Corollary 31 also in L^q for all $q \in [1, \infty)$. If $(\tilde{D}_n)$ is another sequence of dissections with mesh tending to zero, the same construction yields a limit, say $\tilde{\mathbf{X}}$. But

$$\begin{aligned} d_{p\text{-var}}(\mathbf{X}, \tilde{\mathbf{X}}) \le\ & d_{p\text{-var}}\big(\mathbf{X}, S_3\big(X^{D_n}\big)\big) + d_{p\text{-var}}\big(S_3\big(X^{D_n}\big), S_3\big(X^{\tilde{D}_n}\big)\big) \\ & + d_{p\text{-var}}\big(S_3\big(X^{\tilde{D}_n}\big), \tilde{\mathbf{X}}\big) \end{aligned}$$

and the right-hand side converges to zero (in probability, say) as $n \to \infty$ which shows $\mathbf{X} = \tilde{\mathbf{X}}$ a.s. We now show the estimate (16). To this end, let ω^n denote the 2D control given by $|R_{X^{D_n}}|^{\rho'}_{\rho'\text{-var};[[\cdot,\cdot],[\cdot,\cdot]]}$ for $\rho' \in (\rho, 2)$. From Corollary 31

$$\left|d\big(S_3\big(X^{D_n}\big)_s, S_3\big(X^{D_n}\big)_t\big)\right|_{L^q} \le C\sqrt{q}\,\omega^n\big([s, t]^2\big)^{1/(2\rho')}$$

and after sending $n \to \infty$, followed by $\rho' \downarrow \rho$ using Lemma 9, we find

$$\left|d(\mathbf{X}_s, \mathbf{X}_t)\right|_{L^q} \le C\sqrt{q}\,\omega\big([s, t]^2\big)^{1/(2\rho)}$$

and (16) is proved. The statements on $\pi_n(\mathbf{X}_{s,t}), \pi_n(\ln \mathbf{X}_{s,t}) \in n$th Wiener–Itô chaos are immediate from Proposition 21 and closeness of the nth Wiener–Itô chaos under convergence in L^q. We then see that one can switch to equivalent estimates in terms of $\pi_n(\mathbf{X}_{s,t}), \pi_n(\ln \mathbf{X}_{s,t})$ thanks to Recall that Corollary 23), in particular for $n = 1, 2, 3$ and all $s < t$ in $[0, 1]$,

$$\left|\pi_n(\ln \mathbf{X}_{s,t})\right|_{L^2} \le c\,\omega\big([s, t]^2\big)^{n/(2\rho)}.$$

(Regularity, Fernique) An immediate consequence of Proposition 25 applied with 1D control $(s, t) \mapsto \omega([s, t]^2)$.

(Consistency) An immediate consequence of our construction and basic continuity statements of the Young resp. Young–Lyons lift, [26,27]. □

Theorem 37. *Let* $(X, Y) = (X^1, Y^1, \ldots, X^d, Y^d)$ *be a centered continuous Gaussian process such that* (X^i, Y^i) *is independent of* (X^j, Y^j) *when* $i \neq j$. *Let* $\rho \in [1, 2)$ *and assume the covariance of* (X, Y) *is of finite* ρ*-variation dominated by a* 2D *control* ω. *Then, for every* $p > 2\rho$, *there exist positive constants* $\theta = \theta(p, \rho)$ *and* $C_{35} = C(p, \rho, K)$,

[10] A Cauchy criterion for convergence in probability of r.v.s with values in a Polish space is an immediate generalization of the corresponding real-valued case.

with $\omega([0,1]^2) \le K$, such that for all $q \in [1,\infty)$,

$$\left|d_{p\text{-var}}(\mathbf{X},\mathbf{Y})\right|_{L^q} \le C_{35}\sqrt{q}|R_{X-Y}|_\infty^\theta.$$

If $\omega([s,t]^2) \le K|t-s|$ for all $s<t$ in $[0,1]$ we can replace $d_{p\text{-var}}$ by $d_{1/p\text{-Höl}}$ in the preceding line.

Proof. Pick ρ' such that $p > 2\rho' > 2\rho$ and, similarly to the last proof, pass to the limit in Proposition 33. Conclude with Proposition 33. □

Proposition 38 (Young–Wiener integral). *Assume X has covariance R with finite ρ-variation. Let $f \in C^{q\text{-var}}([0,1],\mathbb{R})$, with $q^{-1}+\rho^{-1} > 1$. If X^n is a sequence of Gaussian processes whose covariances are uniformly of finite p-variation and such that $|R_{X^n - X}|_\infty$ converges to 0, then in the supremum topology, $t \to \int_0^t f_u \,\mathrm{d}X_u^n$ converges in L^2. We define this limit to be the integral*

$$t \mapsto \int_0^t f_u \,\mathrm{d}X_u.$$

For all $s<t$ in $[0,1]$, we have the Young–Wiener isometry,

$$\mathbb{E}\left(\left|\int_s^t f_u \,\mathrm{d}X_u\right|^2\right) = \int_{[s,t]^2} f_u f_v \,\mathrm{d}R(u,v),$$

and if $f(s)=0$ we have the Young–Wiener estimate

$$\mathbb{E}\left(\left|\int_s^t f_u \,\mathrm{d}X_u\right|^2\right) \le C_{\rho,q}|f|^2_{q\text{-var};[s,t]}|R|_{\rho\text{-var};[s,t]}. \tag{17}$$

Proof. Proving (17) for X piecewise linear and applying the same methodology developed in this chapter is enough. But for X piecewise linear, it is obvious that

$$\mathbb{E}\left(\left|\int_s^t f_u \,\mathrm{d}X_u\right|^2\right) = \int_{[s,t]^2} f_u f_v \,\mathrm{d}R(u,v).$$

Now, the q-variation of $(u,v) \to f_u f_v$ is of course bounded by $|f|^2_{q\text{-var}}$, so applying Young 2D estimates, we are done. □

Remark 39. *When X is Brownian Motion, $\mathrm{d}R = \delta_{\{s=t\}}$ and we recover the usual Itô isometry.*

4.5. *Almost sure convergence*

Proposition 40. *Let $X = (X_1,\dots,X_d)$ be a centered continuous Gaussian process with independent components, and assume that the covariance of X is of finite ρ-variation dominated by a 2D control ω, for some $\rho < 2$. Then, if $D = (t_i)$ is a subdivision of $[0,1]$, and X^D be the piecewise linear approximation of X. Then, if $p > 2\rho$, there exist positive constants $\theta = \theta(\rho,p)$ and $C = C(\rho,p,K)$, with $\omega([0,1]^2) \le K$, such that for all $q \in [1,\infty)$,*

$$\left|d_{p\text{-var}}\big(\mathbf{X}, S_3\big(X^D\big)\big)\right|_{L^q} \le C\sqrt{q}\max_i \omega\big([t_i,t_{i+1}]^2\big)^\theta.$$

If $\omega([s,t]^2) \le K|t-s|$ for all $s<t$ in $[0,1]$ we have

$$\left|d_{1/p\text{-Höl}}\big(\mathbf{X}, S_3\big(X^D\big)\big)\right|_{L^q} \le C\sqrt{q}\max_i |t_{i+1}-t_i|^\theta.$$

Proof. A simple corollary of Theorems 35, 37 combined with $|R_{X-X^D}|_\infty \le \max_i \omega([t_i, t_{i+1}]^2)^{1/\rho}$. □

As a corollary, we obtain a.s. convergence of dyadic approximations in a Hölder situation. In view of Lemma 16 we have arrived at a substantial generalization of the results in [10].

Corollary 41. *Let $X, \omega, \rho < 2$, $p > 2\rho$ as above and assume $\omega([s,t]^2) \le K|t-s|$ for all $s < t$ in $[0,1]$. If X^{D_n} denote the dyadic piecewise linear approximation of X based on $D_n = \{\frac{k}{2^n}, 0 \le k \le 2^n\}$ then there exist positive constants $\theta = \theta(\rho, p)$ and $C = C(\rho, p, K)$ so that for all $q \in [1, \infty)$*

$$\left|d_{1/p\text{-Höl}}\left(\mathbf{X}, S_3\left(X^{D_n}\right)\right)\right|_{L^q} \le C\sqrt{q}2^{-n\theta/\rho}$$

and as n tends to infinity, $d_{1/p\text{-Höl}}(\mathbf{X}, S_3(X^{D_n})) \to 0$ a.s. and in L^q.

Proof. Only the a.s. convergence statement remains to be seen. But this is a standard Borell–Cantelli argument. □

5. Weak approximations

5.1. *Tightness*

Proposition 42. *Let (X_n) be a sequence of centered, d-dimensional, continuous Gaussian process with independent components, and assume that the covariances of X_n with finite $\rho \in [1,2)$-variation dominated by a 2D control ω, uniformly in n. Let $p > 2\rho$ and let $\mathbf{X}_n$ denote the natural lift of X_n with a.e. sample path in $C_0^{0,p\text{-var}}([0,1], G^3(\mathbb{R}^d))$. Then the family $((\mathbf{X}_n)_*\mathbb{P})$, i.e. the laws of $\mathbf{X}_n$ viewed as Borel measures on the Polish space $C_0^{0,p\text{-var}}([0,1], G^3(\mathbb{R}^d))$, are tight. If ω is Hölder dominated, then tightness holds in $C_0^{0,1/p\text{-Höl}}([0,1], G^3(\mathbb{R}^d))$.*

Proof. Let us fix $p' \in (2\rho, p)$. Define K_R to be the relatively compact set in $C_0^{0,p\text{-var}}([0,1], G^3(\mathbb{R}^d))$,

$$\left\{\mathbf{x}\text{: for all } s < t \text{ in } [0,1]\text{: } \|\mathbf{x}_{s,t}\|^{p'} \le R\left|\omega\left([0,t]^2\right) - \omega\left([0,s]^2\right)\right|\right\}.$$

From the results of Appendix B, there exists real random variables M_n such that (i) for some μ small enough, $\sup_n E(\exp(\mu M_n^2)) < \infty$, (ii) for all $n \ge 1$, for all $s, t \in [0,1]$,

$$\left\|\mathbf{X}_n(s,t)\right\|^{p'} \le M_n\left|\omega\left([0,t]^2\right) - \omega\left([0,s]^2\right)\right|.$$

Hence, there exists $c = c(\mu) > 0$ such that $\sup_n P(\mathbf{X}_n \in K_R) \le \exp(-cR^2)$ which shows tightness in $C_0^{0,p\text{-var}}([0,1], G^3(\mathbb{R}^d))$. Similarly, for Hölder dominated ω we obtain tightness in $C_0^{0,1/p\text{-Höl}}([0,1], G^3(\mathbb{R}^d))$ from the relative compactness of

$$\left\{\mathbf{x}\text{: for all } s < t \text{ in } [0,1]\text{: } \|\mathbf{x}_{s,t}\|^{p'} \le R|t-s|\right\}.$$ □

5.2. *Convergence*

Theorem 43. *Let $\rho \in [1,2)$. Let X_n, X_∞ be continuous Gaussian process with covariance R^n, R^∞ of finite $\rho \in [1,2)$-variation dominated uniformly in n by a 2D control ω, such that*

$$R^n \to R^\infty \quad \text{pointwise on } [0,1]^2.$$

Let $\mathbf{X}_n, \mathbf{X}_\infty$ denote the associated natural $G^3(\mathbb{R}^d)$-valued lifted processes. Then, for any $p > 2\rho$, the processes $\mathbf{X}_n$ converge in distribution to $\mathbf{X}_\infty$ with respect to p-variation topology. If ω is Hölder dominated, then convergence holds with respect to $1/p$-Hölder topology.

Proof. By Prohorov's theorem [2], tightness already implies existence of weak limits as measures on

$$C^{0,p\text{-var}}\big([0,1],G^3(\mathbb{R}^d)\big) \quad \text{resp.} \quad C^{0,1/p\text{-Höl}}\big([0,1],G^3\big(\mathbb{R}^d\big)\big)$$

and it will suffice to establish weak convergence on the space $E := C([0,1],G^3(\mathbb{R}^d))$ with d_∞-metric. By the Portmanteau theorem [2], it suffices to show that for every $f: E \to \mathbb{R}$, bounded and uniformly continuous,

$$\mathbb{E} f(\mathbf{X}_n) \to \mathbb{E} f(\mathbf{X}_\infty). \tag{18}$$

To see this, fix $\varepsilon > 0$, and $\delta = \delta(\varepsilon) > 0$ such that $d_\infty(\mathbf{x},\mathbf{y}) < \delta$ implies $|f(\mathbf{x}) - f(\mathbf{y})| < \varepsilon$. The estimates of Proposition 40) are more than enough to see that there exists a dissection D, with small enough mesh, such that

$$\sup_{0 \le n \le \infty} \mathbb{P}\big(d_\infty\big(\mathbf{X}_n, S_3\big(X_n^D\big)\big) \ge \delta\big) < \varepsilon.$$

Hence,

$$\begin{aligned}
\sup_{0\le n\le\infty} \big|\mathbb{E} f(\mathbf{X}_n) - \mathbb{E} f\big(S_3\big(X_n^D\big)\big)\big| &\le \sup_{0\le n\le\infty} \big|\mathbb{E}\big[\big|f(\mathbf{X}_n) - f\big(S_3\big(X_n^D\big)\big)\big|; d_\infty\big(\mathbf{X}_n, S_3\big(X_n^D\big)\big) \ge \delta\big]\big| \\
&\quad + \sup_{0\le n\le\infty} \big|\mathbb{E}\big[\big|f(\mathbf{X}_n) - f\big(S_3\big(X_n^D\big)\big)\big|; d_\infty\big(\mathbf{X}_n, S_3\big(X_n^D\big)\big) < \delta\big]\big| \\
&\le 2|f|_\infty \sup_{0\le n\le\infty} \mathbb{P}\big(d_\infty\big(\mathbf{X}_n, S_3\big(X_n^D\big)\big) \ge \delta\big) + \varepsilon \\
&\le \big(2|f|_\infty + 1\big)\varepsilon.
\end{aligned}$$

On the other hand, $R^n \to R$ pointwise gives convergence of the finite-dimensional distributions and hence weak convergence of $(X_n^D(t))_{t\in D}$ to $(X_\infty^D(t))_{t\in D}$. The map $(X_n^D(t))_{t\in D} \mapsto f(S_3(X_n^D))$ is easily seen to be continuous and so, for $n \ge n_0(\varepsilon)$ large enough,

$$\big|\mathbb{E} f\big(S_3\big(X_n^D\big)\big) - \mathbb{E} f\big(S_3\big(X_\infty^D\big)\big)\big| \le \varepsilon.$$

The proof is then finished with the triangle inequality,

$$\begin{aligned}
\big|\mathbb{E} f(\mathbf{X}_n) - \mathbb{E} f(\mathbf{X}_\infty)\big| &\le \big|\mathbb{E} f(\mathbf{X}_n) - \mathbb{E} f\big(S_3\big(X_n^D\big)\big)\big| \\
&\quad + \big|\mathbb{E} f\big(S_3\big(X_\infty^D\big)\big) - \mathbb{E} f\big(\mathbf{X}_\infty\big)\big| \\
&\quad + \big|\mathbb{E} f\big(S_3\big(X_n^D\big)\big) - \mathbb{E} f\big(S_3\big(X_\infty^D\big)\big)\big| \\
&\le \big(2|f|_\infty + 1\big)2\varepsilon + \varepsilon.
\end{aligned}$$

□

Example 44. *Set $R(s,t) = \min(s,t)$. The covariance of fractional Brownian Motion is given by*

$$R^H(s,t) = \frac{1}{2}\big(s^{2H} + t^{2H} - |t-s|^{2H}\big).$$

Take a sequence $H_n \uparrow 1/2$. It is easy to see that $R^{H_n} \to R$ pointwise and from our discussion of fractional Brownian Motion, for any $\rho > 1$,

$$\lim_{n\to\infty} \sup \big|R^{H_n}\big|_{\rho\text{-var}} < \infty.$$

It follows that RDE solutions driven by (multidimensional) fractional Brownian Motion with Hurst parameter H_n tend weakly to the usual Stratonovich solution. More elementary, for $H_n \downarrow 1/2$ we see that Young ODE solutions driven by B^{H_n} tend weakly to a Stratonovich solution.

6. Karhunen–Loève approximations

Any choice of an orthonormal basis in $\mathcal{H}$, say $(h^k\colon k \in \mathbb{N})$, yields a L^2-expansion of a Gaussian process X as (a.s. and L^2-convergent) sum of the form $X = \sum_{k\in\mathbb{N}} Z_k h^k$ where $Z_k := \tilde{h}^k := \xi(h^k)$ and $h \in \mathcal{H} \mapsto \tilde{h} \in L^2(\Omega)$ is the classical isometry between $\mathcal{H}$ and the Gaussian subspace in $L^2(\Omega)$; see [23,24], [12], Chapter 3.4. As a reminder, that we work with continuous Gaussian processes with the concrete index set [0, 1], just as for Brownian Motion, we shall refer to L^2-approximation as Karhunen–Loeve (type) approximations, in the same spirit as we prefer to call $\mathcal{H}$ Cameron–Martin space rather than Reproducing Kernel Hilbert Space.

As in previous sections, let $X = (X^i\colon i = 1, \dots, d)$ be a centered continuous Gaussian process, with independent components, each with covariance R of finite ρ-variation for some $\rho \in [1, 2)$ and dominated by some 2D control ω. Let $\mathbf{X}$ be the natural lift of X to a $G^3(\mathbb{R}^d)$-valued process. If $\mathcal{H}_i \subset C([0,1],\mathbb{R})$ denotes the Cameron–Martin space associated to X^i, the Cameron–Martin space to X is identified with $\bigoplus_{i=1}^d \mathcal{H}_i$ and if $(h_i^k)_{k\geq 1}$ is an orthonormal basis for $\mathcal{H}_i$ then $\{(h_i^k(\cdot))_{i=1,\dots,d}, k \geq 1\}$ is an orthonormal basis for $\bigoplus_{i=1}^d \mathcal{H}$. We can write $h^k = (h_1^k, \dots, h_d^k)$.

6.1. *One-dimensional estimates*

The covariance $R = R(s,t)$ of X is a diagonal matrix with d entries. For the purpose of variational regularity of the covariance of a Karhunen–Loeve approximations we may assume that X is in fact 1-dimensional. For any $A \subset \mathbb{N}$ we define

$$\mathcal{F}_A = \sigma(Z_k, k \in A), \qquad X_t^A = \mathbb{E}[X_t|\mathcal{F}_A].$$

If $\omega([a,b] \times [c,d])$ is a (2D) control function which controls the ρ-variation of $R(\cdot,\cdot) = \mathbb{E}[X_\cdot X_\cdot]$ over the indicated rectangle, i.e.

$$|R|^{\rho}_{\rho\text{-var};[a,b]\times[c,d]} \leq \omega\big([a,b] \times [c,d]\big),$$

then clearly, $\mathbb{E}(|X_{s,t}^A|^2) \leq \mathbb{E}(|X_{s,t}|^2) \leq \omega([s,t]^2)^{1/\rho}$. It is then clear from elementary Gaussian estimates that X^A can be taken with continuous sample paths. Moreover, X^A is a Gaussian process in its own right and we shall write R^A for its covariance function,

$$R^A(s,t) = \mathbb{E}\big[X_s^A X_t^A\big].$$

Lemma 45. *Assume that R is of finite ρ-variation, for some $\rho \geq 1$. Then if* $\min\{|A|, |A^c|\} < \infty$

$$\big|R^A\big|_{\rho\text{-var}} < \infty.$$

In particular, if $\rho < 2$, there exists a natural lift of X^A to a $G^3(\mathbb{R}^d)$-valued process denoted by $\mathbf{X}^A$.

Proof. Assume first $|A^c| < \infty$. Then $|h_k \otimes h_k|_{\rho\text{-var};[s,t]^2} \leq |(h_k)|^2_{\rho\text{-var};[s,t]} \leq |R|_{\rho\text{-var};[s,t]^2}$ thanks to Proposition 17 and $|h_k|_{\mathcal{H}} = 1$. It follows that

$$\begin{aligned}
\big|R^A\big|_{\rho\text{-var};[s,t]^2} &= \Big|R - \sum_{k\in A^c} h_k \otimes h_k\Big|_{\rho\text{-var};[s,t]^2} \\
&\leq |R|_{\rho\text{-var};[s,t]^2} + \sum_{k\in A^c} |h_k \otimes h_k|_{\rho\text{-var};[s,t]^2} \\
&\leq \big(1 + \big|A^c\big|\big)|R|_{\rho\text{-var};[s,t]^2}.
\end{aligned}$$

If A is finite, the proof is similar but even easier. $\square$

The interest is in the above lemma is for $\rho \in [1,2)$. To obtain *uniform* estimates valid for all $A \subset \mathbb{N}$, we unfortunately[11] have to work in 2-variation.

Lemma 46. *Assume that R is of finite 2-variation. Then, the R^A has finite 2-variation, uniformly over all $A \subset \mathbb{N}$. More precisely,*

$$\sup_{A\subset\mathbb{N}} \left|R^A\right|_{2\text{-var};[s,t]^2} \le |R|_{2\text{-var};[s,t]^2}.$$

Proof. Let $D = (t_i)$ a subdivision of $[s,t]$ and set $X_i^A = X_{t_i,t_{i+1}}^A$. Let β be a positive semi-definite symmetric matrix, and let us estimate $|\sum_{i,j}\beta_{i,j}\mathbb{E}(X_i^A X_j^A)|$. Now

$$\mathbb{E}\big(X_i^A X_j^A\big) = \sum_{k\in A}\mathbb{E}(Z_k X_i)\mathbb{E}(Z_k X_j) = \frac{1}{2}\sum_{k\in A}\mathbb{E}\big(\big(Z_k^2 - \mathbb{E}\big(Z_k^2\big)\big)X_i X_j\big),$$

so that

$$\sum_{i,j}\beta_{i,j}\mathbb{E}\big(X_i^A X_j^A\big) = \frac{1}{2}\sum_{k\in A}\mathbb{E}\bigg(\big(Z_k^2 - \mathbb{E}\big(Z_k^2\big)\big)\sum_{i,j}\beta_{i,j}X_i X_j\bigg).$$

As β is symmetric, we can write $\beta = P^{\mathrm{T}}\operatorname{diag}(d_1,\dots,d_{\#D})P$, with PP^{T} the identity matrix and (non-negative) eigenvalues (d_i). By simple linear algebra,

$$\sum_{i,j}\beta_{i,j}X_i X_j = (PX)^{\mathrm{T}}\operatorname{diag}(\cdots)(PX) = \sum_i d_i (PX)_i^2$$

and so

$$\begin{aligned}
\sum_{i,j}\beta_{i,j}\mathbb{E}\big(X_i^A X_j^A\big) &= \sum_{k\in A}\sum_i d_i \frac{1}{2}\mathbb{E}\big(\big(Z_k^2 - \mathbb{E}\big(Z_k^2\big)\big)(PX)_i^2\big) \\
&= \sum_i d_i \sum_{k\in A}\mathbb{E}\big(Z_k (PX)_i\big)^2 \\
&\le \sum_i d_i \mathbb{E}\big((PX)_i^2\big) \quad \text{(Parseval inequality)} \\
&= \mathbb{E}\big((PX)^{\mathrm{T}} D (PX)\big) \\
&= \sum_{i,j}\beta_{i,j}\mathbb{E}(X_i X_j) \\
&\le |\beta|_{l^2}|R|_{2\text{-var}} \quad \text{(Hölder inequality).}
\end{aligned}$$

[11]The proof of the following lemma can be extended to showing that if R is of finite ρ-variation, where ρ is an integer greater than 2, then for all $A \subset \mathbb{N}$ and $s < t$, $|R^A|_{\rho\text{-var},[s,t]^2} \le |R|_{\rho\text{-var},[s,t]^2}$. This is done by choosing $\beta_{i,j} = \mathbb{E}(X_i^{n,m}X_j^{n,m})^{\rho-1}$ and indeed if $\rho - 1 \in \mathbb{N}$ then β is a positive symmetric matrix (this is a simple consequence of Hadamard–Schur's lemma). We could not prove (or disprove) this for general $\rho \ge 1$; with β being defined as fractional Hadamard power,

$$\mathbb{E}\big(X_i^{n,m}X_j^{n,m}\big)^{\rho-1}\operatorname{sign}\big[\mathbb{E}\big(X_i^{n,m}X_j^{n,m}\big)\big].$$

If Lemma 46 were to hold true for $\rho \in [1,2)$, the rough path convergence of Karhunen–Loève approximations would follow directly from Theorem 37.

Applying this estimate to $\beta_{i,j} = \mathbb{E}(X_i^A X_j^A)$ we find

$$\sqrt{\sum_{i,j} \left|\mathbb{E}\left(X_i^A X_j^A\right)\right|^2} \leq |R|_{2\text{-var}}.$$

The proof is finished by taking the supremum over all dissections of $[s,t]$. $\square$

6.2. *Uniform bounds on the modulus and convergence*

We now assume that R has finite ρ-variation for some $\rho \in [1,2)$ dominated by some 2D control ω, and we fix $A \subset \mathbb{N}$, finite or with finite complement, so that X^A admits a natural $G^3(\mathbb{R}^d)$-valued lift, denoted $\mathbf{X}^A$. Of course, $\mathbf{X}^{\mathbb{N}} = \mathbf{X}$.

Lemma 47 (Martingale). *For all $s < t$ in $[0,1]$, the following equality holds in $g_3(\mathbb{R}^d) \equiv \ln G^3(\mathbb{R}^d)$,*

$$\begin{aligned}\mathbb{E}\left(\ln(\mathbf{X}_{s,t})|\mathcal{F}_A\right) = \ln\left(\mathbf{X}_{s,t}^A\right) &+ \frac{1}{12}\sum_{i\neq j} X_{s,t}^{A;j} R_{X^{A^c;i}}\begin{pmatrix} s & s \\ t & t\end{pmatrix}[e_i,[e_i,e_j]] \\ &- \frac{1}{2}\sum_{i\neq j}\int_s^t R_{X^{A^c;i}}\begin{pmatrix} u & s \\ t & u\end{pmatrix} \mathrm{d}X_u^{A,j}\left[e_i,[e_i,e_j]\right].\end{aligned}$$

(*The integral which appears in the last line is a Young–Wiener integral in the sense of Proposition* 38.)

Remark 48. *Projection to $g_2(\mathbb{R}^d)$ yields to pleasant equality $\mathbb{E}(\ln(\mathbf{X}_{s,t})|\mathcal{F}_A) = \ln(\mathbf{X}_{s,t}^A)$which explains why martingale arguments* [11,14,16,20] *are enough to discuss the step* 2 *case. That said, the above lemma shows clearly that additional estimates are needed to handle the step* 3 *case.*

Proof. Our proposition at level 1 is $\mathbb{E}(\pi_1(\ln \mathbf{X}_{s,t})|\mathcal{F}_A) = \pi_1(\ln \mathbf{X}_{s,t}^A)$, which is (almost) the definition of X^A. The estimate at level 2 is implies by $\mathbb{E}[\mathbf{X}_{s,t}^{i,j}|\mathcal{F}_A] = (\mathbf{X}^A)_{s,t}^{i,j}$. This is fairly straightforward to prove: one just need to note that conditioning equal L^2-projection is (trivially) L^2-continuous; also recalling that both $\mathbf{X}$ and $\mathbf{X}^A$ are L^2-limit of lifted piecewise linear approximations. Level 3 statements are more involved. We can see as above that for distinct indices i,j,k,

$$\mathbb{E}\left[\mathbf{X}_{s,t}^{i,j,k}|\mathcal{F}_A\right] = \left(\mathbf{X}^A\right)_{s,t}^{i,j,k}.$$

From Proposition 70 in Appendix C, we see that $\mathbb{E}(\ln(\mathbf{X}_{s,t})|\mathcal{F}_A) - \ln(\mathbf{X}_{s,t}^A)$ is equal to

$$\begin{aligned}&\sum_{i\neq j}\mathbb{E}\left(\left\{\mathbf{X}_{s,t}^{i,i,j} + \frac{1}{12}\left|X_{s,t}^i\right|^2 X_{s,t}^j - \frac{1}{2}X_{s,t}^i\mathbf{X}_{s,t}^{i,j}\right\}\,\Big|\,\mathcal{F}_A\right)[e_i,[e_i,e_j]] \\ &\quad - \sum_{i\neq j}\left(\left(\mathbf{X}_{s,t}^A\right)^{i,i,j} + \frac{1}{12}\left|\left(X^A\right)_{s,t}^i\right|^2\left(X^A\right)_{s,t}^j - \frac{1}{2}\left(X^A\right)_{s,t}^i\left(\mathbf{X}^A\right)_{s,t}^{i,j}\right)\left[e_i,[e_i,e_j]\right].\end{aligned}$$

All the three terms can be written as sums (or L^2-limits thereof) involving terms of form $X_{r,s}^i X_{t,u}^i X_{v,w}^j$ and since (write $X_{r,s}^i = X_{r,s}^{A;i} + X_{r,s}^{A^c;i}$ and similarly for the other terms)

$$\mathbb{E}\left(X_{r,s}^i X_{t,u}^i X_{v,w}^j|\mathcal{F}_A\right) - \left(X^A\right)_{r,s}^i\left(X^A\right)_{t,u}^i\left(X^A\right)_{v,w}^j = X_{v,w}^{A;j}\mathbb{E}\left(X_{r,s}^{A^c;i}X_{t,u}^{A^c}{}^{i}\right).$$

After integration, we therefore obtain

$$\mathbb{E}\big(\mathbf{X}_{s,t}^{i,i,j}|\mathcal{F}_A\big)-\big(\mathbf{X}_{s,t}^{A}\big)^{i,i,j}=\frac{1}{2}\int_s^t \mathbb{E}\big(\big|X_{s,u}^{A^c;i}\big|^2\big)\,\mathrm{d}X_u^{A;j}=\frac{1}{2}\int_s^t R_{X^{A^c;i}}\begin{pmatrix} s & s\\ u & u\end{pmatrix}\mathrm{d}X_u^{A,j},$$

$$\mathbb{E}\big(\big|X_{s,t}^{i}\big|^2X_{s,t}^{j}|\mathcal{F}_A\big)-\big(X_{s,t}^{A;i}\big)^2X_{s,t}^{A,j}=X_{s,t}^{A;j}\mathbb{E}\big(\big|X_{s,t}^{A^c,i}\big|^2\big)=X_{s,t}^{A;j}R_{X^{A^c;i}}\begin{pmatrix} s & s\\ t & t\end{pmatrix}$$

and

$$\mathbb{E}\big(X_{s,t}^{i}\mathbf{X}_{s,t}^{i,j}|\mathcal{F}_A\big)-X_{s,t}^{A;i}\mathbf{X}_{s,t}^{A;i,j}=\int_s^t \mathbb{E}\big(X_{s,t}^{A^c;i}X_{s,u}^{A^c;i}\big)\,\mathrm{d}X_u^{A;j}=\int_s^t R_{X^{A^c;i}}\begin{pmatrix} s & s\\ t & u\end{pmatrix}\mathrm{d}X_u^{A,j}.$$

This finishes the proof. □

Proposition 49. *There exists a constant C such that for all $s<t$ in $[0,1]$, $A\subset\mathbb{N}$ and $i,j\in\{1,\dots,d\}$ distinct*

$$\left|\int_s^t R_{X^{A^c;i}}\begin{pmatrix} u & s\\ t & u\end{pmatrix}\mathrm{d}X_u^{A,j}\right|_{L^2}^2\le C\omega\big([s,t]^2\big)^{3/\rho}.$$

Proof. From

$$\int_s^t R_{X^{A^c;i}}\begin{pmatrix} u & s\\ t & u\end{pmatrix}\mathrm{d}X_u^{A,j}=\mathbb{E}\left(\int_s^t R_{X^{A^c;i}}\begin{pmatrix} u & s\\ t & u\end{pmatrix}\mathrm{d}X_u^{j}\,\Big|\,\mathcal{F}_A\right)$$

it suffices to consider the integral with integrator dX^j. We define

$$f(u):=R_{X^{A^c;i}}\begin{pmatrix} u & s\\ t & u\end{pmatrix}$$

and note that $f(s)=0$. It is easy to see that for $u<v$ in $[s,t]$,

$$|f_{u,v}|^2\le|R_{X^{A^c;i}}|^2_{2\text{-var};[u,v]\times[s,t]}+|R_{X^{A^c;i}}|^2_{2\text{-var};[s,t]\times[u,v]}.$$

Noting super-additivity of the right-hand side in $[u,v]$ and using Lemma 46,

$$|f|^2_{2\text{-var};[s,t]}\le 2|R_{X^{A^c;i}}|^2_{2\text{-var};[s,t]^2}\le 2|R_{X^i}|^2_{2\text{-var};[s,t]^2}\le 2\omega\big([s,t]^2\big)^{2/\rho}.$$

Now, f has finite 2-variation and the covariance of the integrator $\mathrm{d}X^j$ has finite ρ-variation, $\rho\in[1,2)$ controlled by ω. Thanks to $1/2+1/\rho>1$ we can conclude with the "Young–Wiener" estimate of Proposition 38. □

Putting the last two results together and using Proposition 25, we obtain the following theorem.

Theorem 50. *For all $s<t$ in $[0,1]$ there exists $C=C(\rho)$ such that*

$$\sup_{A\subset\mathbb{N},\min\{|A|,|A^C|\}<\infty}\mathbb{E}\big(\big\|\mathbf{X}_{s,t}^{A}\big\|^2\big)\le C\omega\big([s,t]^2\big)^{1/\rho}.$$

For $p>2\rho$ and $\omega([0,1]^2)\le K$ there exists $\eta=\eta(p,\rho,K)>0$ such that

$$\sup_{A\subset\mathbb{N},\min\{|A|,|A^C|\}<\infty}\mathbb{E}\big(\exp\eta\big\|\mathbf{X}^{A}\big\|^2_{p\text{-var};[0,1]}\big)<\infty.$$

If $\omega([s,t]^2)\le K|t-s|$ for all $s<t$ in $[0,1]$ we can replace $\|\mathbf{X}^A\|_{p\text{-var};[0,1]}$ by $\|\mathbf{X}^A\|_{1/p\text{-Höl};[0,1]}$.

We now discuss convergence results.

Theorem 51. *Let $A_n = \{1, \dots, n\}$. For any $p > 2\rho$ and $q \in [1, \infty)$,*

$$d_{p\text{-var};[0,1]}\big(\mathbf{X}^{A_n}, \mathbf{X}\big) \to 0 \quad \text{in } L^q(\Omega) \text{ as } n \to \infty, \tag{19}$$

$$\big\|\mathbf{X}^{A_n^c}\big\|_{p\text{-var};[0,1]} \to 0 \quad \text{in } L^q(\Omega) \text{ as } n \to \infty. \tag{20}$$

If ω is Hölder dominated, i.e. $\sup_{0\le s<t\le 1} \omega([s,t]^2)/|t-s|^{1/p} < \infty$, *then*

$$d_{1/p\text{-Höl};[0,1]}\big(\mathbf{X}^{A_n}, \mathbf{X}\big) \to 0 \quad \text{in } L^q(\Omega) \text{ as } n \to \infty, \tag{21}$$

$$\big\|\mathbf{X}^{A_n^c}\big\|_{1/p\text{-Höl};[0,1]} \to 0 \quad \text{in } L^q(\Omega) \text{ as } n \to \infty. \tag{22}$$

Proof. Ad (19), (21): From the results in Appendix A and Theorem 50, it is enough to prove that for any fixed $t \in [0,1]$,

$$d\big(\mathbf{X}_t^{A_n}, \mathbf{X}_t\big) \to 0$$

in L^q or, in fact, in probability (thanks to the uniform L^q-bounds for all $q < \infty$ in Theorem 50). The topology induced by d on $G^3(\mathbb{R}^d)$ is consistent with the manifold topology $G^3(\mathbb{R}^d) \subset T^3(\mathbb{R}^d)$ and in particular with the topology induced from the Euclidean structure on $g^3(\mathbb{R}^d) = \ln(G^3(\mathbb{R}^d))$, seen as global chart for $G^3(\mathbb{R}^d)$. It is therefore enough to show for $N = 1, 2, 3$ we have pointwise convergence,

$$\pi_N\big(\ln\big(\mathbf{X}_t^{A_n}\big) - \ln(\mathbf{X}_t)\big) \to 0 \quad \text{in probability.}$$

By martingale convergence, this is obvious for $N = 1, 2$ but for $N = 3$ we have to handle the correction which we identified in Lemma 47,

$$\left(\frac{1}{12} X_{s,t}^{A;j} R_{X^{A^c;i}}\begin{pmatrix} s & s \\ t & t \end{pmatrix} - \frac{1}{2}\int_s^t R_{X^{A^c;i}}\begin{pmatrix} u & s \\ t & u \end{pmatrix} \mathrm{d}X_u^{A,j}\right)\big[e_i, [e_i, e_j]\big].$$

All we need is pointwise convergence in probability to zero of this expression. Clearly, $X^{A_n^c} = \mathbb{E}[X|\mathcal{F}_{\{n+1,n+2,\dots\}}] \to 0$ a.s. and in all L^q as $n \to \infty$. It follows that $R_{X^{A_n^c;i}} \to 0$ pointwise which takes care of the first summand. The second term is a Young–Wiener integral in the sense of Proposition 38. From our uniform estimates and interpolation, $R_{X^{A_n^c;i}} \to 0$ in $(2+\varepsilon)$-variation. Using notation from the last proposition,

$$\int_s^t f(u)\,\mathrm{d}X_u^{A_n,j} = \mathbb{E}\left(\int_s^t f(u)\,\mathrm{d}X_u^j \,\Big|\, \mathcal{F}_{A_n}\right),$$

and it is enough to show that $\int_s^t f(u)\,\mathrm{d}X_u^j \to 0$ in L^2. Now,

$$|f|^{2+\varepsilon}_{(2+\varepsilon)\text{-var};[s,t]} \le C\,|R_{X^{A_n^c;i}}|^{2+\varepsilon}_{(2+\varepsilon)\text{-var};[s,t]^2} \to 0$$

and using the Young–Wiener estimate for ε chosen small enough (namely such that $(2+\varepsilon)^{-1} + \rho^{-1} > 1$ which is always possible since $\rho \in [1,2)$) we obtain the required convergence in L^2 and hence in probability as required.

Ad (20), (22): As in the first part of the proof, it is enough to show that, for fixed $t \in [0,1]$, $\mathbf{X}_t^{A_n^c} \to 0$ in probability or, equivalently,

$$\ln\big(\mathbf{X}_t^{A_n^c}\big) \to 0 \quad \text{in probability.}$$

We first claim that $\mathbb{E}(\ln(\mathbf{X}_t)|\mathcal{F}_{A_n^c}) \to 0$. Indeed, by backward martingale convergence and Kolmogorov's 0–1 law,

$$\begin{aligned}\mathbb{E}\big(\ln(\mathbf{X}_t)|\mathcal{F}_{A_n^c}\big) &\to \mathbb{E}\left(\ln(\mathbf{X}_t)\middle|\bigcap_k \mathcal{F}_{A_k^c}\right) \quad \text{a.s. and in all } L^q \\ &\overset{\text{a.s.}}{=} \mathbb{E}(\ln \mathbf{X}_t) = 0\end{aligned}$$

(where $\mathbb{E}(\ln \mathbf{X}_t) = 0$ follows from Lemma 47 with $A = \varnothing$). The proof will be finished if we can handle the difference between $\ln(\mathbf{X}_t^{A_n^c})$ and $\mathbb{E}(\ln(\mathbf{X}_t)|\mathcal{F}_{A_n^c})$. But using Lemma 47, this is done in the same way as in the first part of the proof. $\square$

6.3. *Support description*

We recall the standing assumptions. $X = (X^i\colon i = 1, \dots, d)$ is a centered continuous Gaussian process on $[0,1]$, with independent components and finite covariance of finite $\rho \in [1,2)$-variation, dominated by some 2D control ω. From Section 4.4, we know that, for $p \in (2\rho, 4)$, X lifts to a (random) geometric p-rough path $\mathbf{X}$ with a.e. sample path in $C_0^{0,p\text{-var}}([0,1], G^3(\mathbb{R}^d))$. If ω is Hölder dominated we have sample paths in $C_0^{0,1/p\text{-Höl}}([0,1], G^3(\mathbb{R}^d))$. It will be convenient in this section to assume that $\mathbb{P}$ is a Gaussian measure on $C([0,1], \mathbb{R}^d)$ so that $X(\omega) = \omega_t$ can be realized as coordinate process and $\mathbf{X}$ as measurable map from $C([0,1], \mathbb{R}^d)$ into $C_0^{0,p\text{-var}}([0,1], G^3(\mathbb{R}^d))$ resp. $C_0^{0,1/p\text{-Höl}}([0,1], G^3(\mathbb{R}^d))$, defined as

$$\mathbf{X}(\omega) = \lim_{n\to\infty} S_3\big(\omega^{D_n}\big)$$

in probability where ω^{D_n} denotes the piecewise linear approximation based on some dissection (D_n), assuming $|D_n| \to 0$. We shall also make the assumption that $\mathcal{H}$ enjoys *complementary Young regularity* by which we mean that $\mathcal{H} \hookrightarrow C_0^{q\text{-var}}([0,1], \mathbb{R}^d)$ for some $q \geq 1$ with $1/p + 1/q > 1$. Let us recall that the translation of a "smooth" path and its first three iterated integrals, $\mathbf{x} = S_3(x)$, in direction h is defined as $(\mathbf{x}, h) \mapsto S_3(x+h)$. By a closing procedure (cf. [27]) this map extends continuously to $(\mathbf{x}, h) \mapsto T_h\mathbf{x}$, known as *translation operator*, from

$$C_0^{0,p\text{-var}}\big([0,1], G^3\big(\mathbb{R}^d\big)\big) \times C_0^{q\text{-var}}\big([0,1], \mathbb{R}^d\big) \to C_0^{0,p\text{-var}}\big([0,1], G^3\big(\mathbb{R}^d\big)\big)$$

and hence from $C_0^{0,p\text{-var}}([0,1], G^3(\mathbb{R}^d)) \times \mathcal{H} \to C_0^{0,p\text{-var}}([0,1], G^3(\mathbb{R}^d))$. This also holds for $\mathbf{x} \in C_0^{0,1/p\text{-Höl}}([0,1], G^3(\mathbb{R}^d))$.

Lemma 52. *Assume complementary Young regularity of $\mathcal{H}$. Then, for $\mathbb{P}$-almost every ω we have*

$$\forall h \in \mathcal{H}\colon\ \mathbf{X}(\omega + h) = T_h\mathbf{X}(\omega),$$

where T denotes the translation operator for geometric rough paths.

Proof. By switching to a subsequence if needed we may assume that $\mathbf{X}(\omega)$ is defined as $\lim_{n\to\infty} S_3(\omega^{D_n})$ whenever this limit exists (and arbitrarily on the remaining null-set N). Now fix $h \in \mathcal{H}$; using complementary Young regularity we have

$$S_3\big(\omega^{D_n} + h^{D_n}\big) = T_{h^{D_n}} S_3\big(\omega^{D_n}\big) \to T_h\mathbf{X}(\omega) \quad \text{as } n \to \infty$$

and thus see that $\mathbf{X}(\omega + h) = T_h\mathbf{X}(\omega)$ for all h and $\omega \notin N$. $\square$

Lemma 53. *Assume complementary Young regularity of $\mathcal{H}$. Then, for every $h \in \mathcal{H}$ the laws of $\mathbf{X}$ and $T_h\mathbf{X}$ are equivalent.*

Proof. By Cameron–Martin, the law of X and $X+h$, as Borel measures on $C([0,1],\mathbb{R}^d)$ are equivalent. It follows that the image measures under the measurable map $\mathbf{X}(\cdot)$, Borel measures on $C_0^{0,p\text{-var}}([0,1],G^3(\mathbb{R}^d))$ resp. $C_0^{0,1/p\text{-Höl}}([0,1],G^3(\mathbb{R}^d))$, are equivalent. But this says precisely that the laws of $\mathbf{X}$ and $\mathbf{X}(\cdot+h)$ are equivalent and the proof if finished since $\mathbf{X}(\cdot+h)=T_h\mathbf{X}$ almost surely. □

Lemma 54. *Let S,S' be two Polish spaces and μ a Borel measure on S. Assume $x\in\operatorname{supp}\mu$ and f is continuous at x. Then $f(x)\in\operatorname{supp}f_*\mu$. If, in addition, $S'=S$ and $f_*\mu\sim\mu$ then $f(x)\in\operatorname{supp}\mu$.*

Proof. Write $B_\delta(x)$ for an open ball, centered at x of radius $\delta>0$. For every $\varepsilon>0$ there exists δ such that $B_\delta(x)\subset f^{-1}(B_\varepsilon(f(x)))$ and hence $0<\mu(B_\delta(x))\le(f_*\mu)(B_\varepsilon(f(x)))$ so that $f(x)\in\operatorname{supp}f_*\mu$. If $f_*\mu\sim\mu$ then and $0<(f_*\mu)(B_\varepsilon(f(x)))\Longrightarrow 0<\mu(B_\varepsilon(f(x)))$ and so $f(x)\in\operatorname{supp}\mu$. □

Theorem 55. *Let $\mathbf{X}_*\mathbb{P}$ denote the law of $\mathbf{X}$, a Borel measure on the Polish space $C_0^{0,p\text{-var}}([0,1],G^3(\mathbb{R}^d))$ where $p>2\rho$. Assume that $\mathcal{H}$ enjoys complementary Young regularity. Then*[12]

$$\operatorname{supp}[\mathbf{X}_*\mathbb{P}]=\overline{S_3(\mathcal{H})},$$

where support and closure are with respect to p-variation topology. If ω is Hölder dominated, i.e. $\omega([s,t]^2)\le K|t-s|$ for some constant K, we can use $1/p$-Hölder topology instead of p-variation topology.

Proof. As a preliminary remark, note that $S_3(\mathcal{H})$ is meaningful since any $h\in\mathcal{H}$ has finite ρ-variation (Proposition 17) and hence lifts canonicially to a $G^3(\mathbb{R}^d)$-valued paths (of finite ρ-variation) by iterated Young integration.

Step1: $\subset$-inclusion. Since $X^{\{1,\dots,n\}}:=\mathbb{E}[X_\cdot|\mathcal{F}_{\{1,\dots,n\}}]\in\mathcal{H}$ almost surely and converges to $\mathbf{X}$ in the respective rough path metrics, the first inclusion is clear.

Step2: $\supset$-inclusion. The idea is to find at least one fixed $\hat\omega\in C([0,1],\mathbb{R}^d)$ such that $\mathbf{X}(\hat\omega)\in\operatorname{supp}[\mathbf{X}_*\mathbb{P}]$ and such that there exists a (deterministic!) sequence $(g_n)\subset\mathcal{H}$, which can and will depend on $\hat\omega$, such that $T_{-g_n}\mathbf{X}(\hat\omega)=\mathbf{X}(\hat\omega-g_n)\to\mathbf{X}(0)=S_3(0)$ in rough path metric. Having found such an element $\hat\omega$ (with suitable sequence g_n) we can apply Lemma 54 with μ as the law of $\mathbf{X}$, a Borel measure on $S=C_0^{0,p\text{-var}}([0,1],G^3(\mathbb{R}^d))$ resp. $C_0^{0,1/p\text{-Höl}}([0,1],G^3(\mathbb{R}^d))$, $S'=S$ and *continuous* function $f:S\to S$ given by $f:\mathbf{x}\mapsto T_{-g_n}\mathbf{x}$; using that the law of $T_h\mathbf{X}$ is equivalent to the law of $\mathbf{X}$, cf. Lemma 53, we conclude that $T_{-g_n}\mathbf{X}(\hat\omega)\in\operatorname{supp}[\mathbf{X}_*\mathbb{P}]$. This holds true for all n and by closeness of the support, the limit $\mathbf{X}(0)=S_3(0)$ must be in the support. The same argument shows that any further translate $T_hS_3(0)=S_3(h)$ must be in the support and thus

$$\operatorname{supp}[\mathbf{X}_*\mathbb{P}]\supset S_3(\mathcal{H}).$$

Passing the (p-variation resp. $1/p$-Hölder rough path) closure on both sides then finishes the proof. It remains to see how to find $\hat\omega$ with the required properties: $\mathbf{X}(\hat\omega)\in\operatorname{supp}[\mathbf{X}_*\mathbb{P}]$ and $T_{-g_n}\mathbf{X}(\hat\omega)=\mathbf{X}(\hat\omega-g_n)$ hold true for almost every $\hat\omega$ and require no further consideration. Furthermore, Theorem 51 allows us to pick $\hat\omega$ in a set of full measure such that

$$\mathbf{X}(\hat\omega)=\lim_{m\to\infty}S_3\left(\sum_{i=1}^m\xi(h_k)\Big|_{\hat\omega}h_k(\cdot)\right)=\lim_{m\to\infty}\mathbf{X}^{\{1,\dots,m\}}(\hat\omega),$$

$$\mathbf{X}^{\{n+1,n+2,\dots\}}(\hat\omega)\to S_3(0).$$

It now suffices to set $g_n(\cdot)=\sum_{i=1}^n\xi(h_k)|_{\hat\omega}h_k(\cdot)\in\mathcal{H}\hookrightarrow C^{q\text{-var}}$; we then see that

$$\begin{aligned}\mathbf{X}(\hat\omega-g_n)&=T_{-g_n}\mathbf{X}(\hat\omega)=\lim_{m\to\infty}T_{-g_n}\mathbf{X}^{\{1,\dots,m\}}(\hat\omega)=\lim_{m\to\infty}\mathbf{X}^{\{n+1,\dots,m\}}(\hat\omega)\\&=\mathbf{X}^{\{n+1,n+2,\dots\}}(\hat\omega)\to\mathbf{X}(0)=S_3(0),\end{aligned}$$

as required, and this finishes the proof. □

[12] Thanks to $\mathcal{H}\hookrightarrow C^{\rho\text{-var}}$ and $\rho\in[1,2)$, for any $h\in\mathcal{H}$, $S_3(h)$ is canonically defined by iterated Young integration.

Remark 56. (i) *Theorem* 55 *may also be obtained by applying the abstract support theorem of Aida–Kusuoka–Stroock* [1], *Corollary* 1.13.

(ii) *The assumption that $\mathcal{H}$ enjoys complementary Young regularity also appears naturally in the context of Malliavin calculus for Gaussian rough paths* [6]. *Thanks to Proposition* 17, *a sufficient condition is finite ρ-variation of the covariance for $\rho < 3/2$; this covers, from very general principles, Brownian Motion and fractional Brownian Motion with $H > 1/3$. In fact, we can also cover the regime $H \in (1/4, 1/3]$: it suffices to use Besov regularity of $\mathcal{H}^H$, the Cameron–Martin space associated to fractional Brownian Motion, combined with the Besov-variation embedding theorem established in* [18].

(iii) *The assumption that $\mathcal{H}$ enjoys complementary Young regularity can be slightly relaxed. Namely, it suffices to assume that a dense subset of $\mathcal{H}$ has the correct complementary regularity. More precisely, it suffices to assume that for some $q \geq 1$ with $1/p + 1/q > 1$, there exists $(h_n) \subset C^{q\text{-var}}([0,1], \mathbb{R}^d) \cap \mathcal{H}$, which is dense in $\mathcal{H}$. Indeed, using a Gram–Schmidt orthonormalization procedure if necessary, we can assume without loss of generality that the (h_n) form an orthonormal basis in $\mathcal{H}$; it then suffices to run through the proof of Theorem* 55 *using this particular – rather than an arbitrary – orthonormal basis in $\mathcal{H}$.*

(iv) *In fact, we conjecture that Theorem* 55 *holds true without any "complementary regularity" assumption. The problem faced here is that, for fixed $g \in \mathcal{H}$ switching from $\mathbf{X}(\hat{\omega} - g)$ to $\mathbf{X}(\hat{\omega})$ cannot be realized as (continuous) operation T_{-g} on rough path space. We suspect that it will be necessary to construct a paired*[13] *rough path* $(\mathbf{X}, \mathbf{g})$, *such that translation can again be realized as continuous operation, as well as exhibiting Karhunen–Loève approximations as "good" approximations in the sense of* [9]; *an extension of Lemma* 46 *to all $\rho \in [1, 2)$ may also be relevant here. We hope to return to these matters in future work.*

Appendix A: L^q-convergence for rough paths

The following lemma is an elementary consequence of the definition of $\otimes$ and equivalence of homogenous norms.

Lemma 57. *Let $g, h \in G^N(\mathbb{R}^d)$. Then there exists $C = C(N, d)$ such that*

$$\left\| g^{-1} \otimes h \otimes g \right\| \leq C \max\left\{ \|h\|, \|h\|^{1/N} \|g\|^{1-1/N} \right\}.$$

Recall the notions of d_0 and d_∞ as defined in Section 1.1.

Proposition 58 (d_0/d_∞ estimate). *On the path-space $C_0([0,1], G^N(\mathbb{R}^d))$ the distances d_∞ and $d_0 \equiv d_{0\text{-Höl}}$ are locally $1/N$-Hölder equivalent. More precisely, there exists $C = C(N, d)$ such that*

$$d_\infty(\mathbf{x}, \mathbf{y}) \leq d_0(\mathbf{x}, \mathbf{y}) \leq C \max\left\{ d_\infty(\mathbf{x}, \mathbf{y}), d_\infty(\mathbf{x}, \mathbf{y})^{1/N} \left(\|\mathbf{x}\|_\infty + \|\mathbf{y}\|_\infty \right)^{1-1/N} \right\}.$$

Proof. Only the second inequality requires a proof. We write gh instead of $g \otimes h$. For any $s < t$ in $[0, 1]$,

$$\mathbf{x}_{st}^{-1} \mathbf{y}_{s,t} = \mathbf{x}_{st}^{-1} \mathbf{y}_s^{-1} \mathbf{x}_s \mathbf{x}_{st} \mathbf{x}_t^{-1} \mathbf{y}_t \mathbf{x}_t^{-1} \mathbf{y}_t \mathbf{y}_t^{-1} \mathbf{x}_t.$$

By sub-additivity,

$$\begin{aligned} \left\| \mathbf{x}_{st}^{-1} \mathbf{y}_{s,t} \right\| &\leq \left\| \mathbf{x}_{st}^{-1} \mathbf{y}_s^{-1} \mathbf{x}_s \mathbf{x}_{st} \right\| + \left\| \mathbf{x}_t^{-1} \mathbf{y}_t \mathbf{x}_t^{-1} \mathbf{y}_t \mathbf{y}_t^{-1} \mathbf{x}_t \right\| \\ &= \left\| v^{-1} \mathbf{y}_s^{-1} \mathbf{x}_s v \right\| + \left\| w^{-1} \mathbf{x}_t^{-1} \mathbf{y}_t w \right\| \end{aligned}$$

---

[13] The notation $(\mathbf{X}, \mathbf{g})$ is abusive and what we really mean is a $G^3(\mathbb{R}^d \oplus \mathbb{R}^d)$-valued geometric rough path which projects on $\mathbf{X}$ and $S_3(g)$ respectively. Note that $(\mathbf{X}, \mathbf{g})$ contains integrals of form $\int g \, \mathrm{d}X$ which are not well-defined Young-integrals but, in fact, Young–Wiener integrals (i.e. constructed in L^2-sense).

with $v = \mathbf{x}_{st}$ and $w = \mathbf{y}_t^{-1}\mathbf{x}_t$. Note that

$$\left\|\mathbf{y}_t^{-1}\mathbf{x}_t\right\| = \left\|\mathbf{x}_t^{-1}\mathbf{y}_t\right\| = d(\mathbf{x}_t, \mathbf{y}_t)$$

and $\|v\|, \|w\| \le \|\mathbf{x}\|_\infty + \|\mathbf{y}\|_\infty$. The conclusion now follows from Lemma 57. □

We recall the following simple interpolation result [17].

Lemma 59. *For $0 \le \alpha' < \alpha \le 1$ there exists a constant $C = C(\alpha, \alpha')$ such that*

$$d_{\alpha'\text{-Höl}}(\mathbf{x}, \mathbf{y}) \le C\big(\|\mathbf{x}\|_{\alpha\text{-Höl}} \vee \|\mathbf{y}\|_{\alpha\text{-Höl}}\big)^{\alpha'/\alpha} d_0(\mathbf{x}, \mathbf{y})^{1-\alpha'/\alpha}.$$

Similarly, for $1 \le p < p' < \infty$ there exists $C = C(p, p')$ such that

$$d_{p'\text{-var}}(\mathbf{x}, \mathbf{y}) \le C\big(\|\mathbf{x}\|_{p\text{-var}} \vee \|\mathbf{y}\|_{p\text{-var}}\big)^{p/p'} d_0(\mathbf{x}, \mathbf{y})^{1-p/p'}.$$

Corollary 60 (L^q-convergence in rough path metrics). *Let $\mathbf{X}^n, \mathbf{X}^\infty$ be continuous $G^N(\mathbb{R}^d)$-valued process defined on $[0, 1]$. Let $q \in [1, \infty)$ and assume that for some $\alpha \in (0, 1]$, (resp. $p \ge 1$),*

$$\sup_{1\le n\le\infty} \mathbb{E}\big(\left\|\mathbf{X}^n\right\|_{\alpha\text{-Höl}}^q\big) < \infty \quad \Big(\mathit{resp.}\ \sup_{1\le n\le\infty} \mathbb{E}\big(\left\|\mathbf{X}^n\right\|_{p\text{-var}}^q\big) < \infty\Big) \tag{23}$$

and that we have uniform convergence in $L^q(\mathbb{P})$ i.e.

$$d_\infty\big(\mathbf{X}^n, \mathbf{X}^\infty\big) \to 0 \quad \mathit{in}\ L^q(\mathbb{P}). \tag{24}$$

Then $d_{\alpha'\text{-Höl}}(\mathbf{X}^n, \mathbf{X}^\infty)$ for $\alpha' < \alpha$, (resp. $d_{p'\text{-var}}(\mathbf{X}^n, \mathbf{X}^\infty)$ and $p' > p$), converges to zero in $L^q(\mathbb{P})$.

Proof. From the d_0/d_∞ estimate there exists $c_1 > 0$ such that

$$\frac{1}{c_1} d_0\big(\mathbf{X}^n, \mathbf{X}^\infty\big) \le d_\infty\big(\mathbf{X}^n, \mathbf{X}^\infty\big) + d_\infty\big(\mathbf{X}^n, \mathbf{X}^\infty\big)^{1/N}\big(\left\|\mathbf{X}^n\right\|_\infty + \left\|\mathbf{X}^\infty\right\|_\infty\big)^{1-1/N}$$

and so

$$\frac{1}{c_1}\mathbb{E}\big(d_0\big(\mathbf{X}^n, \mathbf{X}^\infty\big)^q\big)^{1/q} \le \mathbb{E}\big(d_\infty\big(\mathbf{X}^n, \mathbf{X}^\infty\big)^q\big)^{1/q} + \mathbb{E}\big[d_\infty\big(\mathbf{X}^n, \mathbf{X}^\infty\big)^{q/N}\big(\left\|\mathbf{X}^n\right\|_\infty + \left\|\mathbf{X}^\infty\right\|_\infty\big)^{q(1-1/N)}\big]^{1/q}.$$

By Hölder's inequality,

$$\mathbb{E}\big[d_\infty\big(\mathbf{X}^n, \mathbf{X}\big)^{q/N}\left\|\mathbf{X}^n\right\|_\infty^{q(1-1/N)}\big] \le \mathbb{E}\big[d_\infty\big(\mathbf{X}^n, \mathbf{X}\big)^q\big]^{1/N}\mathbb{E}\big[\left\|\mathbf{X}^n\right\|_\infty^q\big]^{(1-1/N)}.$$

Since $\|\cdot\|_\infty$ is dominated by Hölder and variation norms, assumption (23) is plentiful to bound $\mathbb{E}(\|\mathbf{X}^n\|_\infty{}^q)$ uniformly in n. We thus obtain convergence of $d_0(\mathbf{X}^n, \mathbf{X}^\infty)$ to 0 in L^q. An almost identical application of Hölder's inequality, now using Lemma 59 instead of the d_0/d_∞ estimate, shows that $d_{\alpha'\text{-Höl}}(\mathbf{X}^n, \mathbf{X}^\infty)$, resp. $d_{p'\text{-var}}(\mathbf{X}^n, \mathbf{X}^\infty)$, converges to zero in $L^q(\mathbb{P})$. □

The assumption (24) can often be weakened to pointwise convergence.

Corollary 61. *Let $\mathbf{X}^n, \mathbf{X}^\infty$ be continuous $G^N(\mathbb{R}^d)$-valued process defined on $[0, 1]$. Let $q \in [1, \infty)$ and assume that we have pointwise convergence in $L^q(\mathbb{P})$ i.e. for all $t \in [0, 1]$,*

$$d\big(\mathbf{X}_t^n, \mathbf{X}_t^\infty\big) \to 0 \quad \mathit{in}\ L^q(\mathbb{P})\ \mathit{as}\ n \to \infty; \tag{25}$$

and uniform Hölder bounds, i.e.

$$\sup_{1\le n\le\infty} \mathbb{E}\big(\|\mathbf{X}^n\|^q_{\alpha\text{-Höl}}\big) < \infty$$

then for $\alpha' < \alpha$,

$$d_{\alpha'\text{-Höl}}\big(\mathbf{X}^n, \mathbf{X}^\infty\big) \to 0 \quad \textit{in } L^q(\mathbb{P}).$$

Proof. From the previous corollary, we only need to show d_∞-convergence in L^q. For any integer m,

$$\begin{aligned} 2^{1-q}\mathbb{E}\big[d_\infty\big(\mathbf{X}^n, \mathbf{X}^\infty\big)^q\big] &\le \mathbb{E}\Big[\sup_{i=1,\dots,m} d\big(\mathbf{X}^n_{i/m}, \mathbf{X}^\infty_{i/m}\big)^q\Big] + \mathbb{E}\Big[\sup_{|t-s|<1/m}\big(\|\mathbf{X}^n_{s,t}\|^q + \|\mathbf{X}^\infty_{s,t}\|^q\big)\Big] \\ &\le \sum_{i=1}^m \mathbb{E}\big(d\big(\mathbf{X}^n_{i/m}, \mathbf{X}^\infty_{i/m}\big)^q\big) + \left(\frac{1}{m}\right)^{\alpha q} \times 2 \sup_{1\le n\le\infty} \mathbb{E}\big[\|\mathbf{X}^n\|^q_{\alpha\text{-Höl}}\big]. \end{aligned}$$

By first choosing m large enough, followed by choosing n large enough we see that $d_\infty(\mathbf{X}^n, \mathbf{X}^\infty) \to 0$ in L^q as required. □

Corollary 62. *Let* $\mathbf{X}^n, \mathbf{X}^\infty$ *be continuous* $G^N(\mathbb{R}^d)$*-valued process defined on* $[0, 1]$. *Let* $q \in [1, \infty)$ *and assume that we have pointwise convergence in* $L^q(\mathbb{P})$ *i.e. for all* $t \in [0, 1]$,

$$d\big(\mathbf{X}^n_t, \mathbf{X}^\infty_t\big) \to 0 \quad \textit{in } L^q(\mathbb{P}) \textit{ as } n \to \infty;$$

uniform p*-variation bounds*,

$$\sup_{1\le n\le\infty} \mathbb{E}\big(\|\mathbf{X}^n\|^q_{p\text{-var}}\big) < \infty \tag{26}$$

and a tightness condition

$$\lim_{\varepsilon\to 0}\sup_n \mathbb{E}\big(\big|\operatorname{osc}\big(\mathbf{X}^n, \varepsilon\big)\big|^q\big) = 0, \tag{27}$$

where $\operatorname{osc}(\mathbf{X}, \varepsilon) \equiv \sup_{|t-s|<\varepsilon} \|\mathbf{X}_{s,t}\|$, *then*

$$d_{p\text{-var}}\big(\mathbf{X}^n, \mathbf{X}^\infty\big) \to 0 \quad \textit{in } L^q(\mathbb{P}).$$

Conditions (26) *and* (27) *are implied by a Kolmogorov type tightness criterion: there exists a* 1*D control function* ω *and a real number* $\theta \ge \frac{1}{2q} + \frac{1}{p}$ *such that for all* $s < t$ *in* $[0, 1]$,

$$\sup_{1\le n\le\infty} \mathbb{E}\big(\big|d\big(\mathbf{X}^n_s, \mathbf{X}^n_t\big)\big|^q\big)^{1/q} \le \omega(s, t)^\theta. \tag{28}$$

Proof. From our criterion for L^q-convergence in rough path metrics, we only need to show d_∞-convergence in L^q, which is an obvious consequence of the inequality

$$d_\infty\big(\mathbf{X}^n, \mathbf{X}^\infty\big) \le \operatorname{osc}\big(\mathbf{X}^n, 1/m\big) + \operatorname{osc}(\mathbf{X}, 1/m) + \sup_{i=1,\dots,m} d\big(\mathbf{X}^n_{i/m}, \mathbf{X}^\infty_{i/m}\big).$$

Finally, the assumption (28) implies (26) and (27) as an application of Corollary 66. (The bound on θ comes from $q_0 = (1/r - 1/p)^{-1}/2$.) □

Remark 63. *One cannot get rid of the tightness condition. Consider* $f_n(t) = 0$ *on* $[0, 1/n]$ *and a triangle peak of height on* $[1/n, 1]$. *Clearly,* $f_n(t) \to 0$ *a.s. (and hence in measure) and the* 1*-variation of* $\{f_n\}$ *is uniformly bounded. Yet,* $f_n \nrightarrow 0$ *in any variation topology which is stronger than the uniform topology.*

Appendix B: Garsia–Rodemich–Rumsey

Similarly to the last appendix, but more quantitatively, we aim for conditions under which $G^N(\mathbb{R}^d)$-valued processes are close in Hölder- (resp. variation-)/$L^q(\mathbb{P})$ sense. When possible, we formulate regularity results in the more general setting of paths (or processes) with values in a Polish space (E, d). The following result is well known (e.g. [35]) for $\mathbb{R}$-valued functions but the arguments extend trivially to the case of E-valued functions.

Theorem 64 (Garsia–Rodemich–Rumsey). *Let Ψ and p be continuous strictly increasing functions on $[0, \infty)$ with $p(0) = \Psi(0) = 0$ and $\Psi(x) \to \infty$ as $x \to \infty$. Given $f \in C([0, 1], E)$, if*

$$\int_0^1 \int_0^1 \Psi\left(\frac{d(f_s, f_t)}{p(|t-s|)}\right) \mathrm{d}s\, \mathrm{d}t \le F, \tag{29}$$

then for $0 \le s < t \le 1$,

$$d(f_s, f_t) \le 8 \int_0^{t-s} \Psi^{-1} x\left(\frac{4F}{u^2}\right) \mathrm{d}p(u).$$

In particular, if $\operatorname{osc}(f, \delta) \equiv \sup_{|t-s|\le\delta} d(f_s, f_t)$ denotes the modulus of continuity of f, we have

$$\operatorname{osc}(f, \delta) \le 8 \int_0^{\delta} \Psi^{-1}\left(\frac{4F}{u^2}\right) \mathrm{d}p(u).$$

Corollary 65. *Let $r \ge 1$ and $\alpha \in [0, 1/r)$. Then, for any fixed $q \ge q_0(r, \alpha)$,*

$$\int_0^1 \int_0^1 \frac{d(f_s, f_t)^q}{|t-s|^{q/r}} \mathrm{d}s\, \mathrm{d}t \le M^q,$$

implies the existence of $C = C(r, \alpha)$ such that $\operatorname{osc}(f, \delta) \le C\delta^{\alpha} M$ and

$$\|f\|_{\alpha\text{-Höl};[0,1]} \le CM.$$

Proof. From Garsia–Rodemich–Rumsey with $\Psi(x) = x^q$, $p(u) = u^{1/r}$ and $F = M^q$ it follows that

$$d(f_s, f_t) \le 8(4F)^{1/q} \int_0^{t-s} u^{-2/q+1/r-1} \mathrm{d}u = \frac{8(4F)^{1/q}}{1/r - 2/q} |t-s|^{1/r-2/q} \le \frac{32M}{1/2r} |t-s|^{\alpha}$$

provided q is large enough so that $0 \le \alpha < 1/r - 2/q$ and $1/r - 2/q > 1/(2r)$. Both statements follow. (One can take $q_0 = (1/r - \alpha)^{-1}/2 \vee 4r$ and $C = 64/r$. Alternatively, at least if $\alpha > 0$, one can take $q_0 = (1/r - \alpha)^{-1}/2$ and $C = 32/\alpha$.) □

Corollary 66. *Let ω be a 1D control function and X a continuous (E, d)-valued stochastic process defined on $[0, 1]$. Assume $r \ge 1$ and $1/p \in [0, 1/r)$. Then, for any fixed $q \ge q_0(r, p)$ (one can take $q_0 = (1/r - 1/p)^{-1}/2$)*

$$\left|d(X_s, X_t)\right|_{L^q(\mathbb{P})} \le M\omega(s, t)^{1/r} \quad \text{for all } s, t \in [0, 1]$$

implies $\operatorname{osc}(X, \delta) \to 0$ in $L^q(\mathbb{P})$ as $\delta \to 0$ and there exists $C = C(r, p)$ such that

$$\left|\|X\|_{p\text{-var};[0,1]}\right|_{L^q(\mathbb{P})} \le CM\omega(0, 1)^{1/r}.$$

If $\omega(s, t) \le t - s$ for all $s, t \in [0, 1]$ then $\|X\|_{p\text{-var};[0,1]}$ above can be replaced by $\|X\|_{1/p\text{-Höl};[0,1]}$.

Proof. We first consider the case of Hölder dominated control $\omega(s,t) \le t-s$. We set $\alpha := 1/p$. From the preceding corollary,

$$\|X\|^q_{\alpha\text{-Höl};[0,1]} \le C^q \int_0^1 \int_0^1 \frac{d(X_s,X_t)^q}{|t-s|^{q/r}} \,\mathrm{d}s\,\mathrm{d}t$$

and taking expectations gives

$$\mathbb{E}\big(\|X\|^q_{\alpha\text{-Höl};[0,1]}\big) \le C^q \int_0^1 \int_0^1 \frac{\mathbb{E}(d(X_s,X_t)^q)}{|t-s|^{q/r}} \,\mathrm{d}s\,\mathrm{d}t \le (CM)^q$$

which shows $|\|X\|_{\alpha\text{-Höl};[0,1]}|_{L^q(\mathbb{P})} \le CM$. The statement on $\operatorname{osc}(X,\delta)$ obvious. We now discuss a general control ω. At the price of replacing M by $M\omega(0,1)^{1/r}$, we assume $\omega(0,1)=1$. The function $\omega(t) := \omega(0,t)$ maps $[0,1]$ continuously and increasingly onto $[0,1]$ and there exists a continuous process Y such that $Y_{\omega(t)} = X_t$ for all $t \in [0,1]$. We then have, for all $s,t \in [0,1]$,

$$\mathbb{E}\big(d(Y_s,Y_t)^q\big)^{1/q} \le M|t-s|^{1/r}.$$

By the Hölder case just discussed, $\operatorname{osc}(Y,\delta) \to 0$ (in L^q) and so $\operatorname{osc}(X,\delta) \to 0$ in L^q by (uniform) continuity of ω. The Hölder case also takes care of the L^q-bound of $\|X\|_{p\text{-var};[0,1]}$, it suffices to note

$$\|X\|_{p\text{-var};[0,1]} = \|Y\|_{p\text{-var};[0,1]} \le \|Y\|_{1/p\text{-Höl}[0,1]}. \qquad \square$$

We now consider paths with values in $G^N(\mathbb{R}^d)$ for which we can of increments, $x_{s,t} \equiv x_s^{-1} \otimes x_t$, and thus of Hölder- and variation distance.

Corollary 67. *Let $r \ge 1$ and $\alpha \in [0,1/r)$. Then, for any $q \ge q_0(r,\alpha)$ and $M > 0, \delta \in (0,1)$,*

$$\int_0^1 \int_0^1 \frac{d(x_s,x_t)^q}{|t-s|^{q/r}} \,\mathrm{d}s\,\mathrm{d}t \le M^q,$$
$$\int_0^1 \int_0^1 \frac{d(y_s,y_t)^q}{|t-s|^{q/r}} \,\mathrm{d}s\,\mathrm{d}t \le M^q,$$
$$\int_0^1 \int_0^1 \frac{d(x_{s,t},y_{s,t})^q}{|t-s|^{q/r}} \,\mathrm{d}s\,\mathrm{d}t \le (\delta M)^q,$$

implies the existence of $C = C(r;N,d), \theta = \theta(r,\alpha;N) > 0$ such that

$$d_{\alpha\text{-Höl}[0,1]}(x,y) \le C\delta^\theta M.$$

Proof. We first note that with $\alpha' = (\alpha + 1/r)/2$ and assuming $q \ge q_0(r,\alpha)$ large enough, Corollary 65 implies

$$\|x\|_{0;[0,1]} \le \|x\|_{\alpha\text{-Höl};[0,1]} \le \|x\|_{\alpha'\text{-Höl};[0,1]} \le c_0 M \tag{30}$$

and the same estimate holds for y. Let us define $z_t = y_t \otimes x_t^{-1}$. Since $z_{s,t} \equiv z_s^{-1} \otimes z_t = x_t \otimes (x_{s,t}^{-1} \otimes y_{s,t}) \otimes x_t^{-1}$ Lemma 57 gives

$$\frac{1}{c_1}\|z_{s,t}\| \le d(x_{s,t},y_{s,t}) \vee d(x_{s,t},y_{s,t})^{1/N}\|x_t\|^{1-1/N}.$$

Dividing by $|t-s|^{1/rN}$ and raising everything to power q yields

$$\left(\frac{1}{c_1}\frac{\|z_{s,t}\|}{|t-s|^{1/(rN)}}\right)^q \le \left(\frac{d(x_{s,t},y_{s,t})}{|t-s|^{1/r}}\right)^q \vee \left(\frac{d(x_{s,t},y_{s,t})^q}{|t-s|^{q/r}}\right)^{1/N} \|x_t\|^{q(1-1/N)}$$

and after integration over $(s,t)\in[0,1]^2$, using Hölder's inequality on the last term, we arrive at

$$\begin{aligned}
&\left(\frac{1}{c_1}\right)^q \int_0^1\int_0^1 \left(\frac{\|z_{s,t}\|}{|t-s|^{1/(rN)}}\right)^q \mathrm{d}s\,\mathrm{d}t \\
&\quad\le \int_0^1\int_0^1 \left(\frac{d(x_{s,t},y_{s,t})}{|t-s|^{1/r}}\right)^q \mathrm{d}s\,\mathrm{d}t \\
&\qquad + \left(\int_0^1\int_0^1 \left(\frac{d(x_{s,t},y_{s,t})^q}{|t-s|^{q/r}}\right)\mathrm{d}s\,\mathrm{d}t\right)^{1/N} \left(\int_0^1\int_0^1 \|x_t\|^q\,\mathrm{d}s\,\mathrm{d}t\right)^{1-1/N} \\
&\quad\le (\delta M)^q + (\delta M)^{q/N}\big(\|x\|_{0\text{-Höl};[0,1]}\big)^{q(1-1/N)} \\
&\quad\le (\delta M)^q + (\delta M)^{q/N}(c_0 M)^{q(1-1/N)} \quad \text{by (30)} \\
&\quad\le \big(c_2\delta^{1/N}M\big)^q.
\end{aligned}$$

We can then apply Corollary 65 to z (with M replaced by $c_1c_2\delta^{1/N}M$) to see that $\|z\|_{0;[0,1]}\le c_3\delta^{1/N}M$. On the other hand, $d(x_{s,t},y_{s,t})=\|x_t^{-1}\otimes z_{s,t}\otimes x_t\|$ and Lemma 57 implies, again using (30),

$$d_0(x,y)\le c_1\max\left\{\|z\|_{0\text{-Höl}},\ \|z\|_{0\text{-Höl}}^{1/N}\cdot\|x\|_{0\text{-Höl}}^{1-1/N}\right\}\le c_4\delta^{1/N^2}M.$$

We now use interpolation, Lemma 59, with Hölder exponents $\alpha<\alpha'$. For $c_5=c_5(\alpha,r)$ and again using (30) we have

$$\begin{aligned}
d_{a\text{-Höl}}(x,y) &\le c_5\big(\|x\|_{\alpha'\text{-Höl}}\vee\|y\|_{\alpha'\text{-Höl}}\big)^{\alpha/\alpha'} d_0(x,y)^{1-\alpha/\alpha'} \\
&\le c_5(c_0M)^{\alpha/\alpha'}\big(c_4\delta^{1/N^2}M\big)^{1-\alpha/\alpha'} \\
&= c_6M\delta^{\theta} \quad \text{with } \theta=\theta(\alpha,r,N):=\frac{\alpha'-\alpha}{\alpha' N^2}.
\end{aligned}$$

The proof is finished. □

Corollary 68. *Assume that X,Y are continuous $G^N(\mathbb{R}^d)$-valued processes defined on $[0,1]$. Assume $r\ge 1$ and $1/p\in[0,1/r)$. Then, for any fixed $q\ge q_0(r,p)$ and $M,\delta\in(0,1)$,*

$$\begin{aligned}
&\mathbb{E}\big(d(X_s,X_t)^q\big)\le\big(M\omega(s,t)^{1/r}\big)^q, \\
&\mathbb{E}\big(d(Y_s,Y_t)^q\big)\le\big(M\omega(s,t)^{1/r}\big)^q, \\
&\mathbb{E}\big(d(X_{s,t},Y_{s,t})^q\big)\le\big(\delta M\omega(s,t)^{1/r}\big)^q,
\end{aligned}$$

implies the existence of $C=C(r;N,d),\theta=\theta(r,p;N)>0$ such that

$$\big|d_{p\text{-var};[0,1]}(X,Y)\big|_{L^q(\mathbb{P})}\le C\delta^{\theta}M.$$

If $\omega(s,t)\le t-s$ for all $s,t\in[0,1]$ then $d_{p\text{-var};[0,1]}$ above can be replaced by $d_{1/p\text{-Höl}[0,1]}$.

Proof. By a (deterministic) time-change argument, exactly as in the proof of Corollary 66, we may assume $\omega(s,t)=t-s$. From Corollary 67 there exists $c_1=c_1(r;N,d),\theta=\theta(r,p;N)$ such that for $q\ge q_0(r,p)$ large enough

$$\begin{aligned}
\left(\frac{1}{c_1\delta^{\theta}}d_{1/p\text{-Höl}[0,1]}(X,Y)\right)^q &\le \int_0^1\int_0^1\frac{d(x_s,x_t)^q}{|t-s|^{q/r}}\,\mathrm{d}s\,\mathrm{d}t+\int_0^1\int_0^1\frac{d(y_s,y_t)^q}{|t-s|^{q/r}}\,\mathrm{d}s\,\mathrm{d}t \\
&\quad+\left(\frac{1}{\delta^q}\int_0^1\int_0^1\frac{d(x_{s,t},y_{s,t})^q}{|t-s|^{q/r}}\,\mathrm{d}s\,\mathrm{d}t\right).
\end{aligned}$$

After taking expectations we see that $(c_1\delta^\theta)^{-q}\mathbb{E}(d_{1/p\text{-Höl}[0,1]}(X,Y)^q) \le 3M^q$ and the proof is easily finished. (One can take $C = 3c_1$.) $\square$

Appendix C: Step 3 Lie algebra

As usual, $e_1, \dots, e_d$ denotes the standard basis in $\mathbb{R}^d$. A vector space basis of the Lie algebra $g_2(\mathbb{R}^d)$ is given by

$$\{(e_i), ([e_i, e_j])_{i<j}\}$$

and if $x : [s,t] \to \mathbb{R}^d$ be a smooth path with signature $S(x)_{s,t} = \mathbf{x}_{s,t}$ then its log-signature satisfies, trivially,

$$\pi_1(\ln \mathbf{x}_{s,t}) = \sum_i x^i_{s,t} e_i \in \mathbb{R}^d$$

and

$$\pi_2(\ln \mathbf{x}_{s,t}) = \frac{1}{2}\sum_{i<j}(\mathbf{x}^{i,j}_{s,t} - \mathbf{x}^{j,i}_{s,t})[e_i, e_j] \in \mathrm{so}(d).$$

We aim for a similar understanding of $g_3(\mathbb{R}^d)$. We leave the following simple technical lemma to the reader:

Lemma 69 (Step 3 Hall expansion). *A vector space basis of the Lie algebra $g_3(\mathbb{R}^d)$ is given by*

$$\{(e_i), ([e_i, e_j])_{i<j}, ([e_i,[e_j,e_k]])_{j\le i<k \text{ or } j<k\le i}\}$$

for $i, j, k \in \{1, \dots, d\}$, known as Philip–Hall Lie basis. For any 3-tensor α, the following identity holds:

$$\sum_{i,j,k}\alpha_{i,j,k}[e_i,[e_j,e_k]] = \sum_{\substack{j<i<k\\ \text{or}\\ j<k<i}}(\alpha_{i,j,k} - \alpha_{i,k,j} + \alpha_{j,i,k} - \alpha_{j,k,i})[e_i,[e_j,e_k]] + \sum_{i\ne j}(\alpha_{i,i,j} - \alpha_{i,j,i})[e_i,[e_i,e_j]].$$

Proposition 70. *Let $x : [s,t] \to \mathbb{R}^d$ be a smooth path with lift $S(x) = \mathbf{x}$. Then its log-signature projected to the third level, $\pi_3(\ln \mathbf{x}_{s,t})$, expands to in the Hall-basis as follows.*

$$\begin{aligned}\pi_3(\ln \mathbf{x}_{s,t}) = {}& \frac{1}{6}\sum_{\substack{j<i<k\\ \text{or}\\ j<k<i}}(\mathbf{x}^{i,j,k}_{s,t} + \mathbf{x}^{j,i,k}_{s,t} - 2\mathbf{x}^{i,k,j}_{s,t} + \mathbf{x}^{k,i,j}_{s,t} - 2\mathbf{x}^{j,k,i}_{s,t} + \mathbf{x}^{k,j,i}_{s,t})[e_i,[e_j,e_k]]\\ &+\sum_{i\ne j}\left\{\mathbf{x}^{i,i,j}_{s,t} + \frac{1}{12}|x^i_{s,t}|^2 x^j_{s,t} - \frac{1}{2}x^i_{s,t}\mathbf{x}^{i,j}_{s,t}\right\}[e_i,[e_i,e_j]].\end{aligned}$$

This identity remains valid for (weak) geometric rough paths.

Proof. Without loss of generalities $x : [0,1] \to \mathbb{R}^d$ and $x(0) = 0$. The signature of concatenated paths is given by the group product in the free group. Specializing to the step 3 group (viewed as subset of the enveloping tensor algebra), the smooth path $x = x|_{[0,t+\mathrm{d}t]}$ is the concatenation of $x|_{[0,t]}$ and $x|_{[t,t+\mathrm{d}t]}$. We have $S(x_{t+\mathrm{d}t}) = S(x_t) \otimes \exp(\mathrm{d}x_t)$ and by sending $\mathrm{d}t \to 0$ it is easy to (re-)derive the usual control ODE for lifted paths

$$\mathrm{d}\mathbf{x}_t = U_i(\mathbf{x})\,\mathrm{d}x^i,$$

where $\mathbf{x}_t = S(x)_t$ and $U_i(\mathbf{x}) = \mathbf{x} \otimes e_i$. To understand the evolution in the step 3 Lie algebra we write

$$\mathbf{z}_i(t) = \pi_i(\ln S(x)_t), \quad i = 1, 2, 3$$

and using the Baker–Campbell–Hausdorff formula, we obtain[14],

$$\begin{aligned}
\mathrm{d}\mathbf{z}_1(t) &= \mathrm{d}x_t,\\
\mathrm{d}\mathbf{z}_2(t) &= \frac{1}{2}\big[\mathbf{z}_1(t), \mathrm{d}x_t\big],\\
\mathrm{d}\mathbf{z}_3(t) &= \frac{1}{2}\big[\mathbf{z}_2(t), \mathrm{d}x_t\big] + \frac{1}{12}\big[\mathbf{z}_1(t), \big[\mathbf{z}_1(t), \mathrm{d}x_t\big]\big],
\end{aligned}$$

which integrates iteratively to

$$\begin{aligned}
\mathbf{z}_1(t) &= x_t,\\
\mathbf{z}_2(t) &= \frac{1}{2}\int_{0<u<v<t} [\mathrm{d}x_u, \mathrm{d}x_v],\\
\mathbf{z}_3(t) &= \frac{1}{4}\int_{0<u<v<w<t} \big[[\mathrm{d}x_u, \mathrm{d}x_v], \mathrm{d}x_w\big] + \frac{1}{12}\int_{0<u<t} \big[x_u, [x_u, dx_u]\big].
\end{aligned}$$

In particular, the log-signature of x projected to the third level is precisely $\mathbf{z}_3(1)$ and given by

$$\begin{aligned}
&\frac{1}{4}\int_{0<u<v<w<t} \big[[\mathrm{d}x_u, \mathrm{d}x_v], \mathrm{d}x_w\big] + \frac{1}{12}\int_{0<u<t} \big[x_u, [x_u, \mathrm{d}x_u]\big]\\
&\quad = \frac{1}{4}\sum_{i,j,k} \mathbf{x}^{i,j,k}\big[[e_i, e_j], e_k\big] + \frac{1}{12}\sum_{i,j,k}\big(\mathbf{x}^{i,j,k} + \mathbf{x}^{j,i,k}\big)\big[e_i, [e_j, e_k]\big]\\
&\quad = \frac{1}{12}\sum_{i,j,k}\big(-3\mathbf{x}^{j,k,i} + \mathbf{x}^{i,j,k} + \mathbf{x}^{j,i,k}\big)\big[e_i, [e_j, e_k]\big].
\end{aligned}$$

Using the step 3 Hall expansion lemma, a few lines of computations give

$$\begin{aligned}
6\mathbf{z}_3(1) = &\sum_{\substack{j<i<k\\ \text{or}\\ j<k<i}} \big(\mathbf{x}_t^{i,j,k} + \mathbf{x}_t^{j,i,k} - 2\mathbf{x}_t^{i,k,j} + \mathbf{x}_t^{k,i,j} - 2\mathbf{x}_t^{j,k,i} + \mathbf{x}_t^{k,j,i}\big)\big[e_i, [e_j, e_k]\big]\\
&+ \sum_{i\neq j}\big(-2\mathbf{x}_t^{i,j,i} + \mathbf{x}_t^{i,i,j} + \mathbf{x}_t^{j,i,i}\big)\big[e_i, [e_i, e_j]\big].
\end{aligned}$$

When x is defined on $[s,t]$ the last expression is, of course,

$$\frac{1}{6}\big(\mathbf{x}_{s,t}^{i,i,j} - 2\mathbf{x}_{s,t}^{i,j,i} + \mathbf{x}_{s,t}^{j,i,i}\big)$$

and can be simplified with some calculus. We have

$$\begin{aligned}
\mathbf{x}_{s,t}^{i,i,j} &= \frac{1}{2}\int_s^t \big|x_{s,u}^i\big|^2 \mathrm{d}x_u^j,\\
\mathbf{x}_{s,t}^{i,j,i} + \mathbf{x}_{s,t}^{j,i,i} &= \int_{s<u<t} x_{s,u}^i x_{s,u}^j \,\mathrm{d}x_u^i\\
&= \frac{1}{2}\big|x_{s,t}^i\big|^2 x_{s,t}^j - \frac{1}{2}\int_{s<u<t}\big|x_{s,u}^i\big|^2 \mathrm{d}x_{s,u}^j \quad \text{(by integration by part)},
\end{aligned}$$

[14]A recursion formula for $\mathbf{z}_n$ appears in Chen's seminal work [8].

$$\begin{aligned}
\mathbf{x}_{s,t}^{j,i,i} &= \int_{s<u_1<u_2<u_3<t} \mathrm{d}x_{u_1}^j \,\mathrm{d}x_{u_2}^i \,\mathrm{d}x_{u_3}^i \\
&= \frac{1}{2}\int_{s<u<t} \left|x_{u,t}^i\right|^2 \mathrm{d}x_u^j \\
&= \frac{1}{2}\left|x_{s,t}^i\right|^2 x_{s,t}^j - x_{s,t}^i \int_{s<u<t} x_{s,u}^i \,\mathrm{d}x_{s,u}^j + \frac{1}{2}\int_s^t \left|x_{s,u}^i\right|^2 \mathrm{d}x_u^j
\end{aligned}$$

and therefore

$$\begin{aligned}
\mathbf{x}_{s,t}^{i,i,j} - 2\mathbf{x}_{s,t}^{i,j,i} + \mathbf{x}_{s,t}^{j,i,i} &= \mathbf{x}_{s,t}^{i,i,j} - 2\big(\mathbf{x}_{s,t}^{i,j,i} + \mathbf{x}_{s,t}^{j,i,i}\big) + 3\mathbf{x}_{s,t}^{j,i,i} \\
&= 3\int_s^t \left|x_{s,u}^i\right|^2 \mathrm{d}x_u^j + \frac{1}{2}\left|x_{s,t}^i\right|^2 x_{s,t}^j - 3x_{s,t}^i \mathbf{x}_{s,t}^{i,j}.
\end{aligned}$$

For the final statement, it suffices to remark that a (weak) geometric rough path is, in particular, a pointwise limit of smooth paths. We also note the following simple result. □

Lemma 71. *Let a, b be two elements of the Lie algebra $g_3(\mathbb{R}^d)$ and write $a^i = \pi_i(a)$, $b_i = \pi_i(b)$. Then*

$$\begin{aligned}
&\left|\pi_2\big(\ln\big(\mathrm{e}^{-a} \otimes \mathrm{e}^{b}\big)\big)\right| \le \left|b^2 - a^2\right| + \left|b^1 - a^1\right|\left|b^1\right|, \\
&\left|\pi_3\big(\ln\big(\mathrm{e}^{-a} \otimes \mathrm{e}^{b}\big)\big)\right| \le \left|b^3 - a^3\right| + \left|b^2 - a^2\right|\left|b^1\right| + \left|b^1 - a^1\right|\left(\left|b^2\right| + \frac{1}{3}\left|a^1\right|^2 + \frac{1}{3}\left|b^1\right|^2\right).
\end{aligned}$$

Proof. We only deal with the level 3 estimate and leave the (similar, but easier) step 2 estimate to the reader. From the Campbell–Baker–Hausdorff formula,

$$\ln\big(\mathrm{e}^{-a} \otimes \mathrm{e}^{b}\big) = -a + b + \frac{1}{2}[-a, b] + \frac{1}{12}\big[a, [a, b]\big] + \frac{1}{12}\big[b, [b, -a]\big].$$

By antisymmetry of the bracket, $[-a, b] = [b - a, b]$ so that

$$\pi_3[-a, b] = \pi_3[b - a, b] = \pi_3\big((b - a) \otimes b\big) - \pi_3\big(b \otimes (b - a)\big)$$

and since

$$\left|\pi_3\big((b - a) \otimes b\big)\right| = \left|\big(b^2 - a^2\big) \otimes b^1 + \big(b^1 - a^1\big) \otimes b^2\right|$$

(and similar for $\pi_3(b \otimes (b - a))$) we see that

$$\left|\pi_3\left(\frac{1}{2}[-a, b]\right)\right| \le \left|\big(b^2 - a^2\big)\right|\left|b^1\right| + \left|b^1 - a^1\right|\left|b^2\right|.$$

The same reasoning applies to $[a, [a, b]] = [a, [a, b - a]]$ and $[b, [b, -a]] = [b, [b, b - a]]$ and gives $|\pi_3[a, [a, b]]| \le 4|a^1|^2|b^1 - a^1|$, $|\pi_3[b, [b, -a]]| \le 4|b^1|^2|b^1 - a^1|$. Combing these estimates finishes the proof. □

Acknowledgement

The authors would like to thank an anonymous referee for his careful reading and in particular for pointing out a problem in the original proof of Theorem 55.

References

[1] S. Aida, S. Kusuoka and D. Stroock. On the support of Wiener functionals. In *Asymptotic Problems in Probability Theory: Wiener Functionals and Asymptotics (Sanda/Kyoto, 1990)* 3–34. *Pitman Res. Notes Math. Ser.* **284**. Longman Sci. Tech., Harlow, 1993. MR1354161

[2] P. Billingsley. *Convergence of Probability Measures*. Wiley, New York, 1968. MR0233396

[3] C. Borell. Tail probabilities in Gauss space. In *Vector Space Measures and Applications (Proc. Conf., Univ. Dublin, Dublin, 1977), II* 73–82. *Lecture Notes in Phys.* **77**. Springer, Berlin, 1978. MR0502400

[4] C. Borell. On polynomial chaos and integrability. *Probab. Math. Statist.* **3** (1984) 191–203. MR0764146

[5] C. Borell. On the Taylor series of a Wiener polynomial. In *Seminar Notes on Multiple Stochastic Integration, Polynomial Chaos and Their Integration.* Case Western Reserve University, Cleveland, 1984.

[6] T. Cass and P. Friz. Densities for rough differential equations under Hoermander's condition. *Ann. of Math.* Accepted, 2008. Available at http://pjm.math.berkeley.edu/scripts/coming.php?jpath=annals.

[7] T. Cass, P. Friz and N. Victoir. Non-degeneracy of Wiener functionals arising from rough differential equations. *Trans. Amer. Math. Soc.* **361** (2009) 3359–3371.

[8] K.-T. Chen. Integration of paths, geometric invariants and a generalized Baker–Hausdorff formula. *Ann. of Math. (2)* **65** (1957) 163–178. MR0085251

[9] L. Coutin, P. Friz and N. Victoir. Good rough path sequences and applications to anticipating stochastic calculus. *Ann. Probab.* **35** (2007) 1172–1193. MR2319719

[10] L. Coutin and Z. Qian. Stochastic analysis, rough path analysis and fractional Brownian motions. *Probab. Theory Related Fields* **122** (2002) 108–140. MR1883719

[11] L. Coutin and N. Victoir. Enhanced Gaussian processes and applications. Preprint, 2005.

[12] J.-D. Deuschel and D. W. Stroock. *Large Deviations*. *Pure and Applied Mathematics* **137**. Academic Press, Boston, MA, 1989. MR0997938

[13] R. M. Dudley and R. Norvaiša. *Differentiability of Six Operators on Nonsmooth Functions and p-variation*. *Lecture Notes in Mathematics* **1703**. Springer, Berlin, 1999. (With the collaboration of Jinghua Qian). MR1705318

[14] D. Feyel and A. de La Pradelle. Curvilinear integrals along enriched paths. *Electron. J. Probab.* **11** (2006) 860–892. MR2261056

[15] P. Friz and H. Oberhauser. Isoperimetry and rough path regularity. arXiv-preprint, 2007.

[16] P. Friz and N. Victoir. Approximations of the Brownian rough path with applications to stochastic analysis. *Ann. Inst. H. Poincaré Probab. Statist.* **41** (2005) 703–724. MR2144230

[17] P. Friz and N. Victoir. A note on the notion of geometric rough paths. *Probab. Theory Related Fields* **136** (2006) 395–416. MR2257130

[18] P. Friz and N. Victoir. A variation embedding theorem and applications. *J. Funct. Anal.* **239** (2006) 631–637. MR2261341

[19] P. Friz and N. Victoir. Large deviation principle for enhanced Gaussian processes. *Ann. Inst. H. Poincaré Probab. Statist.* **43** (2007) 775–785.

[20] P. K. Friz. Continuity of the Itô-map for Hölder rough paths with applications to the support theorem in Hölder norm. In *Probability and Partial Differential Equations in Modern Applied Mathematics* 117–135. *IMA Vol. Math. Appl.* **140**. Springer, New York, 2005. MR2202036

[21] M. Gubinelli. Controlling rough paths. *J. Funct. Anal.* **216** (2004) 86–140. MR2091358

[22] N. C. Jain and D. Monrad. Gaussian measures in B_p. *Ann. Probab.* **11** (1983) 46–57. MR0682800

[23] M. Ledoux. Isoperimetry and Gaussian analysis. In *Lectures on Probability Theory and Statistics (Saint-Flour, 1994)* 165–294. *Lecture Notes in Math.* **1648**. Springer, Berlin, 1996. MR1600888

[24] M. Ledoux and M. Talagrand. *Probability in Banach spaces*. *Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)]* **23**. Springer, Berlin, 1991. (Isoperimetry and processes.) MR1102015

[25] A. Lejay. An introduction to rough paths. In *Séminaire de Probabilités XXXVII* 1–59. *Lecture Notes in Math.* **1832**. Springer, Berlin, 2003. MR2053040

[26] T. Lyons. Differential equations driven by rough signals. *Rev. Mat. Iberoamericana* **14** (1998) 215–310. MR1654527

[27] T. Lyons and Z. Qian. *System Control and Rough Paths*. *Oxford Mathematical Monographs.* Oxford Univ. Press, 2002. MR2036784

[28] A. Millet and M. Sanz-Solé. Large deviations for rough paths of the fractional Brownian motion. *Ann. Inst. H. Poincaré Probab. Statist.* **42** (2006) 245–271. MR2199801

[29] A. Millet and M. Sanz-Sole. Approximation of rough paths of fractional Brownian motion. In *Seminar on Stochastic Analysis, Random Fields and Application* 275–303. *Progress in Probability* **59**. Birkhäuser, Basel, 2008. MR2401962

[30] J. Musielak and W. Orlicz. On generalized variations. I. *Studia Math.* **18** (1959) 11–41. MR0104771

[31] J. Musielak and Z. Semadeni. Some classes of Banach spaces depending on a parameter. *Studia Math.* **20** (1961) 271–284. MR0137985

[32] D. Nualart. *The Malliavin Calculus and Related Topics*. *Probability and Its Applications*. Springer, New York, 1995. MR1344217

[33] D. Revuz and M. Yor. *Continuous Martingales and Brownian Motion*, 3rd edition. *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]* **293**. Springer, Berlin, 1999.

[34] M. Schreiber. Fermeture en probabilité de certains sous-espaces d'un espace L^2. Application aux chaos de Wiener. *Z. Wahrsch. Verw. Gebiete* **14** (1969/70) 36–48. MR0413290

[35] D. W. Stroock and S. R. S. Varadhan. *Multidimensional Diffusion Processes*. *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]* **233**. Springer, Berlin, 1979. MR0532498

[36] N. Towghi. Multidimensional extension of L. C. Young's inequality. *JIPAM. J. Inequal. Pure Appl. Math.* **3** (2002), article 22 (electronic).

[37] L. C. Young. An inequality of Hölder type connected with Stieltjes integration. *Acta Math.* **67** (1936) 251–282. MR1555421