SMOOTHNESS OF THE DENSITY FOR SOLUTIONS TO GAUSSIAN ROUGH DIFFERENTIAL EQUATIONS

THOMAS CASS, MARTIN HAIRER, CHRISTIAN LITTERER, AND SAMY TINDEL

Abstract. We consider stochastic differential equations of the form $dY_t = V(Y_t)\,dX_t + V_0(Y_t)\,dt$ driven by a multi-dimensional Gaussian process. Under the assumption that the vector fields V_0 and $V = (V_1, \ldots, V_d)$ satisfy Hörmander's bracket condition, we demonstrate that Y_t admits a smooth density for any $t \in (0, T]$, provided the driving noise satisfies certain non-degeneracy assumptions. Our analysis relies on relies on an interplay of rough path theory, Malliavin calculus, and the theory of Gaussian processes. Our result applies to a broad range of examples including fractional Brownian motion with Hurst parameter $H > 1/4$, the Ornstein-Uhlenbeck process and the Brownian bridge returning after time T.

1. Introduction

Over the past decade our understanding of stochastic differential equations (SDEs) driven by Gaussian processes has evolved considerably. As a natural counterpart to this development, there is now considerable interest in investigating the probabilistic properties of solutions to these equations. Consider an SDE of the form

$$dY_t = V(Y_t)dX_t + V_0(Y_t)\,dt, \qquad Y(0) = y_0 \in \mathbb{R}^e, \tag{1.1}$$

driven by an $\mathbb{R}^d$-valued continuous Gaussian process X along C_b^∞-vector fields V_0 and $V = (V_1, \ldots, V_d)$ on $\mathbb{R}^e$. Once the existence and uniqueness of Y has been settled, it is natural to ask about the existence of a smooth density of Y_t for $t > 0$. In the context of diffusion processes, the theory is classical and goes back to Hörmander [22] for an analytical approach, and Malliavin [27] for a probabilistic approach.

For the case where X is fractional Brownian motion, this question was first addressed by Nualart and Hu [23], where the authors show the existence and smoothness of the density when the vector fields are elliptic, and the driving Gaussian noise is fractional Brownian motion (fBM) for $H > 1/2$. Further progress was achieved in [1] where, again for the regime $H > 1/2$, the density was shown to be smooth under Hörmander's celebrated bracket condition. Rougher noises are not directly amenable to the analysis put forward in these two papers. Additional ingredients have since gradually become available with the development of a broader theory of (Gaussian) rough paths (see [25], [8], [12]). The papers [7] and [6] used this technology to establish the existence of a density under fairly general assumptions on the Gaussian driving noises. These papers however fall

The research of Martin Hairer was supported by EPSRC grant EP/D071593/1, a Wolfson Research Merit award of the Royal Society, and a Philip Leverhulme Prize.

The research of Christian Litterer is supported by EPSRC grant EP/H000100/1.

The research of Christian Litterer was partially supported by a grant of the European Research Council (ERC grant nr. 258237).

S. Tindel is member of the BIGS (Biology, Genetics and Statistics) team at INRIA.

short of proving the smoothness of the density, because the proof demands far more quantitative estimates than were available at the time.

More recently, decisive progress was made on two aspects which obstructed the extension of this earlier work. First, the paper [5] established sharp tail estimates on the Jacobian of the flow $J^{\mathbf{X}}_{t\leftarrow 0}(y_0)$ driven by a wide class of (rough) Gaussian processes. The tail turns out to decay quickly enough to allow to conclude the finiteness of all moments for $J^{\mathbf{X}}_{t\leftarrow 0}(y_0)$. Second, [21] obtained a general, deterministic version of the key Norris lemma (see also [24] for some recent work in the context of fractional Brownian motion). The lemma of Norris first appeared in [29] and has been interpreted as a quantitative version of the Doob-Meyer decomposition. Roughly speaking, it ensures that there cannot be too many cancellations between martingale and bounded variation parts of the decomposition. The work [21] however shows that the same phenomenon arises in a purely deterministic setting, provided that the one-dimensional projections of the driving process are sufficiently and uniformly rough. This intuition is made precise through the notion of the "modulus of Hölder roughness". Together with an analysis of the higher order Malliavin derivatives of the flow of (1.1), also carried out in [21], these two results yield a Hörmander-type theorem for fractional Brownian motion if $H > 1/3$.

In this paper we aim to realise the broader potential of these developments by generalising the analysis to a wide class of Gaussian processes. This class includes fractional Brownian motion with Hurst parameter $H \in (\frac{1}{4}, \frac{1}{2}]$, the Ornstein-Uhlenbeck process, and the Brownian bridge. Instead of focusing on particular examples of processes, our approach aims to develop a general set of conditions on X under which Malliavin-Hörmander theory still works.

The probabilistic proof of Hörmander's theorem is intricate, and hard to summarise in a few lines, see [18] for a relatively short exposition. However, let us highlight some basic features of the method in order to see where our main contributions lie:

(i) At the centre of the proof of Hörmander's theorem is a quantitative estimate on the non-degeneracy of the Malliavin covariance matrix $C_T(\omega)$. Our effort in this direction consists in a direct and instructive approach, which reveals an additional structure of the problem. In particular, the conditional variance of the process plays an important role, which does not appear to have been noticed so far. More specifically, following [7] we study the Malliavin covariance matrix as a 2D Young integral against the covariance function $R(s,t)$. This provides the convenient representation:

$$v^T C_t(\omega) v = \int_{[0,t]\times[0,t]} f_s(v;\omega) f_r(v;\omega)\, dR(s,r),$$

for some γ-Hölder continuous $f(v;\omega)$, which avoids any detours via the fractional calculus that are specific to fBM. Compared to the setting of [6] we have to impose some additional assumptions on $R(s,t)$, but our more quantitative approach allows us in return to relax the zero-one law condition required in this paper.

(ii) An essential step in the proof is achieved when one obtains some lower bounds on $v^T C_t v$ in terms of $|f|_{\infty;[0,t]}$. Towards this aim we prove a novel interpolation inequality, which lies at the heart of this paper. It is explicit and also sharp in the sense that it collapses to a well-known inequality for the space $L^2([0,T])$ in the case of Brownian motion. Furthermore, this result should be important in other applications in the area, for example in establishing bounds on the density function (see [2] for a first step in this direction) or studying small-time asymptotics.

(iii) Hörmander's theorem also relies on an accurate analysis and control of the higher order Malliavin derivatives of the flow $J^{\mathbf{X}}_{t\leftarrow 0}(y_0)$. This turns out the be notationally cumbersome, but structurally quite similar to the technology already developed for fBm. For this step we therefore rely as much as possible on the analysis performed in [21]. The integrability results in [5] then play the first of two important roles in showing that the flow belongs to the Shigekawa-Sobolev space $\mathbb{D}^{\infty}(\mathbb{R}^e)$.

(iv) Finally, an induction argument that allows to transfer the bounds from the interpolation inequality to the higher order Lie brackets of the vector fields has to be set up. This induction requires another integrability estimate for $J^{\mathbf{X}}_{t\leftarrow 0}(y_0)$, plus a Norris type lemma allowing to bound a generic integrand A in terms of the resulting noisy integral $\int A\,dX$ in the rough path context. This is the content of our second main contribution, which can be seen as a generalisation of the Norris Lemma from [21] to a much wider range of regularities and Gaussian structures for the driving process X. Namely, we extend the result of [21] from p-rough paths with $p<3$ to general p under the same "modulus of Hölder roughness" assumption. It is interesting to note that the argument still only requires information about the roughness of the path itself and not its lift.

Let us further comment on the Gaussian assumptions allowing the derivation of the interpolation inequality briefly described in Step (ii) above. First, we need a standing assumption that regards the regularity of $R(s,t)$ (expressed in terms of its so called 2D ρ-variation, see [12]) and complementary Young regularity of X and its Cameron-Martin space. This is a standard assumption in the theory of Gaussian rough paths. The first part of the condition guarantees the existence of a natural lift of the process to a rough path. The complementary Young regularity in turn is necessary to perform Malliavin calculus, and allows us to obtain the integrability estimates for $J^{\mathbf{X}}_{t\leftarrow 0}(y_0)$ in [5].

In order to understand the assumptions on which our central interpolation inequality hinges, let us mention that it emerges from the need to prove lower bounds of the type:

$$\int_{[0,T]\times[0,T]} f_s f_t \, dR(s,t) \geq C\, |f|^{a}_{\gamma;[0,T]} \, |f|^{2-a}_{\infty;[0,T]}\,, \tag{1.2}$$

for some exponents γ and a, and all γ-Hölder continuous functions f. After viewing the integral in (1.2) along a sequence of discrete-time approximations to the integral, relation (1.2) relies on solving a sequence of finite dimensional partially constrained quadratic programming (QP) problems. These (QP) problem involve some matrices Q whose generic element can be written as $Q^{ij}=E[X^1_{t_i,t_{i+1}}X^1_{t_j,t_{j+1}}]$, where $X^1_{t_i,t_{i+1}}$ designates the increment $X^1_{t_{i+1}}-X^1_{t_i}$ of the first component of X. Interestingly enough, some positivity properties of Schur complements computed within the matrix Q play a prominent role in the resolution of the aforementioned (QP) problems. In order to guarantee these positivity properties, we shall make two non-degeneracy type assumptions on the conditional variance and covariance structure of our underlying process X^1 (see Conditions 2 and 3 below). This is obviously quite natural, since Schur complements are classically related to conditional variances in elementary Gaussian analysis. We also believe that our conditions essentially characterise the class of processes for which we can quantify the non-degeneracy of $C_T(\omega)$ in terms of the conditional variance of the process X.

The outline of the article is as follows. In Section 2, we give a short overview of the elements of the theory of rough paths required for our analysis. Section 3 then states our main result. In Section 4, we demonstrate how to verify the non-degeneracy assumptions required on the driving process in a number of concrete examples. The remainder of the article is devoted to the proofs. First, in Section 5, we state and prove our general version of Norris's lemma and we apply it to the class of Gaussian processes we have in mind. In Section 6, we then provide the proof of an interpolation inequality of the type (1.2). In Section 7 we obtain bounds on the derivatives of the solution with respect to its initial condition, as well as on its Malliavin derivative. Finally, we combine all of these ingredients in Section 8 in order to prove our main theorem.

2. Rough paths and Gaussian processes

In this section we introduce some basic notation concerning rough paths, following the exposition in [5]. In particular, we recall the conditions needed to ensure that a given Gaussian process has a natural rough path lift.

For $N \in \mathbb{N}$, recall that the truncated algebra $T^N(\mathbb{R}^d)$ is defined by $T^N(\mathbb{R}^d) = \bigoplus_{n=0}^N (\mathbb{R}^d)^{\otimes n}$, with the convention $(\mathbb{R}^d)^{\otimes 0} = \mathbb{R}$. $T^N(\mathbb{R}^d)$ is equipped with a straightforward vector space structure, plus an operation $\otimes$ defined by

$$\pi_n(g \otimes h) = \sum_{k=0}^N \pi_{n-k}(g) \otimes \pi_k(h), \qquad g, h \in T^N(\mathbb{R}^d),$$

where π_n designates the projection on the nth tensor level. Then $(T^N(\mathbb{R}^d), +, \otimes)$ is an associative algebra with unit element $\mathbf{1} \in (\mathbb{R}^d)^{\otimes 0}$.

At its most fundamental, we will study continuous $\mathbb{R}^d$-valued paths parameterised by time on a compact interval $[0, T]$; we denote the set of such functions by $C([0, T], \mathbb{R}^d)$. We write $x_{s,t} := x_t - x_s$ as a shorthand for the increments of a path. Using this notation we have

$$\|x\|_\infty := \sup_{t \in [0,T]} |x_t|, \qquad \|x\|_{p\text{-var};[0,T]} := \left(\sup_{D[0,T]=(t_j)} \sum_{j: t_j \in D[0,T]} \left| x_{t_j, t_{j+1}} \right|^p \right)^{1/p},$$

for $p \geq 1$, the uniform norm and the p-variation semi-norm respectively. We denote by $C^{p\text{-var}}([0, T], \mathbb{R}^d)$ the linear subspace of $C([0, T], \mathbb{R}^d)$ consisting of the continuous paths that have finite p-variation. Of interest will also be the set of γ-Hölder continuous function, denoted by $C^\gamma([0, T], \mathbb{R}^d)$, which consists of functions satisfying

$$\|x\|_{\gamma;[0,T]} := \sup_{0 \leq s < t \leq T} \frac{|x_{s,t}|}{|t-s|^\gamma} < \infty.$$

For $s < t$ and $n \geq 2$, consider the simplex $\Delta^n_{st} = \{(u_1, \ldots, u_n) \in [s, t]^n;\ u_1 < \cdots < u_n\}$, while the simplices over $[0, 1]$ will be denoted by Δ^n. A continuous map $\mathbf{x} : \Delta^2 \to T^N(\mathbb{R}^d)$ is called a multiplicative functional if for $s < u < t$ one has $\mathbf{x}_{s,t} = \mathbf{x}_{s,u} \otimes \mathbf{x}_{u,t}$. An important example arises from considering paths x with finite variation: for $0 < s < t$ we set

$$\mathbf{x}^n_{s,t} = \sum_{1 \leq i_1, \ldots, i_n \leq d} \left(\int_{\Delta^n_{st}} dx^{i_1} \cdots dx^{i_n} \right) e_{i_1} \otimes \cdots \otimes e_{i_n}, \tag{2.1}$$

where $\{e_1, \ldots, e_d\}$ denotes the canonical basis of $\mathbb{R}^d$, and then define the *signature* of x as

$$S_N(x) : \Delta^2 \to T^N(\mathbb{R}^d), \qquad (s, t) \mapsto S_N(x)_{s,t} := 1 + \sum_{n=1}^N \mathbf{x}^n_{s,t}.$$

$S_N(x)$ will be our typical example of multiplicative functional. Let us also add the following two remarks:

(i) A rough path (see Definition 2.1 below), as well as the signature of any smooth function, takes values in the strict subset $G^N(\mathbb{R}^d) \subset T^N(\mathbb{R}^d)$ given by the "group-like elements"

$$G^N(\mathbb{R}^d) = \exp^{\oplus}\left(L^N(\mathbb{R}^d)\right),$$

where $L^N(\mathbb{R}^d)$ is the linear span of all elements that can be written as a commutator of the type $a \otimes b - b \otimes a$ for two elements in $T^N(\mathbb{R}^d)$.

(ii) It is sometimes convenient to think of the indices $w = (i_1, \dots, i_n)$ in (2.1) as words based on the alphabet $\{1, \dots, d\}$. We shall then write $\mathbf{x}^w$ for the iterated integral $\int_{\Delta^n_{st}} dx^{i_1} \cdots dx^{i_n}$.

More generally, if $N \geq 1$ we can consider the set of such group-valued paths

$$\mathbf{x}_t = \left(1, \mathbf{x}_t^1, \dots, \mathbf{x}_t^N\right) \in G^N\left(\mathbb{R}^d\right).$$

Note that the group structure provides a natural notion of increment, namely $\mathbf{x}_{s,t} := \mathbf{x}_s^{-1} \otimes \mathbf{x}_t$, and we can describe the set of "norms" on $G^N\left(\mathbb{R}^d\right)$ which are homogeneous with respect to the natural scaling operation on the tensor algebra (see [12] for definitions and details). One such example is the Carnot-Caratheodory (CC) norm (see [12]), which we denote by $\|\cdot\|_{CC}$; the precise one used is mostly irrelevant in finite dimensions because they are all equivalent. The subset of these so-called homogeneous norms which are symmetric and sub-additive (again, see [12]) gives rise to genuine metrics on $G^N\left(\mathbb{R}^d\right)$, for example d_{CC} in the case of the CC norm. In turn these metrics give rise to a notion of homogenous p-variation metrics $d_{p\text{-var}}$ on the $G^N\left(\mathbb{R}^d\right)$-valued paths. Fixing attention to the CC norm, we will use the following homogenous p-variation and γ-Hölder variation semi-norms:

$$\|\mathbf{x}\|_{p\text{-var};[s,t]} = \max_{i=1,\dots,\lfloor p\rfloor} \left(\sup_{D[s,t]=(t_j)} \sum_{j:t_j \in D[s,t]} \left\|\mathbf{x}_{t_j,t_{j+1}}\right\|_{CC}^p \right)^{1/p}, \tag{2.2}$$

$$\|\mathbf{x}\|_{\gamma;[s,t]} = \sup_{(u,v)\in\Delta^2_{st}} \frac{\|\mathbf{x}_{u,v}\|_{CC}}{|v-u|^{\gamma}}.$$

Metrics on path spaces which are not homogenous will also feature. The most important will be the following

$$\mathcal{N}_{\mathbf{x},\gamma;[s,t]} := \sum_{i=1}^{N} \sup_{(u,v)\in\Delta^2_{st}} \frac{|\mathbf{x}^k_{u,v}|_{(\mathbb{R}^d)^{\otimes k}}}{|v-u|^{k\gamma}}, \tag{2.3}$$

which will be written simply as $\mathcal{N}_{\mathbf{x},\gamma}$ when $[s,t]$ is clear from the context.

Definition 2.1. *The space of weakly geometric p-rough paths (denoted $WG\Omega_p(\mathbb{R}^d)$) is the set of paths $\mathbf{x} : \Delta^2 \to G^{\lfloor p\rfloor}(\mathbb{R}^d)$ such that (2.2) is finite.*

We will also work with the space of geometric p-rough paths, which we denote by $G\Omega_p(\mathbb{R}^d)$, defined as the $d_{p\text{-var}}$-closure of

$$\left\{S_{\lfloor p\rfloor}(x) : x \in C^{1\text{-var}}\left([0,T], \mathbb{R}^d\right)\right\}.$$

Analogously, if $\gamma > 0$ and $N = [1/\gamma]$ we define $C^{0,\gamma}([0,T]; G^N(\mathbb{R}^d))$ to be the linear subspace of $G\Omega_N(\mathbb{R}^d)$ consisting of paths $\mathbf{x} : \Delta^2 \to G^N(\mathbb{R}^d)$ such that

$$\lim_{n\to\infty} \|\mathbf{x} - S_N(x_n)\|_{\gamma;[0,T]} = 0$$

for some sequence $(x_n)_{n=1}^{\infty} \subset C^{\infty}([0,T]; \mathbb{R}^d)$.

In the following we will consider RDEs driven by paths $\mathbf{x}$ in $WG\Omega_p(\mathbb{R}^d)$, along a collection of vector fields $V = (V_1, \dots, V_d)$ on $\mathbb{R}^e$, as well as a deterministic drift along V_0. From the point of view of existence and uniqueness results, the appropriate way to measure the regularity of the V_i's turns out to be the notion of Lipschitz-γ (short: Lip-γ) in the sense of Stein (see [12] and [26]). This notion provides a norm on the space of such vector fields (the Lip-γ norm), which we denote $|\cdot|_{\text{Lip-}\gamma}$. For the collection V of vector fields we will often make use of the shorthand

$$|V|_{\text{Lip-}\gamma} = \max_{i=1,\dots,d} |V_i|_{\text{Lip-}\gamma}\,,$$

and refer to the quantity $|V|_{\text{Lip-}\gamma}$ as the Lip-γ norm of V.

A theory of such Gaussian rough paths has been developed by a succession of authors (see [8, 14, 7, 11]) and we will mostly work within their framework. To be more precise, we will assume that $X_t = \left(X_t^1, \dots, X_t^d\right)$ is a continuous, centred (i.e. mean zero) Gaussian process with i.i.d. components on a complete probability space $(\Omega, \mathcal{F}, P)$. Let $\mathcal{W} = C([0,T], \mathbb{R}^d)$ and suppose that $(\mathcal{W}, \mathcal{H}, \mu)$ is the abstract Wiener space associated with X. The function $R : [0,T] \times [0,T] \to \mathbb{R}$ will denote the covariance function of any component of X, i.e.:

$$R(s,t) = E\left[X_s^1 X_t^1\right].$$

Following [14], we recall some basic assumptions on the covariance function of a Gaussian process which are sufficient to guarantee the existence of a natural lift of a Gaussian rough process to a rough path. We recall the notion of *rectangular increments* of R from [15], these are defined by

$$R\begin{pmatrix} s,t \\ u,v \end{pmatrix} := E\left[\left(X_t^1 - X_s^1\right)\left(X_v^1 - X_u^1\right)\right].$$

The existence of a lift for X is ensured by insisting on a sufficient rate of decay for the correlation of the increments. This is captured, in a very general way, by the following two-dimensional ρ-variation constraint on the covariance function.

Definition 2.2. *Given $1 \le \rho < 2$, we say that R has finite (two-dimensional) ρ-variation if*

$$V_\rho\left(R; [0,T] \times [0,T]\right)^\rho := \sup_{\substack{D=(t_i)\in\mathcal{D}([0,T]) \\ D'=\left(t'_j\right)\in\mathcal{D}([0,T])}} \sum_{i,j} \left| R\begin{pmatrix} t_i, t_{i+1} \\ t'_j, t'_{j+1} \end{pmatrix} \right|^\rho < \infty. \tag{2.4}$$

If a process has a covariance function with finite ρ-variation for $\rho \in [1,2)$ in the sense of Definition 2.2, [14, Thm 35] asserts that $(X_t)_{t\in[0,T]}$ lifts to a geometric p-rough path provided $p > 2\rho$. Moreover, there is a unique *natural lift* which is the limit, in the $d_{p\text{-var}}$-induced topology, of the canonical lift of piecewise linear approximations to X.

A related take on this notion is obtained by enlarging the set of partitions of $[0,T]^2$ over which the supremum is taken in (2.4). Recall from [15] that a *rectangular partition* of the square $[0,T]^2$ is a collection $\{A_i : i \in I\}$ of rectangles of the form $A_i = [s_i, t_i] \times [u_i, v_i]$, whose union equals $[0,T]^2$ and which have pairwise disjoint interiors. The collection of rectangular partitions is denoted $\mathcal{P}_{\text{rec}}\left([0,T]^2\right)$, and R is said to have *controlled ρ-variation* if

$$|R|^\rho_{\rho\text{-var};[0,T]^2} := \sup_{\substack{\{A_i:i\in I\}\in\mathcal{P}_{\text{rec}}\left([0,T]^2\right) \\ A_i=[s_i,t_i]\times[u_i,v_i]}} \sum_{i,j} \left| R\begin{pmatrix} s_i, t_i \\ u_i, v_i \end{pmatrix} \right|^\rho < \infty. \tag{2.5}$$

We obviously have $V_\rho(R; [0,T]^2) \le |R|_{\rho\text{-var};[0,T]^2}$, and it is shown in [15] that for every $\epsilon > 0$ there exists $c_{p,\epsilon}$ such that $|R|_{\rho\text{-var};[0,T]^2} \le c_{p,\epsilon V_{\rho+\epsilon}(R;[0,T]^2)}$. The main advantage of the quantity (2.5) compared to (2.4) is that the map

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

is a 2D control in the sense of [15].

Definition 2.3. *Given $1 \leq \rho < 2$, we say that R has finite (two-dimensional) Hölder-controlled ρ-variation if $V_\rho(R;[0,T]\times[0,T]) < \infty$, and if there exists $C > 0$ such that for all $0 \leq s \leq t \leq T$ we have*

$$V_\rho(R;[s,t]\times[s,t]) \leq C(t-s)^{1/\rho}. \tag{2.6}$$

Remark 2.4. *This is (essentially) without loss of generality compared to Definition 2.2. To see this, we note that if R also has controlled ρ-variation in the sense of (2.5), then we can introduce the deterministic time-change $\tau : [0,T] \to [0,T]$ given by $\tau = \sigma^{-1}$, where $\sigma : [0,T] \to [0,T]$ is the strictly increasing function defined by*

$$\sigma(t) := \frac{T\,|R|^{\rho}_{\rho\text{-}var;[0,t]^2}}{|R|^{\rho}_{\rho\text{-}var;[0,T]^2}}. \tag{2.7}$$

It is then easy to see that $\tilde{R}$, the covariance function of $\tilde{X} = X \circ \tau$, is Hölder-controlled in the sense of Definition 2.3.

Two important consequences of assuming that R has finite Hölder-controlled ρ-variation are: (i) $\mathbf{X}$ has $1/p-$Hölder sample paths for every $p > 2\rho$, and (ii) by using [12, Thm 15.33] we can deduce that

$$E\left[\exp\left(\eta\|\mathbf{X}\|^2_{1/p;[0,T]}\right)\right] < \infty \quad \text{for some} \quad \eta > 0, \tag{2.8}$$

i.e $\|\mathbf{X}\|^2_{1/p;[0,T]}$ has a Gaussian tail.

The mere existence of this lift is unfortunately not sufficient to apply the usual concepts of Malliavin calculus. In addition, it will be important to require a complementary (Young) regularity of the sample paths of X and the elements of its Cameron-Martin space. The following assumption captures both of these requirements.

Condition 1. *Let $(X_t)_{t\in[0,T]} = \left(X^1_t,\ldots,X^d_t\right)_{t\in[0,T]}$ be a Gaussian process with i.i.d. components. Suppose that the covariance function has finite Hölder-controlled ρ-variation for some $\rho \in [1,2)$. We will assume that X has a natural lift to a geometric p-rough path and that $\mathcal{H}$, the Cameron-Martin space associated with X, has Young-complementary regularity to X in the following sense: for some $q \geq 1$ satisfying $1/p + 1/q > 1$, we have the continuous embedding*

$$\mathcal{H} \hookrightarrow C^{q\text{-}var}\left([0,T],\mathbb{R}^d\right).$$

The following theorem appears in [14] as Proposition 17 (cf. also the recent note [15]); it shows how the assumption $V_\rho\big(R;[0,T]^2\big) < \infty$ allows us to embed $\mathcal{H}$ in the space of continuous paths with finite ρ variation. The result is stated in [14] for one-dimensional Gaussian processes, but the generalisation to arbitrary finite dimensions is straightforward.

Theorem 2.5 ([14])**.** *Let $(X_t)_{t\in[0,T]} = \left(X^1_t,\ldots,X^d_t\right)_{t\in[0,T]}$ be a mean-zero Gaussian process with independent and identically distributed components. Let R denote the covariance function of (any) one of the components. Then if R is of finite ρ-variation for some $\rho \in [1,2)$ we can embed $\mathcal{H}$ in the space $C^{\rho\text{-}var}\left([0,T],\mathbb{R}^d\right)$, in fact*

$$|h|_{\mathcal{H}} \geq \frac{|h|_{\rho\text{-}var;[0,T]}}{\sqrt{V_\rho(R;[0,T]\times[0,T])}}. \tag{2.9}$$

Remark 2.6 ([13])**.** *Writing $\mathcal{H}^H$ for the Cameron-Martin space of fBM for H in $(1/4, 1/2)$, the variation embedding in* [13] *gives the stronger result that*

$$\mathcal{H}^H \hookrightarrow C^{q\text{-}var}\left([0,T], \mathbb{R}^d\right) \quad \textit{for any} \quad q > (H + 1/2)^{-1}.$$

Theorem 2.5 and Remark 2.6 provide sufficient conditions for a process to satisfy the fundamental Condition 1, which we summarise in the following remark.

Remark 2.7. *As already observed, the requirement that R has finite 2D ρ-variation, for some $\rho \in [1, 2)$, implies both that X lifts to a geometric p-rough path for all $p > 2\rho$ and also that $\mathcal{H} \hookrightarrow C^{\rho\text{-}var}\left([0,T], \mathbb{R}^d\right)$ (Theorem 2.5). Complementary regularity of $\mathcal{H}$ in the above condition thus can be obtained by $\rho \in [1, 3/2)$, which covers for example BM, the OU process and the Brownian bridge (in each case with $\rho = 1$). When X is fBm, we know that X admit a lift to $G\Omega_p\left(\mathbb{R}^d\right)$ if $p > 1/H$, and Remark 2.6 therefore ensures the complementary regularity of X and $\mathcal{H}$ if $H > 1/4$.*

3. Statement of the main theorem

We will begin the section by laying out and providing motivation for the assumptions we impose on the driving Gaussian signal X. We will then end the section with a statement of the central theorem of this paper, which is a version of Hörmander's Theorem for Gaussian RDEs. First, we give some notation which will feature repeatedly.

Notation 1. *We define*

$$\mathcal{F}_{a,b} := \sigma\left(X_{v,v'} : a \leq v \leq v' \leq b\right)$$

to be the σ-algebra generated by the increments of X between times a and b.

The following condition aims to capture the non-degeneracy of X, it will feature prominently in the sequel.

Condition 2 (non-determinacy-type condition)**.** *Let $(X_t)_{t\in[0,T]}$ be a continuous Gaussian process. Suppose that the covariance function R of X has finite Hölder-controlled ρ-variation for some ρ in $[1, 2)$. We assume that there exists $\alpha > 0$ such that*

$$\inf_{0\leq s<t\leq T} \frac{1}{(t-s)^{\alpha}} \operatorname{Var}\left(X_{s,t} | \mathcal{F}_{0,s} \vee \mathcal{F}_{t,T}\right) > 0. \tag{3.1}$$

Whenever this condition is satisfied we will call α the ***index of non-determinism*** *if it is the smallest value of α for which (3.1) is true.*

Remark 3.1. *It is worthwhile making a number of comments. Firstly, notice that the conditional variance*

$$\operatorname{Var}\left(X_{s,t} | \mathcal{F}_{0,s} \vee \mathcal{F}_{t,T}\right)$$

is actually deterministic by Gaussian considerations. Then for any $[s,t] \subseteq [0,S] \subseteq [0,T]$, the law of total variance can be used to show that

$$\operatorname{Var}\left(X_{s,t} | \mathcal{F}_{0,s} \vee \mathcal{F}_{t,S}\right) \geq \operatorname{Var}\left(X_{s,t} | \mathcal{F}_{0,s} \vee \mathcal{F}_{t,T}\right).$$

It follows that if (3.1) holds on $[0,T]$, then it will also hold on any interval $[0,S] \subseteq [0,T]$ provided $S > 0$.

Note that Condition 2 implies the existence of $c > 0$ such that

$$\mathrm{Var}\left(X_{s,t}|\mathcal{F}_{0,s} \vee \mathcal{F}_{t,T}\right) \geq c\,(t-s)^{\alpha}.$$

This is reminiscent of (but not equivalent to) other notions of non-determinism which have been studied in the literature. For example, it should be compared to the similar notion introduced in [3], where it was exploited to show the existence of a smooth local time function (see also the subsequent work of Cuzick et al. [9] and [10]). In the present context, Condition 2 is also related to the following condition: for any f of finite p-variation over $[0,T]$

$$\int_0^T f_s dh_s = 0 \quad \forall h \in \mathcal{H} \qquad \Rightarrow \qquad f = 0 \quad \text{a.e. on } [0,T]. \tag{3.2}$$

This has been used in [6] to prove the existence of the density for Gaussian RDEs. In some sense, our Condition 2 is the quantitative version of (3.2). In this paper when we speak of a non-degenerate Gaussian process $(X_t)_{t\in[0,T]}$ we will mean the following:

Definition 3.2. *Let $(X_t)_{t\in[0,T]}$ be a continuous, real-valued Gaussian process. If $D = \{t_i : i = 0, 1, \ldots, n\}$ is any partition of $[0,T]$ let $(Q^D_{ij})_{1\leq i,j\leq n}$ denote the $n \times n$ matrix given by the covariance matrix of the increments of X along D, i.e.*

$$Q^D_{ij} = R\begin{pmatrix} t_{i-1}, t_i \\ t_{j-1}, t_j \end{pmatrix}. \tag{3.3}$$

We say that X is non-degenerate if Q^D is positive definite for every partition D of $[0,T]$.

Remark 3.3. *An obvious example of a 'degenerate' Gaussian process is a bridge processes which return to zero in $[0,T]$. This is plainly ruled out by an assumption of non-degeneracy.*

It is shown in [7] that non-degeneracy is implied by (3.2). Thus non-degeneracy is a weaker condition than (3.2). It also has the advantage of being formulated more tangibly in terms of the covariance matrix. The next lemma shows that Condition 2 also implies that the process is non-degenerate.

Lemma 3.4. *Let $(X_t)_{t\in[0,T]}$ be a continuous Gaussian process which satisfies Condition 2 then X is non-degenerate.*

Proof. Fix a partition D of $[0,T]$, and denote the covariance matrix along this partition by Q with entries as in (3.3). If Q is not positive definite, then for some non-zero vector $\lambda = (\lambda_1, \ldots, \lambda_n) \in \mathbb{R}^n$ we have

$$0 = \lambda^T Q \lambda = E\left[\left(\sum_{i=1}^n \lambda_i X_{t_{i-1},t_i}\right)^2\right]. \tag{3.4}$$

Suppose, without loss of generality, that $\lambda_j \neq 0$. Then from (3.4) we can deduce that

$$X_{t_{j-1},t_j} = \sum_{i\neq j}^n \frac{\lambda_i}{\lambda_j} X_{t_{i-1},t_i}$$

with probability one. This immediately implies that

$$\mathrm{Var}\left(X_{t_{j-1},t_j}|\mathcal{F}_{0,t_{j-1}} \vee \mathcal{F}_{t_j,T}\right) = 0,$$

which contradicts (3.1). □

A crucial step in the proof of the main theorem is to establish lower bounds on the eigenvalues of the Malliavin covariance matrix in order to obtain moment estimates for its inverse. In the setting we have adopted, it transpires that these eigenvalues can be bounded from below by some power of the 2D Young integral:

$$\int_{[0,T]^2} f_s f_t dR\left(s,t\right), \tag{3.5}$$

for some suitable (random) function $f \in C^{p\text{-var}}\left([0,T],\mathbb{R}^d\right)$. By considering the Riemann sum approximations to (3.5), the problem of finding a lower bound can be re-expressed in terms of solving a sequence of finite-dimensional constrained quadratic programming problems. By considering the dual of these problems, we can simplify the constraints which appear considerably; they become non-negativity constraints, which are much easier to handle. Thus the dual problem has an explicit solution subject to a dual feasibility condition. The following condition is what emerges as the limit of the dual feasibility conditions for the discrete approximations.

Condition 3. *Let $(X_t)_{t\in[0,T]}$ be a continuous, real-valued Gaussian process. We will assume that X has non-negative conditional covariance in that for every $[u,v]\subseteq[s,t]\subseteq[0,S]\subseteq[0,T]$ we have*

$$\operatorname{Cov}\left(X_{s,t},X_{u,v}|\mathcal{F}_{0,s}\vee\mathcal{F}_{t,S}\right)\geq 0. \tag{3.6}$$

In Section 6 we will prove a novel interpolation inequality. The significance of Condition 3 will become clearer when we work through the details of that section. For the moment we content ourselves with an outline. Firstly, for a finite partition D of the interval $[0,T]$ one can consider the discretisation of the process X_t conditioned on the increments in $D\cap([0,s]\cup[t,T])$. Let Q^D be the corresponding covariance matrix of the increments (see (3.3)). Then the conditional covariance $\operatorname{Cov}\left(X^D_{s,t},X^D_{u,v}|\mathcal{F}^D_{0,s}\vee\mathcal{F}^D_{t,T}\right)$ of the discretised process can be characterised in terms of a Schur complement Σ of the matrix Q^D. Utilising this relation, the condition

$$\operatorname{Cov}\left(X^D_{s,t},X^D_{u,v}|\mathcal{F}^D_{0,s}\vee\mathcal{F}^D_{t,T}\right)>0$$

is precisely what ensures that the row sums for Σ are non-negative. Conversely, if for any finite partition D all Schur complements of the matrix Q^D have non-negative row sums Condition 3 is satisfied. This relation motivates an alternative sufficient condition that implies Condition 3, which has the advantage that it may be more readily verified for a given Gaussian process. In order to state the condition, recall that an $n\times n$ real matrix Q is diagonally dominant if

$$Q_{ii}\geq\sum_{j\neq i}|Q_{ij}| \quad \text{for every } i\in\{1,2,\ldots,n\}. \tag{3.7}$$

Condition 4. *Let $(X_t)_{t\in[0,T]}$ be a continuous real-valued Gaussian process. For every $[0,S]\subseteq[0,T]$ we assume that X has diagonally dominant increments on $[0,S]$. By this we mean that for every partition $D=\{t_i : i=0,1,\ldots,n\}$ of $[0,S]$, the $n\times n$ matrix $\left(Q^D_{ij}\right)_{1\leq i,j\leq n}$ with entries*

$$Q^D_{ij}=E\left[X_{t_{i-1},t_i}X_{t_{j-1},t_j}\right]=R\begin{pmatrix} t_{i-1},t_i \\ t_{j-1},t_j\end{pmatrix}$$

is diagonally dominant.

Diagonal dominance is obviously in general a stronger assumption than requiring that a covariance matrix has positive row sums. Consequently, Condition 4 is particularly useful for negatively correlated processes, when diagonal dominance of the increments and positivity of row sums are the same. The condition can then be expressed succinctly as

$$E\left[X_{0,S}X_{s,t}\right] \geq 0 \qquad \forall\ [s,t] \subseteq [0,S] \subseteq [0,T].$$

In fact, it turns out that Condition 4 implies Condition 3. This is not obvious a priori, and ultimately depends on two nice structural features. The first is the observation from linear algebra that the property of diagonal dominance is preserved under taking Schur complements (see [32] for a proof of this). The second results from the interpretation of the Schur complement (in the setting of Gaussian vectors) as the covariance matrix of a certain conditional distribution. We will postpone the proof of this until Section 6 when these properties will be used extensively.

The final condition we will impose is classical, namely Hörmander's condition on the vector fields defining the RDE.

Condition 5 (Hörmander)**.** *We assume that*

$$span\left\{V_1,\ldots,V_d,[V_i,V_j],[V_i,[V_j,V_k]],\ldots\ :\ i,j,k,\ldots=0,1,\ldots,d\right\}|_{y_0} = \mathcal{T}_{y_0}\mathbb{R}^e \cong \mathbb{R}^e. \tag{3.8}$$

We are ready to formulate our main theorem.

Theorem 3.5. *Let $(X_t)_{t\in[0,T]} = \left(X_t^1,\ldots,X_t^d\right)_{t\in[0,T]}$ be a continuous Gaussian process, with i.i.d. components associated to the abstract Wiener space $(\mathcal{W},\mathcal{H},\mu)$. Assume that some (and hence every) component of X satisfies:*

1. *Condition 1, for some $\rho \in [1,2)$;*
2. *Condition 2, with index of non-determinancy $\alpha < 2/\rho$;*
3. *Condition 3, i.e. it has non-negative conditional covariance.*

Fix $p > 2\rho$, and let $\mathbf{X} \in G\Omega_p\left(\mathbb{R}^d\right)$ denote the canonical lift of X to a Gaussian rough path. Suppose $V = (V_1,\ldots,V_d)$ is a collection of C^∞-bounded vector fields on $\mathbb{R}^e$, and let $(Y_t)_{t\in[0,T]}$ be the solution to the RDE

$$dY_t = V\left(Y_t\right)d\mathbf{X}_t + V_0\left(Y_t\right)dt, \quad Y\left(0\right) = y_0.$$

Assume that the collection $(V_0,V_1,\ldots,V_d)$ satisfy Hörmander's condition, Condition 5, at the starting point y_0. Then random variable Y_t has a smooth density with respect to Lebesgue measure on $\mathbb{R}^e$ for every $t \in (0,T]$.

4. Examples

In this section we demonstrate how the conditions on X we introduced in the last section can be checked for a number of well known processes. We choose to focus on three particular examples: fractional Brownian motion (fBm) with Hurst parameter $H > 1/4$, the Ornstein Uhlenbeck (OU) process and the Brownian bridge. Together these encompass a broad range of Gaussian processes that one encounters in practice. Of course there are many more examples, but these should be checked on a case-by-case basis by analogy with our presentation for these three core examples. We first remark that Conditions 1 is straight forward to check in these cases (see e.g. [12] and [7]). We will therefore commence with a verification of the non-determinism condition i.e. Condition 2.

4.1. **Non-determinism-type condition.** Recall that the Cameron-Martin space $\mathcal{H}$ is defined to be the the completion of the linear space of functions of the form

$$\sum_{i=1}^{n} a_i R\left(t_i,\cdot\right), \quad a_i \in \mathbb{R} \text{ and } t_i \in [0,T],$$

with respect to the inner product

$$\left\langle \sum_{i=1}^{n} a_i R(t_i,\cdot), \sum_{j=1}^{m} b_j R(s_j,\cdot) \right\rangle_{\mathcal{H}} = \sum_{i=1}^{n}\sum_{j=1}^{m} a_i b_j R(t_i,s_j).$$

Some authors prefer instead to work with the set of step functions $\mathcal{E}$

$$\mathcal{E} = \left\{ \sum_{i=1}^{n} a_i 1_{[0,t_i]} : a_i \in \mathbb{R}, t_i \in [0,T] \right\},$$

equipped with the inner product

$$\left\langle 1_{[0,t]}, 1_{[0,s]} \right\rangle_{\tilde{\mathcal{H}}} = R(s,t).$$

If $\tilde{\mathcal{H}}$ denote the completion of ξ w.r.t $\langle\cdot,\cdot\rangle_{\tilde{\mathcal{H}}}$, then it is obvious that the linear map $\phi : \mathcal{E} \to \mathcal{H}$ defined by

$$\phi\left(1_{[0,t]}\right) = R(t,\cdot) \tag{4.1}$$

extends to an isometry between $\tilde{\mathcal{H}}$ and $\mathcal{H}$. We also recall that $\tilde{\mathcal{H}}$ is isometric to the Hilbert space $H^1(Z) \subseteq L^2(\Omega,\mathcal{F},P)$ which is defined to be the $|\cdot|_{L^2(\Omega)}$-closure of the set:

$$\left\{ \sum\nolimits_{i=1}^{n} a_i Z_{t_i} : a_i \in \mathbb{R},\ t_i \in [0,T],\ n \in \mathbb{N} \right\}.$$

In particular, we have that $\left|1_{[0,t]}\right|_{\tilde{\mathcal{H}}} = |Z_t|_{L^2(\Omega)}$. We will now prove that Condition 2 holds whenever it is the case that $\tilde{\mathcal{H}}$ embeds continuously in $L^q([0,T])$ for some $q \geq 1$. Hence, Condition 2 will simplify in many cases to showing that

$$|\tilde{h}|_{L^q[0,T]} \leq C|\tilde{h}|_{\tilde{\mathcal{H}}},$$

for some $C > 0$ and all $\tilde{h} \in \tilde{\mathcal{H}}$.

Lemma 4.1. *Suppose $(Z_t)_{t\in[0,T]}$ is a continuous real-valued Gaussian processes. Assume that for some $q \geq 1$ we have $\tilde{\mathcal{H}} \hookrightarrow L^q([0,T])$. Then Z satisfies Condition 2 with index of non-determinancy less than or equal to $2/q$; i.e.*

$$\inf_{0\leq s<t\leq T} \frac{1}{(t-s)^{2/q}} \operatorname{Var}\left(Z_{s,t}|\mathcal{F}_{0,s} \vee \mathcal{F}_{t,T}\right) > 0.$$

Proof. Fix $[s,t] \subseteq [0,T]$ and for brevity let $\mathcal{G}$ denote the σ-algebra $\mathcal{F}_{0,s} \vee \mathcal{F}_{t,T}$. Then, using the fact that $\operatorname{Var}(Z_{s,t}|\mathcal{G})$ is deterministic and positive, we have

$$\begin{aligned}
\operatorname{Var}(Z_{s,t}|\mathcal{G}) = \|\operatorname{Var}(Z_{s,t}|\mathcal{G})\|_{L^2(\Omega)} &= E\left[E\left[(Z_{s,t} - E[Z_{s,t}|\mathcal{G}])^2 |\mathcal{G}\right]^2\right]^{1/2} \\
&= E\left[(Z_{s,t} - E[Z_{s,t}|\mathcal{G}])^2\right] = \|Z_{s,t} - E[Z_{s,t}|\mathcal{G}]\|^2_{L^2(\Omega)} \\
&= \inf_{Y\in L^2(\Omega,\mathcal{G},P)} \|Z_{s,t} - Y\|^2_{L^2(\Omega)}.
\end{aligned}$$

We can therefore find sequence of random variables $(Y_n)_{n=1}^{\infty} \subset L^2(\Omega,\mathcal{G},P)$ such that

$$\|Z_{s,t} - Y_n\|^2_{L^2(\Omega)} = E\left[(Z_{s,t} - Y_n)^2\right] \downarrow \operatorname{Var}(Z_{s,t}|\mathcal{G}). \tag{4.2}$$

Moreover because $E[Z_{s,t}|\mathcal{G}]$ belongs to the closed subspace $H^1(Z)$, we can assume that Y_n has the form

$$Y_n = \sum_{i=1}^{k_n} a_i^n Z_{t_i^n, t_{i+1}^n}$$

for some sequence of real numbers

$$\{a_i^n : i = 1, \ldots, k_n\},$$

and a collection of subintervals

$$\{[t_i^n, t_{i+1}^n] : i = 1, \ldots, k_n\}$$

which satisfy $[t_i^n, t_{i+1}^n] \subseteq [0,s] \cup [s,T]$ for every $n \in \mathbb{N}$.

We now exhibit a lower bound for $\|Z_{s,t} - Y_n\|^2_{L^2(\Omega)}$ which is independent of n (and hence from (4.2) will apply also to $\mathrm{Var}(Z_{s,t}|\mathcal{G})$). Let us note that the isometry between the $H^1(Z)$ and $\tilde{\mathcal{H}}$ gives that

$$\|Z_{s,t} - Y_n\|^2_{L^2(\Omega)} = |\tilde{h}_n|^2_{\tilde{\mathcal{H}}}, \tag{4.3}$$

where

$$\tilde{h}_n(\cdot) := \sum_{i=1}^{k_n} a_i^n 1_{[t_i^n, t_{i+1}^n]}(\cdot) + 1_{[s,t]}(\cdot) \in \xi.$$

The embedding $\tilde{\mathcal{H}} \hookrightarrow L^q([0,T])$ then shows that

$$|\tilde{h}_n|^2_{\tilde{\mathcal{H}}} \geq c|\tilde{h}|^2_{L^q[0,T]} \geq c(t-s)^{2/q}.$$

The result follows immediately from this together with (4.2) and (4.3). □

Checking that $\tilde{\mathcal{H}}$ embeds continuously in a suitable $L^q([0,T])$ space is something which is readily done for our three examples. This is what the next lemma shows.

Lemma 4.2. *If $(Z_t)_{t\in[0,T]}$ is fBm with Hurst index $H \in (0,1/2)$ and $q \in [1,2)$ then $\tilde{\mathcal{H}} \hookrightarrow L^q([0,T])$. If $(Z_t)_{t\in[0,T]}$ is the (centred) Ornstein-Uhlenbeck process or the Brownian bridge (returning to zero after time T) then $\tilde{\mathcal{H}} \hookrightarrow L^2([0,T])$.*

Proof. The proof for each of the three examples has the same structure. We first identify an isometry K^* which maps $\tilde{\mathcal{H}}$ surjectively onto $L^2[0,T]$. (The operator K^* is of course different for the three examples). We then prove that the inverse $(K^*)^{-1}$ is a bounded linear operator when viewed as a map from $L^2[0,T]$ into $L^q[0,T]$. For fBm this is shown via the Hardy-Littlewood lemma (see [30]). For the OU process and the Brownian bridge it follows from a direct calculation on the operator K^*. Equipped with this fact, we can deduce that

$$|\tilde{h}|_{L^q[0,T]} = \left|(K^*)^{-1} K^*\tilde{h}\right|_{L^q[0,T]} \leq \left|(K^*)^{-1}\right|_{L^2\to L^q} \left|K^*\tilde{h}\right|_{L^2[0,T]} = \left|(K^*)^{-1}\right|_{L^2\to L^q} |\tilde{h}|_{\tilde{\mathcal{H}}},$$

which concludes the proof. □

As an immediate corollary of the last two lemmas we can conclude that the (centred) Ornstein-Uhlenbeck process and the Brownian bridge (returning to zero after time T) both satisfy Condition 2 with index of non-determinism no greater than unity. In the case of fBm $(Z_t^H)_{t\in[0,T]}$ the scaling properties of Z^H enable us to say more about the non-determinism index than can be obtained by an immediate application of Lemmas 4.1 and 4.2. To see this, note that for fixed $[s,t] \subseteq [0,T]$ we can introduce a new process

$$\tilde{Z}_u^H := (t-s)^{-H} Z^H_{u(t-s)}.$$

$\tilde{Z}$ defines another fBm, this time on the interval $[0, T(t-s)^{-1}] =: [0, \tilde{T}]$. Let $u = s(t-s)^{-1}, v = t(t-s)^{-1}$ and denote by $\tilde{\mathcal{F}}_{a,b}$ the σ-algebra generated by the increments of $\tilde{Z}$ in $[a, b]$. Scaling then allows us to deduce that

$$\mathrm{Var}(Z_{s,t}|\mathcal{F}_{0,s} \vee \mathcal{F}_{t,T}) = (t-s)^{2H} \mathrm{Var}(\tilde{Z}_{u,v}|\tilde{\mathcal{F}}_{0,u} \vee \tilde{\mathcal{F}}_{v,\tilde{T}}). \tag{4.4}$$

By construction $u - v = 1$. And since $\tilde{Z}$ is fBm it follows from Lemmas 4.1 and 4.2 that

$$\inf_{\substack{[u,v]\subseteq[0,\tilde{T}],\\ |u-v|=1}} \mathrm{Var}(\tilde{Z}_{u,v}|\tilde{\mathcal{F}}_{0,u} \vee \tilde{\mathcal{F}}_{v,\tilde{T}}) > 0. \tag{4.5}$$

It follows from (4.4) and (4.5) that Z^H satisfies Condition 2 with index of non-determinancy no greater than $2H$.

4.2. **Non-negativity of the conditional covariance.** We finally verify that our example processes also satisfy Condition 3. We first consider the special case of process with negatively correlated increments.

4.2.1. *Negatively correlated increments.* From our earlier discussion, it suffices to check that Condition 4 holds. In other words, that Q^D is diagonally dominant for every partition D. This amounts to showing that

$$E\left[Z_{t_{i-1},t_i}\, Z_{0,T}\right] \geq 0$$

for every $0 \leq t_{i-1} < t_i \leq T$. It is useful to have two general conditions on R which will guarantee that (i) the increments of Z are negatively correlated, and (ii) diagonal dominance is satisfied. Here is a simple characterisation of these properties:
Negatively correlated increments: If $i < j$, write

$$Q_{ij} = E\left[Z_{t_{i-1},t_i}\, Z_{t_{j-1},t_j}\right] = \int_{t_{i-1}}^{t_i} \int_{t_{j-1}}^{t_j} \partial^2_{ab} R(a,b)\, da\, db,$$

so that a sufficient condition for $Q_{ij} < 0$ is $\partial^2_{ab} R(a,b) \leq 0$ for $a < b$. This is trivially verified for fBm with $H < 1/2$. Note that $\partial^2_{ab} R(a,b)$ might have some nasty singularities on the diagonal, but the diagonal is avoided here.
Diagonal dominance: If we assume negatively correlated increments, then diagonal dominance is equivalent to $\sum_j Q_{ij} > 0$. Moreover, if we assume Z_0 is deterministic and Z is centred we get

$$\sum_j Q_{ij} = E\left[Z_{t_{i-1},t_i}\, Z_T\right] = \int_{t_{i-1}}^{t_i} \partial_a R(a,T)\, da,$$

so that a sufficient condition for $\sum_j Q_{ij} > 0$ is $\partial_a R(a,b) \geq 0$ for $a < b$. This is again trivially verified for fBm with $H < 1/2$.

Example 4.3. *In the case where $(Z_t)_{t\in[0,T]}$ is the Brownian bridge, which returns to zero at time $T' > T$ we have*

$$R(a,b) = a(T' - b), \quad \textit{for } a < b.$$

It is then immediate that $\partial^2_{ab} R(a,b) = -1 < 0$ and $\partial_a R(a,b) = T' - b > 0$. Similarly, for the centred Ornstein-Uhlenbeck process we have

$$R(a,b) = 2e^{-b}\sinh(a), \quad \textit{for } a < b.$$

From which it follows that $\partial^2_{ab} R(a,b) = -2e^{-b}\cosh(a) < 0$ and $\partial_a R(a,b) = 2e^{-b}\cosh(a) > 0$.

4.2.2. *Without negatively correlated increments.* In the three examples we were able to check Condition 3 by using the negative correlation of the increments and showing explicitly the diagonal dominance. In the case where the increments have positive or mixed correlation we may have to check the weaker condition, Condition 3, directly. An observation that might be useful in this regard is the following geometrical interpretation. Recall that we want to want to check that

$$\mathrm{Cov}\left(Z_{s,t}, Z_{u,v}|\mathcal{F}_{0,s} \vee \mathcal{F}_{t,T}\right) \geq 0.$$

For simplicity, let $X = Z_{s,t}$, $Y = Z_{u,v}$ and $\mathcal{G} = \mathcal{F}_{0,s} \vee \mathcal{F}_{t,T}$. The map $P_{\mathcal{G}} : Z \mapsto E\left[Z|\mathcal{G}\right]$ then defines a projection from the Hilbert space $L^2\left(\Omega, \mathcal{F}, P\right)$ onto the closed subspace $L^2\left(\Omega, \mathcal{G}, P\right)$, which gives the orthogonal decomposition

$$L^2\left(\Omega, \mathcal{F}, P\right) = L^2\left(\Omega, \mathcal{G}, P\right) \oplus L^2\left(\Omega, \mathcal{G}, P\right)^{\perp}.$$

A simple calculation then yields

$$\mathrm{Cov}\left(X, Y|\mathcal{G}\right) = E\left[\mathrm{Cov}\left(X, Y|\mathcal{G}\right)\right] = E\left[\left(I - P_{\mathcal{G}}\right) X \left(I - P_{\mathcal{G}}\right) Y\right] = \left\langle P_{\mathcal{G}}^{\perp} X, P_{\mathcal{G}}^{\perp} Y\right\rangle_{L^2(\Omega)},$$

where $P_{\mathcal{G}}^{\perp}$ is the projection onto $L^2\left(\Omega, \mathcal{G}, P\right)^{\perp}$. In other words, $\mathrm{Cov}\left(X, Y|\mathcal{G}\right) \geq 0$ if and only if $\cos\theta \geq 0$, where θ is the angle between the projections $P_{\mathcal{G}}^{\perp} X$ and $P_{\mathcal{G}}^{\perp} Y$ of, respectively, X and Y onto the orthogonal complement of $L^2\left(\Omega, \mathcal{G}, P\right)$.

5. A Norris-type Lemma

In this section we generalise a deterministic version of the Norris Lemma obtained in [21] for p rough paths with $1 < p < 3$ to general $p > 1$. It is interesting to note that the assumption on the driving noise we make is consistent with [21]. In particular, it still only depends on the roughness of the basic path and not the rough path lift.

5.1. **Norris's lemma.** To simplify the notation, we will assume that $T = 1$ in this subsection; all the work will therefore be done on the interval $[0, 1]$. Our Norris type lemma relies on the notion of controlled process, which we proceed to define now. Recall first the definition contained in [16] for second order rough paths: whenever $\mathbf{x} \in C^{0,\gamma}([0, 1]; G^N(\mathbb{R}^d))$ with $\gamma > 1/3$, the space $\mathcal{Q}_{\mathbf{x}}(\mathbb{R})$ of controlled processes is the set of functions $y \in C^{\gamma}([0, 1]; \mathbb{R})$ such that the increment y_{st} can be decomposed as

$$y_{st} = y_s^i x_{s,t}^i + r_{s,t},$$

where the remainder term r satisfies $|r_{s,t}| \leq c_y |t - s|^{2\gamma}$ and where we have used the summation over repeated indices convention. Notice that y has to be considered in fact as a vector $(y, y^1, \ldots, y^d)$.

In order to generalize this notion to lower values of γ, we shall index our controlled processes by words based on the alphabet $\{1, \ldots, d\}$. To this end, we need the following additional notations:

Notation 2. *Let $w = (i_1, \ldots, i_n)$ and $\bar{w} = (j_1, \ldots, j_m)$ be two words based on the alphabet $\{1, \ldots, d\}$. Then $|w| = n$ denotes the length of w, and $w\bar{w}$ stands for the concatenation $(i_1, \ldots, i_n, j_1, \ldots, j_m)$ of w and $\bar{w}$. For $L \geq 1$, $\mathcal{W}_L$ designates the set of words of length at most L.*

Let us now turn to the definition of controlled process based on a rough path:

Definition 5.1. *Let $\mathbf{x} \in C^{0,\gamma}([0, 1]; G^N(\mathbb{R}^d))$, with $\gamma > 0$, $N = [1/\gamma]$. A controlled path based on $\mathbf{x}$ is a family $(y^w)_{w \in \mathcal{W}_{N-1}}$ indexed by words of length at most $N - 1$, such that for any word $w \in \mathcal{W}_{N-2}$ we have:*

$$y_{s,t}^w = \sum_{\bar{w} \in \mathcal{W}_{N-1-|w|}} y_s^{w\bar{w}} \mathbf{x}_{s,t}^{\bar{w}} + r_{s,t}^w, \quad \textit{where} \quad |r_{s,t}^w| \leq c_y |t - s|^{(N-|w|)\gamma}. \tag{5.1}$$

In particular, for $w=\emptyset$, we get a decomposition for the increment $y_{s,t}$ of the form

$$y_{s,t} = \sum_{\bar w\in\mathcal{W}_{N-1}} y_s^{\bar w}\mathbf{x}_{s,t}^{\bar w} + r_{st}^y, \quad \textit{where} \quad |r_{s,t}^y| \le c_y|t-s|^{N\gamma}. \tag{5.2}$$

The set of controlled processes is denoted by $\mathcal{Q}_{\mathbf{x}}^\gamma$, and the norm on $\mathcal{Q}_{\mathbf{x}}^\gamma$ is given by

$$\|y\|_{\mathcal{Q}_{\mathbf{x}}^\gamma} = \sum_{w\in\mathcal{W}_{N-1}} \|y^w\|_\gamma.$$

We next recall the definition of θ-Hölder-roughness introduced in [21].

Definition 5.2. *Let $\theta\in(0,1)$. A path $x:[0,T]\to\mathbb{R}^d$ is called θ-Hölder rough if there exists a constant $c>0$ such that for every s in $[0,T]$, every ϵ in $(0,T/2]$, and every ϕ in $\mathbb{R}^d$ with $|\phi|=1$, there exists t in $[0,T]$ such that $\epsilon/2<|t-s|<\epsilon$ and*

$$|\langle\phi, x_{s,t}\rangle| > c\epsilon^\theta.$$

The largest such constant is called the modulus of θ-Hölder roughness, and is denoted $L_\theta(x)$.

A first rather straightforward consequence of this definition is that if a rough path $\mathbf{x}$ happens to be Hölder rough, then the *derivative processes* y^w in the decomposition (5.1) of a controlled path y is uniquely determined by y. This can be made quantitative in the following way:

Proposition 5.3. *Let $\mathbf{x}\in C^{0,\gamma}([0,1];G^N(\mathbb{R}^d))$, with $\gamma>0$ and $N=[1/\gamma]$. We also assume that x is a θ-Hölder rough path with $\theta<2\gamma$. Let y be a $\mathbb{R}$-valued controlled path defined as in Definition 5.1, and set $\mathcal{Y}_n(y)=\sup_{|w|=n}\|y^w\|_\infty$. Then there exists a constant M depending only on d such that the bound*

$$\mathcal{Y}_n(y) \le \frac{M\,[\mathcal{Y}_{n-1}(y)]^{1-\frac{\theta}{2\gamma}}}{L_\theta(x)}\left(1+\sum_{|\bar w|=2}^{N-n}\left(\|y^{w\bar w}\|_\infty\|\mathbf{x}^{\bar w}\|_{|\bar w|\gamma}\right)^{\frac{\theta}{|\bar w|\gamma}} + \sum_{|\bar w|=n-1}\|r^w\|_{(N-n+1)\gamma}^{\frac{\theta}{(N-n+1)\gamma}}\right) \tag{5.3}$$

holds for every controlled rough path $\mathcal{Q}_{\mathbf{x}}^\gamma$.

Proof. Start from decomposition (5.1), and recast it as

$$y_{s,t}^w = \sum_{j=1}^d y_s^{wj}\mathbf{x}_{s,t}^{(j)} + \sum_{2\le|\bar w|\le N-1-|w|} y_s^{w\bar w}\mathbf{x}_{s,t}^{\bar w} + r_{s,t}^w,$$

where we have set wj for the concatenation of the word w and the word (j) for notational sake. This identity easily yields

$$\sup_{|t-s|\le\varepsilon}\left|\sum_{j=1}^d y_s^{wj}\mathbf{x}_{s,t}^{(j)}\right| \le 2\|y^w\|_\infty + \sum_{2\le|\bar w|\le N-1-|w|}\|y^{w\bar w}\|_\infty\|\mathbf{x}^{\bar w}\|_{\gamma|\bar w|}\,\varepsilon^{|\bar w|\,\gamma} + \|r^w\|_{\gamma(N-|w|)}\,\varepsilon^{(N-|w|)\,\gamma} \tag{5.4}$$

Since x is θ-Hölder rough by assumption, for every $j\le d$, there exists $v=v(j)$ with $\varepsilon/2\le|v-s|\le\varepsilon$ such that

$$\left|\sum_{j=1}^d y_s^{wj}\mathbf{x}_{s,t}^{(j)}\right| > L_\theta(x)\,\varepsilon^\theta|\left(y_s^{w1},\ldots,y_s^{wd}\right)|. \tag{5.5}$$

Combining both (5.4) and (5.5) for all words w of length $n-1$, we thus obtain that

$$\mathcal{Y}_n(y) \le \frac{c}{L_\theta(x)} \Bigg[\mathcal{Y}_{n-1}(y)\, \varepsilon^{-\theta} + \sup_{|w|=n} \Bigg(\sum_{2\le|\bar{w}|\le N-1-|w|} \|y^{w\bar{w}}\|_\infty \|\mathbf{x}^{\bar{w}}\|_{\gamma|\bar{w}|}\, \varepsilon^{|\bar{w}|\,\gamma-\theta} + \|r^w\|_{\gamma(N-|w|)}\, \varepsilon^{(N-|w|)\,\gamma-\theta} \Bigg) \Bigg].$$

One can optimise the right hand side of the previous inequality over ε, by choosing ε such that the term $\mathcal{Y}_{n-1}(y)\,\varepsilon^{-\theta}$ is of the same order as the other ones. The patient reader might verify that our claim (5.3) is deduced from this elementary computation. □

Remark 5.4. *Definition 5.1 and Proposition 5.3 can be generalized straightforwardly to d-dimensional controlled processes. In particular, if y is a d dimensional path, decomposition* (5.2) *becomes*

$$y^i_{s,t} = \sum_{\bar{w}\in\mathcal{W}_{N-1}} y^{i,\bar{w}}_s \mathbf{x}^{\bar{w}}_{s,t} + r^{i,y}_{s,t}, \quad \text{where} \quad |r^{i,y}_{s,t}| \le c_y|t-s|^{N\gamma}, \tag{5.6}$$

for all $i=1,\dots,d$.

We now show how the integration of controlled processes fits into the general rough paths theory. For this we will use the non-homogeneous norm $\mathcal{N}_{\mathbf{x},\gamma} = \mathcal{N}_{\mathbf{x},\gamma,[0,1]}$ introduced in (2.3).

Proposition 5.5. *Let y be a d-dimensional controlled process, given as in Definition 5.1 and whose increments can be written as in* (5.6)*. Then* $(\mathbf{x},\mathbf{y})$ *is a geometrical rough path in $G^N(\mathbb{R}^{2d})$. In particular, for $(s,t)\in\Delta^2$, the integral $I_{st} \equiv \int_s^t y^i_s\, dx^i_s$ is well defined and admits the decomposition*

$$I_{s,t} = \sum_{j=1}^d \Big(y^j_s x^j_{s,t} + \sum_{\bar{w}\in\mathcal{W}_{N-1}} y^{\bar{w}}_s \mathbf{x}^{\bar{w}j}_{s,t} \Big) + r^I_{s,t}, \tag{5.7}$$

where $|r^I_{s,t}| \le \mathcal{N}_{\mathbf{x}} \|\mathbf{y}\|_\gamma |t-s|^{(N+1)\gamma}$.

Proof. Approximate x and y by smooth functions x^m, y^m, while preserving the controlled process structure (namely $y^m \in \mathcal{Q}_{\mathbf{x}^m}$). Then one can easily check that (x^m, y^m) admits a signature, and that $I^m_{s,t} \equiv \int_s^t y^{m,i}_s\, dx^{m,i}_s$ can be decomposed as (5.7). Limits can then be taken thanks to [17], which ends the proof.

□

The following theorem is a version of Norris' Lemma, and constitutes the main result of this section.

Theorem 5.6. *Let $\mathbf{x}$ be a geometric rough path of order $N\ge 1$ based on the $\mathbb{R}^d$-valued function x. We also assume that x is a θ-Hölder rough path with $2\gamma > \theta$. Let y be a $\mathbb{R}^d$-valued controlled path of the form given in Definition 5.1, $b\in C^\gamma([0,1])$, and set*

$$z_t = \sum_{i=1}^d \int_0^t y^i_s\, dx^i_s + \int_0^t b_s\, ds = I_{st} + \int_0^t b_s\, ds\,.$$

Then, there exist constants $r>0$ and $q>0$ such that, setting

$$\mathcal{R} = 1 + L_\theta(x)^{-1} + \mathcal{N}_{\mathbf{x},\gamma} + \|\mathbf{y}\|_{\mathcal{Q}^\gamma_{\mathbf{x}}} + \|b\|_{C^\gamma}\,, \tag{5.8}$$

one has the bound

$$\|y\|_\infty + \|b\|_\infty \le M\mathcal{R}^q\, \|z\|^r_\infty\,,$$

for a constant M depending only on T, d and y.

Proof. We shall divide this proof in several steps. In the following computations, κ will designate a certain power for $\mathcal{R}$ and M will stand for a multiplicative constant, whose exact values are irrelevant and can change from line to line.

Step 1: Bounds on y. Combining (5.7), the bound on r^I given at Proposition 5.5 and the definition of $\mathcal{R}$, we easily get the relation

$$\|z\|_\infty \le M\mathcal{R}^\kappa.$$

We now resort to relation (5.3) applied to the controlled path z and for $n = 1$, which means that $\mathcal{Y}_n(z) \asymp \|y\|_\infty$ and $\mathcal{Y}_{n-1}(z) \asymp \|z\|_\infty$. With the definition of $\mathcal{R}$ in mind, this yields the bound

$$\|y\|_\infty \le M\, \|z\|_\infty^{1-\frac{\theta}{2\gamma}} \mathcal{R}^\kappa, \tag{5.9}$$

which corresponds to our claim for y.

Along the same lines and thanks to relation (5.3) for $n > 1$, we iteratively get the bounds

$$\mathcal{Y}_n(y) \le M\, \|z\|_\infty^{(1-\frac{\theta}{2\gamma})^n} \mathcal{R}^\kappa, \tag{5.10}$$

which will be useful in order to complete the bound we have announced for b.

Step 2: Bounds on r^I and I. In order to get an appropriate bound on r, it is convenient to consider $\mathbf{x}$ as a rough path with Hölder regularity $\beta < \gamma$, still satisfying the inequality $2\beta > \theta$. Notice furthermore that $\mathcal{N}_{\mathbf{x},\beta} \le \mathcal{N}_{\mathbf{x},\gamma}$. Consider now $w \in \mathcal{W}_n$. According to (5.10) we have

$$\|y^w\|_\infty \le M\, \|z\|_\infty^{(1-\frac{\theta}{2\gamma})^n} \mathcal{R}^\kappa,$$

while $\|y^w\|_\gamma \le M\mathcal{R}$ by definition. Hence, invoking the inequality

$$\|y^w\|_\beta \le 2\|y^w\|_\gamma^{\frac{\beta}{\gamma}} \|y^w\|_\infty^{1-\frac{\beta}{\gamma}},$$

which follows immediately from the definition of the Hölder norm, we obtain the bound

$$\|y^w\|_\beta \le M\, \|z\|_\infty^{(1-\frac{\theta}{2\gamma})^n(1-\frac{\beta}{\gamma})} \mathcal{R}^\kappa,$$

which is valid for all $w \in \mathcal{W}_n$ and all $n \le N-1$. Summing up, we end up with the relation

$$\|\mathbf{y}\|_\beta \le M\, \|z\|_\infty^{(1-\frac{\theta}{2\gamma})^{N-1}(1-\frac{\beta}{\gamma})} \mathcal{R}^\kappa.$$

Now according to Proposition 5.5, we get $r^I_{s,t} \le \mathcal{N}_{\mathbf{x},\beta}\|\mathbf{y}\|_\beta |t-s|^{(N+1)\beta}$ and the above estimate yields

$$\|r^I\|_{(N+1)\beta} \le M\, \|z\|_\infty^{(1-\frac{\theta}{2\gamma})^{N-1}(1-\frac{\beta}{\gamma})} \mathcal{R}^\kappa.$$

Plugging this estimate into the decomposition (5.7) of I_{st} we end up with

$$\|I\|_\infty \le M\, \|z\|_\infty^{(1-\frac{\theta}{2\gamma})^{N-1}(1-\frac{\beta}{\gamma})} \mathcal{R}^\kappa. \tag{5.11}$$

Step 3: Bound on b. According to the bound (5.11) we have just obtained, we obviously have

$$\Big\| \int_0^\cdot b_s\, ds \Big\|_\infty \le M\, \|z\|_\infty^{(1-\frac{\theta}{2\gamma})^{N-1}(1-\frac{\beta}{\gamma})} \mathcal{R}^\kappa .$$

Once again we use an interpolation inequality to strengthen this bound. Indeed, we have (see [**?**, Lemma 6.14] for further details)

$$\|\partial_t f\|_\infty \le M \|f\|_\infty \max\Big(\frac{1}{T}, \|f\|_\infty^{-\frac{1}{\gamma+1}} \|\partial_t f\|_\gamma^{\frac{1}{\gamma+1}}\Big),$$

and applying this inequality to $f_t = \int_0^t b_s\,ds$, it follows that

$$\|b\|_\infty \leq M\,\|z\|_\infty^{(1-\frac{\theta}{2\gamma})^{N-1}(1-\frac{\beta}{\gamma})(\frac{\gamma}{\gamma+1})}\mathcal{R}^\kappa\,. \tag{5.12}$$

Gathering the bounds (5.9) and (5.12), our proof is now complete. □

5.2. **Small-ball estimates for $\mathbf{L}_\theta\left(\mathbf{X}\right)$.** We now take X to be a Gaussian process satisfying Condition 2. As the reader might have noticed, equation (5.8) above involves the random variable $L_\theta\left(X\right)^{-1}$, for which we will need some tail estimates. The non-determinism condition naturally gives rise to such estimates as the following lemma makes clear.

Lemma 5.7. *Suppose $(X_t)_{t\in[0,T]}$ is a zero-mean, $\mathbb{R}^d$-valued, continuous Gaussian process with i.i.d. components, with each component having a continuous covariance function R. Suppose that one (and hence every) component of X satisfies Condition 2. Let $\alpha_0 > 0$ be the index of non-determinism for X and suppose $\alpha \geq \alpha_0$. Then there exist positive and finite constants C_1 and C_2 such that for any interval $I_\delta \subseteq [0,T]$ of length δ and $0 < x < 1$ we have*

$$P\left(\inf_{|\phi|=1}\sup_{s,t\in I_\delta}|\langle\phi, X_{s,t}\rangle| \leq x\right) \leq C_1\exp\left(-C_2\delta x^{-2/\alpha}\right). \tag{5.13}$$

Proof. The proof is similar to Theorem 2.1 of Monrad and Rootzen [28]; we need to adapt it because our non-determinism condition is different.

We start by introducing two simplifications. Firstly, for any ϕ in $\mathbb{R}^d$ with $|\phi| = 1$ we have

$$\left(\langle\phi, X_t\rangle\right)_{t\in[0,T]} \overset{D}{=} \left(X_t^1\right)_{t\in[0,T]}\,, \tag{5.14}$$

which implies that

$$P\left(\sup_{s,t\in I_\delta}|\langle\phi, X_{s,t}\rangle| \leq x\right) = P\left(\sup_{s,t\in I_\delta}\left|X_{s,t}^1\right| \leq x\right). \tag{5.15}$$

We will prove that the this probability is bounded above by

$$\exp\left(-c\delta x^{2/\alpha}\right),$$

for a positive real constant c, which will not depend on T, δ or x. The inequality (5.13) will then follow by a well-known compactness argument (see [21] and [29]). The second simplification is to assume that $\delta = 1$. We can justify this by working with the scaled process

$$\tilde{X}_t = \delta^{\alpha/2}X_{t/\delta},$$

which is still Gaussian process only now parametrised on the interval $[0,\tilde{T}] := [0, T\delta]$. Furthermore, the scaled process also satisfies Condition 2 since

$$\begin{aligned}\operatorname{Var}\left(\tilde{X}_{s,t}|\tilde{\mathcal{F}}_{0,s}\vee\tilde{\mathcal{F}}_{t,\tilde{T}}\right) &= \delta^\alpha\operatorname{Var}\left(X_{s/\delta,t/\delta}|\mathcal{F}_{0,s/\delta}\vee\mathcal{F}_{t/\delta,T}\right)\\ &\geq c\delta^\alpha\left(\frac{t-s}{\delta}\right)^\alpha = c\left(t-s\right)^\alpha.\end{aligned}$$

Thus, if we can prove the result for intervals of length 1, we can deduce the bound on (5.15) we want from the identity

$$P\left(\sup_{s,t\in I_\delta}\left|X_{s,t}^1\right| \leq x\right) = P\left(\sup_{s,t\in I_1}\left|\tilde{X}_{s,t}^1\right| \leq \frac{x}{\delta^{\alpha/2}}\right).$$

To conclude the proof, we begin by defining the natural number $n := \lfloor x^{-2/\alpha} \rfloor \geq 1$ and the dissection $D(I) = \{t_i : i = 0, 1, \ldots, n+1\}$ of $I = I_1$, given by

$$t_i = \inf I + i x^{2/\alpha}, \quad i = 0, 1, \ldots, n$$
$$t_{n+1} = \inf I + 1 = \sup I.$$

Then it is trivial to see that

$$P\left(\sup_{s,t \in I} \left|X^1_{s,t}\right| \leq x\right) \leq P\left(\max_{i=1,2,\ldots,n} \left|X^1_{t_{i-1},t_i}\right| \leq x\right). \tag{5.16}$$

To estimate (5.16) we successively condition on the components of

$$(X^1_{t_0,t_1}, \ldots, X^1_{t_{n-1},t_n}).$$

More precisely, the distribution of $X^1_{t_{n-1},t_n}$ conditional on $(X^1_{t_0,t_1}, \ldots, X^1_{t_{n-2},t_{n-1}})$ is Gaussian with a variance σ^2. Condition 2 ensures that σ^2 is bounded below by cx^2. When Z is a Gaussian random variable with fixed variance, $P(|Z| \leq x)$ will be maximised when the mean is zero. We therefore obtain the following upper bound

$$P\left(\sup_{s,t \in I} \left|X^1_{s,t}\right| \leq x\right) \leq \left(\int_{-x/\sigma}^{x/\sigma} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}y^2\right) dy\right)^n.$$

Using $x/\sigma \leq \sqrt{c}$, we can finally deduce that

$$P\left(\sup_{s,t \in I} \left|X^1_{s,t}\right| \leq x\right) \leq \exp(-Cn) \leq \exp\left(-\frac{Cx^{-2/\alpha}}{2}\right),$$

where $C := \log\left[2\Phi\left(\sqrt{c}\right) - 1\right]^{-1} \in (0, \infty)$. □

Corollary 5.8. *Suppose $(X_t)_{t \in [0,T]}$ is a zero-mean, $\mathbb{R}^d$-valued, continuous Gaussian process with i.i.d. components satisfying the conditions of Lemma 5.7. Then for every $\theta > \alpha/2$, the path $(X_t)_{t \in [0,T]}$ is almost surely θ-Hölder rough. Furthermore, for $0 < x < 1$ there exist positive finite constants C_1 and C_2 such that the modulus of θ-Hölder roughness, $L_\theta(X)$, satisfies*

$$P(L_\theta(X) < x) \leq C_1 \exp\left(-C_2 x^{-2/\alpha}\right).$$

In particular, under these assumptions we have that $L_\theta(X)^{-1}$ is in $\cap_{p>0} L^p(\Omega)$.

Proof. The argument of [21] applies in exactly the same way to show that $L_\theta(X)$ is bounded below by

$$\frac{1}{2 \cdot 8^\theta} D_\theta(X),$$

where

$$D_\theta(X) := \inf_{\|\phi\|=1} \inf_{n \geq 1} \inf_{k \leq 2^n} \sup_{s,t \in I_{k,n}} \frac{|\langle \phi, X_{s,t} \rangle|}{(2^{-n}T)^\theta}$$

and $I_{k,n} := [(k-1)2^{-n}T, k2^{-n}T]$. We can deduce that for any $x \in (0,1)$

$$P(D_\theta(X) < x) \leq \sum_{n=1}^{\infty} \sum_{k=1}^{2^n} P\left(\inf_{\|\phi\|=1} \sup_{s,t \in I_{k,n}} \frac{|\langle \phi, X_{s,t} \rangle|}{(2^{-n}T)^\theta} < x\right),$$

whereupon we can apply Lemma 5.7 to yield

$$P\left(D_\theta\left(X\right)<x\right)\leq c_1\sum_{n=1}^{\infty}2^n\exp\left(-c_2 2^{-n(1-2\theta/\alpha)}T^{-2\theta/\alpha}x^{-2/\alpha}\right).$$

By exploiting the fact that $\theta>\alpha/2$, we can then find positive constants c_3 and c_4 such that

$$\begin{aligned}P\left(D_\theta\left(X\right)<x\right)&\leq c_3\sum_{n=1}^{\infty}\exp\left(-c_4 n x^{-2/\alpha}\right)=c_3\frac{\exp\left(-c_4x^{-2/\alpha}\right)}{1-\exp\left(-c_4x^{-2/\alpha}\right)}\\&\leq c_5\exp\left(-c_4x^{-2/\alpha}\right),\end{aligned}$$

which concludes the proof. □

6. An interpolation inequality

Under the standing assumptions on the Gaussian process, the *Malliavin covariance matrix* of the random variable $U_{t\leftarrow 0}^{\mathbf{X}}\left(y_0\right)\equiv Y_t$ can be represented as a $2D$ Young integral (see [7])

$$C_t=\sum_{i=1}^{d}\int_{[0,t]^2}J_{t\leftarrow s}^{\mathbf{X}}\left(y_0\right)V_i\left(Y_s\right)\otimes J_{t\leftarrow s'}^{\mathbf{X}}\left(y_0\right)V_i\left(Y_{s'}\right)dR\left(s,s'\right). \tag{6.1}$$

In practice, showing the smoothness of the density boils down to getting integrability estimates on the inverse of $\inf_{\|v\|=1}v^TC_Tv$, the smallest eigenvalue of C_T. For this reason we will be interested in

$$v^TC_Tv=\sum_{i=1}^{d}\int_{[0,T]^2}\left\langle v,J_{t\leftarrow s}^{\mathbf{X}}\left(y_0\right)V_i\left(Y_s\right)\right\rangle\left\langle v,J_{t\leftarrow s'}^{\mathbf{X}}\left(y_0\right)V_i\left(Y_{s'}\right)\right\rangle dR\left(s,s'\right).$$

We will return to study the properties of C_T more extensively in Section . For the moment, we look to generalise this perspective somewhat. Suppose $f:[0,T]\to\mathbb{R}$ is some (deterministic) real-valued Hölder-continuous function, where γ is Young-complementary to ρ, 2D-variation regularity of R. Our aim in this section is elaborate on the non-degeneracy of the 2D Young integral

$$\int_{[0,T]}f_sf_tdR\left(s,t\right). \tag{6.2}$$

More precisely, what we want is to use Conditions 2 and 3 to give a quantitative version of the non-degeneracy statement:

$$\int_{[0,T]}f_sf_tdR\left(s,t\right)=0\Rightarrow f\equiv 0. \tag{6.3}$$

To give an idea of the type of estimate we might aim for, consider the case where $R\equiv R^{BM}$ is the covariance function of Brownian motion. The 2D Young integral (6.2) then collapses to the square of the L^2-norm of f :

$$\left|\int_{[0,T]}f_sf_tdR^{BM}\left(s,t\right)\right|=\left|f\right|_{L^2[0,T]}^2, \tag{6.4}$$

and the interpolation inequality (Lemma A3 of [20]) gives

$$\left\|f\right\|_{\infty;[0,T]}\leq 2\max\left(T^{-1/2}\left|f\right|_{L^2[0,T]},\left|f\right|_{L^2[0,T]}^{2\gamma/(2\gamma+1)}\left\|f\right\|_{\gamma\text{-Höl};[0,T]}^{1/(2\gamma+1)}\right). \tag{6.5}$$

Therefore, in the setting of Brownian motion at least, (6.5) and (6.4) quantifies (6.3). The problem is that the proof of (6.5) relies heavily properties of the L^2-norm, in particular we use the fact that

$$\text{if } f(u) \geq c > 0 \text{ for all } u \in [s,t] \text{ then } |f|_{L^2[s,t]} \geq c(t-s)^{1/2}.$$

We cannot expect for this to naively generalise to inner products resulting from other covariance functions. We therefore have to re-examine the proof of the inequality (6.5) with this generalisation in mind.

It is easier to first consider a discrete version of the problem. Suppose D is some (finite) partition of $[0,T]$. Then the Riemann sum approximation to (6.2) along D can be written as

$$f(D)^T Qf(D),$$

where Q is the the matrix (3.3) and $f(D)$ the vector with entries given by the values of f at the points in the partition. The next sequence of results is aimed at addressing the following question:

Problem 6.1. *Suppose $|f|_{\infty;[s,t]} \geq 1$ for some interval $[s,t] \subseteq [0,T]$. Can we find a positive lower bound $f(D)^T Qf(D)$ which holds uniformly over some sequence of partitions whose mesh tends to zero?*

The next lemma is the first step towards securing an answer.

Lemma 6.2. *Let $(Q_{ij})_{i,j\in\{1,2,\dots,n\}}$ be a real $n \times n$ positive definite matrix and k be an integer in $\{1,\dots,n\}$. Suppose Q has the block decomposition*

$$Q = \begin{pmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{pmatrix}, \quad \textit{with} \quad Q_{11} \in \mathbb{R}^{k,k}, Q_{12} \in \mathbb{R}^{k,n-k}, Q_{21} \in \mathbb{R}^{n-k,k}, Q_{22} \in \mathbb{R}^{n-k,n-k}.$$

Let S denote the Schur complement of Q_{11} in Q, i.e. S is the $(n-k)\times(n-k)$ matrix given by

$$S = Q_{22} - Q_{21}Q_{11}^{-1}Q_{12}.$$

Assume S has non-negative row sums:

$$\sum_{j=1}^{n-k} S_{ij} \geq 0 \textit{ for all } i \in \{1,\dots,n-k\}, \tag{6.6}$$

and $b > 0$ is such that $\mathbf{b} = (b,\dots,b)$ is in $\mathbb{R}^{n-k}$. Then the infimum of the quadratic form $x^T Qx$ over the subset

$$\mathcal{C} = \{(x_1,\dots,x_n) : (x_{k+1},\dots,x_n) \geq \mathbf{b}\}$$

is given by

$$\inf_{x\in\mathcal{C}} x^T Qx = \mathbf{b}^T S\mathbf{b} = b^2 \sum_{i,j=1}^{n-k} S_{ij}.$$

Proof. Without loss of generality we may assume that $b = 1$. We can then reformulate the statement as describing the infimal value for the following constrained quadratic programming problem:

$$\min x^T Qx \qquad \text{subject to} \qquad Ax \geq (\mathbf{0}_{\mathbb{R}^k}, \mathbf{1}_{\mathbb{R}^{n-k}})$$

where $\mathbf{0}_{\mathbb{R}^k} := (0,\dots,0) \in \mathbb{R}^k$, $\mathbf{1}_{\mathbb{R}^{n-k}} := (1,\dots,1) \in \mathbb{R}^{n-k}$ and A the $n \times n$ matrix which is defined (in the standard basis) by

$$Ax = A(x_1,\dots,x_n)^T = (\mathbf{0}_{\mathbb{R}^k}, x_{k+1},\dots,x_n)^T.$$

The Lagrangian function of this quadratic programming problem (see e.g. [4] page 215) is given by

$$L(x,\lambda) = x^T Qx + \lambda^T \left(-Ax + (\mathbf{0}_{\mathbb{R}^k}, \mathbf{1}_{\mathbb{R}^{n-k}})\right).$$

Solving for

$$\nabla_x L(x,\lambda) = 2Qx - A^T\lambda = 0$$

and using the strict convexity of the function we deduce that $x^* = 1/2Q^{-1}A^T\lambda$ is the minimiser of L. Hence, the (Lagrangian) dual function $g(\lambda) := \inf_x L(x,\lambda)$ is given by

$$g(\lambda) = -\frac{1}{4}\lambda^T AQ^{-1}A^T\lambda + \lambda^T (\mathbf{0}_{\mathbb{R}^k}, \mathbf{1}_{\mathbb{R}^{n-k}})$$

and the dual problem consists of

$$\max g(\lambda) \qquad \text{subject to} \qquad \lambda \geq 0.$$

As Q^{-1} is positive definite the function g is strictly concave and the local maximum $\lambda^* = 2Qb > 0$ that is obtained by solving $\nabla_\lambda g(\lambda) = 0$ with

$$\nabla_\lambda g(\lambda) = -\frac{1}{2}AQ^{-1}A^T\lambda + \lambda^T (\mathbf{0}_{\mathbb{R}^k}, \mathbf{1}_{\mathbb{R}^{n-k}}) \tag{6.7}$$

is also the unique global maximum. Writing Q^{-1} in block form

$$Q^{-1} = \begin{pmatrix} (Q^{-1})_{11} & (Q^{-1})_{12} \\ (Q^{-1})_{21} & (Q^{-1})_{22} \end{pmatrix},$$

and using the definition of A it is easy to see that the vector

$$\lambda^* := \left(\mathbf{0}_{\mathbb{R}^k}, 2\left(Q^{-1}\right)_{22}^{-1}\mathbf{1}_{\mathbb{R}^{n-k}}\right) \tag{6.8}$$

solves (6.7). We need to check that this vector is feasible for the dual problem, since then strong duality holds (see e.g. [4] pages 226-227) and the optimal values for the dual and primal problems coincide.

In order to check that (6.8) is feasible we need to show $\lambda^* \geq 0$. To do this, we first remark that a straight-forward calculation gives the inverse of the sub-block $\left(Q^{-1}\right)_{22}$ as the Schur complement of Q_{11} in Q; that is

$$\left(Q^{-1}\right)_{22}^{-1} = S = Q_{22} - Q_{21}Q_{11}^{-1}Q_{12}.$$

Condition (6.6) gives immediately that for every $i \in \{k+1,\ldots,n\}$

$$\lambda_i^* = 2\sum_{j=1}^{n-k} S_{(i-k)j} \geq 0$$

and hence $\lambda^* \geq 0$. It now follows from strong duality that we have

$$\inf_{x\in\mathcal{C}} x^T Qx = \min_{\lambda\in\mathbb{R}^n_+} g(\lambda) = \mathbf{1}^T_{\mathbb{R}^{n-k}} S\mathbf{1}_{\mathbb{R}^{n-k}},$$

as required. □

Suppose now that Q arises as the covariance matrix of the increments of a Gaussian process along some partition. In light of the previous lemma, we need to know when the Schur complement of some sub-block of Q will satisfy condition (6.6). In the context of Gaussian vectors, these Schur complements have a convenient interpretation; they are the covariance matrices which result from partially conditioning on some of the components. This identification motivates the positive conditional covariance condition (Condition 3).

In order to present the proof of the interpolation inequality as transparently as possible, we first gather together some relevant technical comments. To start with, suppose we have two sets of real numbers

$$D = \{t_i : i = 0, 1, \ldots, n\} \subset \tilde{D} = \left\{\tilde{t}_i : i = 0, 1, \ldots, \tilde{n}\right\} \subseteq [0, T]$$

ordered in such a way that $0 \leq t_0 < t_1 < \ldots < t_n \leq T$, and likewise for $\tilde{D}$. Suppose s and t be real numbers with $s < t$ and let Z be a continuous Gaussian process. We need to consider how the variance of the increment $Z_{s,t}$ changes when we condition on

$$\mathcal{F}^D := \sigma\left(Z_{t_{i-1},t_i} : i = 1, \ldots, n\right),$$

compared to conditioning the larger σ-algebra

$$\mathcal{F}^{\tilde{D}} := \sigma\left(Z_{\tilde{t}_{i-1},\tilde{t}_i} : i = 1, \ldots, \tilde{n}\right).$$

To simplify the notation a little we introduce

$$\mathcal{G} = \sigma\left(Z_{\tilde{t}_{i-1},\tilde{t}_i} : \left\{\tilde{t}_{i-1}, \tilde{t}_i\right\} \cap \tilde{D} \setminus D \neq \emptyset\right),$$

so that

$$\mathcal{F}^{\tilde{D}} = \mathcal{F}^D \vee \mathcal{G}.$$

Because

$$\left(Z_{s,t}, Z_{t_0,t_1}, \ldots, Z_{t_{\tilde{n}-1},t_{\tilde{n}}}\right) \in \mathbb{R}^{\tilde{n}+1} \tag{6.9}$$

is Gaussian, the joint distribution of $Z_{s,t}$ and the vector (6.9) conditional on $\mathcal{F}^D$ (or indeed $\mathcal{F}^{\tilde{D}}$) is once again Gaussian, with a random mean but a deterministiccovariance matrix. A simple calculation together with the *law of total variance* gives that

$$\begin{aligned}\operatorname{Var}\left(Z_{s,t}|\mathcal{F}^D\right) &= E\left[\operatorname{Var}\left(Z_{s,t}|\mathcal{F}^D \vee \mathcal{G}\right)\right] + \operatorname{Var}\left(E\left[Z_{s,t}|\mathcal{F}^D \vee \mathcal{G}\right]\right) \\ &\geq E\left[\operatorname{Var}\left(Z_{s,t}|\mathcal{F}^D \vee \mathcal{G}\right)\right] = \operatorname{Var}\left(Z_{s,t}|\mathcal{F}^{\tilde{D}}\right),\end{aligned}$$

which is the comparison we sought. We condense these observations into the following lemma.

Lemma 6.3. *Let $(Z_t)_{t\in[0,T]}$ be a Gaussian process, and suppose that D and $\tilde{D}$ are two partitions of $[0, T]$ with $D \subseteq \tilde{D}$. Then for any $[s, t] \subseteq [0, T]$ we have*

$$\operatorname{Var}\left(Z_{s,t}|\mathcal{F}^D\right) \geq \operatorname{Var}\left(Z_{s,t}|\mathcal{F}^{\tilde{D}}\right).$$

Our aim is to show how the optimisation problem of Lemma 6.2 can be used to exhibit lower bounds on 2D Young integrals with respect to R. In order to do this we need to take a detour via two technical lemmas. The first is the following continuity result for the conditional covariance, which we need approximate when passing to a limit from a discrete partition. The situation we will often have is two subintervals $[s, t] \subseteq [0, S]$ of $[0, T]$, and a sequence of sets $(D_n)_{n=1}^{\infty}$of the form

$$D_n = D_n^1 \cup D_n^2.$$

$\left(D_n^1\right)_{n=1}^{\infty}$ and $\left(D_n^2\right)_{n=1}^{\infty}$ here will be nested sequences of partitions of $[0, s]$ and $[t, S]$ respectively with *mesh* $\left(D_n^i\right) \to 0$ as $n \to \infty$ for $i = 1, 2$. If

$$\mathcal{F}^D := \sigma\left(Z_{u,v} : \{u, v\} \subseteq D\right),$$

then we can define a filtration $(\mathcal{G}_n)_{n=1}^{\infty}$ by $\mathcal{G}_n := \mathcal{F}^{D_n^1} \vee \mathcal{F}^{D_n^2}$ and ask about the convergence of

$$\operatorname{Cov}\left(Z_{p,q} Z_{u,v}|\mathcal{G}_n\right)$$

as $n \to \infty$ for subintervals $[p,q]$ and $[u,v]$ are subintervals of $[0,S]$. The following lemma records the relevant continuity statement.

Lemma 6.4. *For any p,q,u,v such that $[p,q]$ and $[u,v]$ are subintervals of $[0,S] \subseteq [0,T]$ we have*

$$\text{Cov}\left(Z_{p,q}Z_{u,v}|\mathcal{G}_n\right) \to \text{Cov}\left(Z_{p,q}Z_{u,v}|\mathcal{F}_{0,s} \vee \mathcal{F}_{t,S}\right),$$

as $n \to \infty$.

Proof. The martingale convergence theorem gives

$$\text{Cov}\left(Z_{p,q}Z_{u,v}|\mathcal{G}_n\right) \to \text{Cov}\left(Z_{p,q}Z_{u,v}| \vee_{n=1}^{\infty} \mathcal{G}_n\right), \text{ a.s. and in } L^p \text{ for all } p \geq 1.$$

The continuity of Z and the fact that $mesh\,(D_n) \to 0$ easily implies that, modulo null sets, one has $\vee_{n=1}^{\infty}\mathcal{G}_n = \mathcal{F}_{0,s} \vee \mathcal{F}_{t,T}$.

□

We now introduce another condition on Z, which we will later discard. This condition is virtually the same as Condition 3, the only difference being that we insist on the strict positivity of the conditional variance.

Condition 6. *Let $(Z_t)_{t\in[0,T]}$ be a real-valued continuous Gaussian process. We will assume that for every $[u,v] \subseteq [s,t] \subseteq [0,S] \subseteq [0,T]$ we have*

$$\text{Cov}\left(Z_{s,t}, Z_{u,v}|\mathcal{F}_{0,s} \vee \mathcal{F}_{t,S}\right) > 0. \tag{6.10}$$

The second technical lemma we need will apply whenever we work with a Gaussian process that satisfies Condition 6. It delivers a nested sequence of partitions, with mesh tending to zero, and such that the discretisation of Z along each partition will satisfy the dual feasibility condition (i.e. (6.6) in Lemma 6.2).

Lemma 6.5. *Let $(Z_t)_{t\in[0,T]}$ be a continuous Gaussian process that satisfies Condition 6. Then for every $0 \leq s < t \leq S \leq T$ there exists a nested sequence of partitions*

$$(D_m)_{m=1}^{\infty} = \left(\{t_i^m : i = 0,1,\ldots,n_m\}\right)_{m=1}^{\infty}$$

of $[0,S]$ such that:

(1) *$mesh(D_m) \to 0$ as $m \to \infty$;*
(2) *$\{s,t\} \subseteq D_m$ for all m;*
(3) *If Z_1^m and Z_2^m are the jointly Gaussian vectors*

$$Z_1^m = \left(Z_{t_i^m,t_{i+1}^m} : t_i^m \in D_m \cap ([0,s) \cup [t,S))\right),$$
$$Z_2^m = \left(Z_{t_i^m,t_{i+1}^m} : t_i^m \in D_m \cap [s,t)\right),$$

with respective covariance matrices Q_{11}^m and Q_{22}^m. Then the Gaussian vector (Z_1^m, Z_2^m) has a covariance matrix of the form

$$Q^m = \begin{pmatrix} Q_{11}^m & Q_{12}^m \\ (Q_{12}^m)^T & Q_{22}^m \end{pmatrix},$$

and the Schur complement of Q_{11}^m in Q^m has non-negative row sums.

Proof. See the appendix. □

The next result shows how we can bound from below the $2D$ Young integral of a Hölder-continuous f against R. The lower bound thus obtained is expressed in terms of the minimum of f, and the conditional variance of the Gaussian process.

Proposition 6.6. *Suppose $R : [0,T]^2 \to \mathbb{R}$ is the covariance function of some continuous Gaussian process $(Z_t)_{t\in[0,T]}$. Suppose R has finite 2D ρ-variation for some ρ in $[1,2)$, and that Z is non-degenerate and has a positive conditional covariance (i.e. satisfies Condition 3)). Let $\gamma \in (0,1)$ be such that $1/\rho + \gamma > 1$ and assume $f \in C^\gamma([0,T],\mathbb{R})$. Then for every $[s,t] \subseteq [0,T]$ we have the following lower bound on the 2D-Young integral of f against R :*

$$\int_{[0,T]^2} f_u f_v dR(u,v) \geq \left(\inf_{u\in[s,t]} |f(u)|^2\right) \mathrm{Var}\left(Z_{s,t} | \mathcal{F}_{0,s} \vee \mathcal{F}_{t,T}\right).$$

Remark 6.7. *We emphasise again that $\mathcal{F}_{a,b}$ is the σ-algebra generated by the increments of the form $Z_{u,v}$ for $u,v \in [a,b]$.*

Proof. Fix $[s,t] \subseteq [0,T]$, and take $b := \inf_{u\in[s,t]} |f(u)|$.

Step 1: We first note that there is no loss of generality in assuming the stronger Condition 6 instead of Condition 3. To see this, let $(B_t)_{t\in[0,T]}$ be a Brownian motion, which is independent of $(Z_t)_{t\in[0,T]}$, and for every $\epsilon > 0$ define the perturbed process

$$Z_t^\epsilon := Z_t + \epsilon B_t.$$

It is easy to check that Z^ϵ satisfies the conditions in the statement. Let $\mathcal{F}_{p,q}^\epsilon$ be the σ-algebra generated by the increments $Z_{u,v}^\epsilon$ between times p and q (note that $\mathcal{F}_{p,q}^\epsilon$ actually equals $\mathcal{F}_{p,q} \vee \sigma(B_{l,m} : u \leq l < m \leq q)$), and note that we have

$$\mathrm{Cov}\left(Z_{s,t}^\epsilon, Z_{u,v}^\epsilon | \mathcal{F}_{0,s}^\epsilon \vee \mathcal{F}_{t,T}^\epsilon\right) = \mathrm{Cov}\left(Z_{s,t}, Z_{u,v} | \mathcal{F}_{0,s} \vee \mathcal{F}_{t,T}\right) + \epsilon^2 (u-v) > 0$$

for every $0 \leq s < u < v \leq t \leq T$. It follows that Z^ϵ satisfies Condition 6. Let R^ϵ denote the covariance function of Z^ϵ. If we could prove the result with the additional hypothesis of Condition 6, then it would follow that

$$\begin{aligned}\int_{[0,T]^2} f_u f_v dR^\epsilon(u,v) &\geq b^2 \mathrm{Var}\left(Z_{s,t}^\epsilon | \mathcal{F}_{0,s}^\epsilon \vee \mathcal{F}_{t,T}^\epsilon\right) \\ &= b^2 \mathrm{Var}\left(Z_{s,t} | \mathcal{F}_{0,s} \vee \mathcal{F}_{t,T}\right) + b^2\epsilon^2 (t-s). \qquad (6.11)\end{aligned}$$

Because

$$\int_{[0,T]^2} f_u f_v dR^\epsilon(u,v) = \int_{[0,T]^2} f_u f_v dR(u,v) + \epsilon^2 |f|_{L^2[0,T]}^2,$$

the result for Z will then follow from (6.11) by letting ϵ tend to zero.

Step 2: We now prove the result under the additional assumption of Condition 6. By considering $-f$ if necessary we may assume that f is bounded from below by b on $[s,t]$. Since we now assume Condition 6 we can use Lemma 6.5 to obtain a nested sequence of partitions $(D_r)_{r=1}^\infty$ such that $\{s,t\} \subset D_r$ for all r, $\mathrm{mesh}(D_r) \to 0$ as $r \to \infty$, and such that the dual feasibility condition (property 3 in the Lemma 6.5) holds. Suppose $D = \{t_i : i = 0,1,\ldots,n\}$ is any partition of $[0,T]$ in this sequence (i.e. $D = D_r$ for some r). Then for some $l < m \in \{0,1,\ldots,n-1\}$ we have $t_l = s$ and $t_m = t$. Denote by $f(D)$ the column vector

$$f(D) = (f(t_0),\ldots,f(t_{n-1}))^T \in \mathbb{R}^n,$$

and $Q = (Q_{i,j})_{1\leq i,j<n}$ the symmetric $n \times n$ matrix with entries

$$Q_{ij} = R\begin{pmatrix} t_{i-1}, t_i \\ t_{j-1}, t_j \end{pmatrix} = E\left[Z_{t_{i-1},t_i} Z_{t_{j-1},t_j}\right].$$

From the non-degeneracy of Z it follows that Q is positive definite. The Riemann sum approximation to the 2D integral of f against R along the partition D can be written as

$$\sum_{i=1}^{n}\sum_{j=1}^{n} f_{t_{i-1}} f_{t_{j-1}} R\begin{pmatrix} t_{i-1}, t_i \\ t_{j-1}, t_j \end{pmatrix} = \sum_{i=1}^{n}\sum_{j=1}^{n} f_{t_{i-1}} f_{t_{j-1}} Q_{i,j} = f(D)^T Q f(D). \tag{6.12}$$

If necessary, we can ensure that that last $m-l$ components of $f(D)$ are bounded below by b. To see this, we simply permute its coordinates using any bijective map $\tau : \{1, \ldots, n\} \to \{1, \ldots, n\}$ which has the property that

$$\tau(l+j) = n - m + l + j, \text{ for } j = 0, 1, \ldots, m-l.$$

Fix one such map τ, and let $f_\tau(D)$ denote the vector resulting from applying τ to the coordinates of $f(D)$. Similarly, let $Q_\tau = \left(Q^\tau_{i,j}\right)_{1\leq i,j<n}$ be the $n \times n$ matrix

$$Q^\tau_{ij} = Q_{\tau(i)\tau(j)},$$

and note that Q^τ is the covariance matrix of the Gaussian vector

$$Z = \left(Z_{t_{\tau(1)-1},t_{\tau(1)}}, \ldots, Z_{t_{\tau(n)-1},t_{\tau(n)}}\right).$$

A simple calculation shows that

$$f(D)^T Q f(D) = f_\tau(D)^T Q_\tau f_\tau(D).$$

We can apply Lemma 6.2 because condition (6.6) is guaranteed to hold by the properties of the sequence $(D_r)_{r=1}^\infty$. We deduce that

$$f(D)^T Q f(D) = f_\tau(D)^T Q_\tau f_\tau(D) \geq b^2 \sum_{i,j=1}^{m-l} S_{ij}, \tag{6.13}$$

where S is the $(m-l)\times(m-l)$ matrix obtained by taking the Schur complement of the leading principal $(n-m+l)\times(n-m+l)$ minor of $\tilde{Q}$. As already mentioned, the distribution of a Gaussian vector conditional on some of its components remains Gaussian; the conditional covariance is described by a suitable Schur complement . In this case, this means we have that

$$S = \operatorname{Cov}\left[\left(Z_{t_l,t_{l+1}}, \ldots, Z_{t_{m-1},t_m}\right) | Z_{t_{j-1},t_j},\ j \in \{1, \ldots, l\} \cup \{m+1, \ldots, n\}\right]. \tag{6.14}$$

If we define

$$\mathcal{F}^D := \sigma\left(Z_{t_{j-1},t_j} :\ j \in \{1, \ldots, l\} \cup \{m+1, \ldots, n\}\right),$$

to be the σ-algebra generated by the increments of Z in $D \setminus [s,t]$, then using (6.14) we arrive at

$$\begin{aligned}
\sum_{i,j=1}^{m-l} S_{ij} &= \sum_{i,j=1}^{m-l-1} E\left[\left(Z_{t_{l+i-1},t_{l+i}}\right)\left(Z_{t_{l+j-1},t_{l+j}}\right) | \mathcal{F}^D\right] \\
&\quad - \sum_{i,j=1}^{m-l-1} E\left[\left(Z_{t_{l+i-1},t_{l+i}}\right) | \mathcal{F}^D\right] E\left[\left(Z_{t_{l+j-1},t_{l+j}}\right) | \mathcal{F}^D\right] \\
&= E\left[(Z_{s,t})^2 | \mathcal{F}^D\right] - E\left[Z_{s,t} | \mathcal{F}^D\right]^2 = \operatorname{Var}\left(Z_{s,t} | \mathcal{F}^D\right).
\end{aligned} \tag{6.15}$$

To finish the proof we note that $\mathcal{F}^D \subseteq \mathcal{F}_{0,s} \vee \mathcal{F}_{t,T}$, and exploit the monotonicity of the conditional variance described by Lemma 6.3 to give

$$\operatorname{Var}\left(Z_{s,t}|\mathcal{F}^D\right) \geq \operatorname{Var}\left(Z_{s,t}|\mathcal{F}_{0,s} \vee \mathcal{F}_{t,T}\right). \tag{6.16}$$

Then by combining (6.16), (6.15) and (6.13) in (6.12) we obtain

$$\sum_{i=1}^{n}\sum_{j=1}^{n} f_{t_{i-1}} f_{t_{j-1}} Q_{i,j} \geq b^2 \operatorname{Var}\left(Z_{s,t}|\mathcal{F}_{0,s} \vee \mathcal{F}_{t,T}\right).$$

Because this inequality holds for any $D \in (D_r)_{r=1}^{\infty}$, we can apply it for $D = D_r$ and let $r \to \infty$ to give:

$$\int_{[0,T]^2} f_u f_v dR\left(u,v\right) \geq b^2 \operatorname{Var}\left(Z_{s,t}|\mathcal{F}_{0,s} \vee \mathcal{F}_{t,T}\right),$$

whereupon the proof is complete. □

We now deliver on a promise we made in Section 3 by proving that the diagonal dominance of the increments implies the positivity of the conditional covariance.

Corollary 6.8. *Let $(Z_t)_{t\in[0,T]}$ be a real-valued continuous Gaussian process. If Z satisfies Condition 4 then it also satisfies Condition 3.*

Proof. Fix $s < t$ in $[0,T]$, let $(D_n)_{n=1}^{\infty}$ be a sequence of partitions having the properties described in the statement of Lemma 6.4 and suppose $[u,v] \subseteq [s,t]$. From the conclusion of Lemma 6.4 we have that

$$\operatorname{Cov}\left(Z_{s,t}Z_{u,v}|\mathcal{G}_n\right) \to \operatorname{Cov}\left(Z_{s,t}Z_{u,v}|\mathcal{F}_{0,s} \vee \mathcal{F}_{t,T}\right) \tag{6.17}$$

as $n \to \infty$. Let Z_n be the Gaussian vector whose components consist of the increments of Z over all the consecutive points in the partition $D_n \cup \{s,u,v,t\}$. Let Q denote the covariance matrix of Z_n. The left hand side of (6.17) is the sum of all the entries in some row of a particular Schur complement of Q. Z is assumed to have diagonally dominant increments. Any such Schur complement of Q will therefore be diagonally dominant, since diagonal dominance is preserved under Schur-complementation (see [32]). As diagonally dominant matrices have non-negative row sums it follows that $\operatorname{Cov}\left(Z_{s,t}Z_{u,v}|\mathcal{G}_n\right)$ is non-negative, and hence the limit in (6.17) is too. □

We are now in a position to generalise the L^2-interpolation inequality (6.5) stated earlier.

Theorem 6.9 (interpolation)**.** *Let $(Z_t)_{t\in[0,T]}$ be a continuous Gaussian process with covariance function $R : [0,T]^2 \to \mathbb{R}$. Suppose R has finite two-dimensional ρ-variation for some ρ in $[1,2)$. Assume that Z is non-degenerate in the sense of Definition 3.2, and has positive conditional covariance (i.e. satisfies Condition 3). Suppose $f \in C\left([0,T],\mathbb{R}\right)$ with $\gamma + 1/\rho > 1$. Then for every $0 < S \leq T$ at least one of the following inequalities is always true:*

$$\|f\|_{\infty;[0,S]} \leq 2E\left[Z_S^2\right]^{-1/2}\left(\int_{[0,S]^2} f_s f_t dR\left(s,t\right)\right)^{1/2}, \tag{6.18}$$

or, for some interval $[s,t] \subseteq [0,S]$ of length at least

$$\left(\frac{\|f\|_{\infty;[0,S]}}{2\,\|f\|_{\gamma;[0,S]}}\right)^{1/\gamma},$$

we have

$$\frac{1}{4}\|f\|_{\infty;[0,S]}^{2}\operatorname{Var}\left(Z_{s,t}|\mathcal{F}_{0,s}\vee\mathcal{F}_{t,S}\right)\leq\int_{[0,S]^2}f_v f_{v'}dR\left(v,v'\right). \tag{6.19}$$

Proof. We take $S=T$, the generalisation to $0<S<T$ needing only minor changes. f is continuous and therefore achieves its maximum in $[0,T]$. Thus, by considering $-f$ if necessary, we can find $t\in[0,T]$ such that

$$f\left(t\right)=\|f\|_{\infty;[0,T]}.$$

There are two possibilities which together are exhaustive. In the first case f never takes any value less than half its maximum, i.e.

$$\inf_{u\in[0,T]}f\left(u\right)\geq\frac{1}{2}\|f\|_{\infty;[0,T]}.$$

Hence we can can apply Proposition 6.6 to deduce (6.18). In the second case, there exists $u\in[0,T]$ such that $f\left(u\right)=2^{-1}\|f\|_{\infty;[0,T]}$. Then, assuming that $u<t$ (the argument for $u>t$ leads to the same outcome), we can define

$$s=\sup\left\{v<t:f\left(v\right)\leq\frac{1}{2}\|f\|_{\infty;[0,T]}\right\}.$$

By definition f is then bounded below by $\|f\|_{\infty;[0,T]}/2$ on $[s,t]$. The Hölder continuity of f gives a lower bound on the length of this interval in an elementary way

$$\frac{1}{2}\|f\|_{\infty;[0,T]}=\left|f\left(t\right)-f\left(s\right)\right|\leq\|f\|_{\gamma;[0,T]}\left|t-s\right|^{\gamma},$$

which yields

$$\left|t-s\right|\geq\left(\frac{\|f\|_{\infty;[0,T]}}{2\|f\|_{\gamma;[0,T]}}\right)^{1/\gamma}.$$

Another application of Proposition 6.6 then gives (6.19). □

Corollary 6.10. *Assume Condition 2 so that the ρ-variation of R is Hölder-controlled, and for some $c>0$ and some $\alpha\in(0,1)$ we have the lower bound on the conditional variance:*

$$\operatorname{Var}\left(Z_{s,t}|\mathcal{F}_{0,s}\vee\mathcal{F}_{t,T}\right)\geq c\left(t-s\right)^{\alpha}.$$

Theorem 6.9 then allows us to bound $\|f\|_{\infty;[0,T]}$ above by the maximum of

$$2E\left[Z_T^2\right]^{-1/2}\left(\int_{[0,T]^2}f_s f_t dR\left(s,t\right)\right)^{1/2}$$

and

$$\frac{2}{\sqrt{c}}\left(\int_{[0,T]^2}f_s f_t dR\left(s,t\right)\right)^{\gamma/(2\gamma+\alpha)}\|f\|_{\gamma;[0,T]}^{\alpha/(2\gamma+\alpha)}.$$

Proof. This is immediate from Theorem 6.9. □

In particular, if Z is a Brownian motion we have $\operatorname{Var}\left(Z_{s,t}|\mathcal{F}_{0,s}\vee\mathcal{F}_{t,T}\right)=(t-s)$, hence Corollary 6.10 shows that

$$\|f\|_{\infty;[0,T]}\leq 2\max\left(T^{-1/2}\left|f\right|_{L^2[0,T]},\left|f\right|_{L^2[0,T]}^{2\gamma/(2\gamma+1)}\|f\|_{\gamma;[0,T]}^{1/(2\gamma+1)}\right),$$

which is exactly (6.5). We have therefore achieved out goal of generalising this inequality.

7. Malliavin differentiability of the flow

7.1. **High order directional derivatives.** Let $\mathbf{x}$ be in $WG\Omega_p\left(\mathbb{R}^d\right)$ and suppose that the vector fields $V=(V_1,\ldots,V_d)$ and V_0 are smooth and bounded. For $t\in[0,T]$ we let $U^{\mathbf{x}}_{t\leftarrow 0}(\cdot)$ denote the map defined by

$$U^{\mathbf{x}}_{t\leftarrow 0}(\cdot): y_0 \mapsto y_t,$$

where y is the solution to the RDE

$$dy_t = V(y_t)\,d\mathbf{x}_t + V_0(y_t)\,dt, \quad y(0)=y_0. \tag{7.1}$$

It is well-known (see [12]) that the flow (i.e. the map $y_0\mapsto U^{\mathbf{x}}_{t\leftarrow 0}(y_0)$) is differentiable; its derivative (or Jacobian) is the linear map

$$J^{\mathbf{x}}_{t\leftarrow 0}(y_0)(\cdot) \equiv \frac{d}{d\epsilon} U^{\mathbf{x}}_{t\leftarrow 0}(y_0+\epsilon\cdot)\bigg|_{\epsilon=0} \in L(\mathbb{R}^e,\mathbb{R}^e).$$

If we let $\Phi^{\mathbf{x}}_{t\leftarrow 0}(y_0)$ denote the pair

$$\Phi^{\mathbf{x}}_{t\leftarrow 0}(y_0) = \left(U^{\mathbf{x}}_{t\leftarrow 0}(y_0), J^{\mathbf{x}}_{t\leftarrow 0}(y_0)\right) \in \mathbb{R}^e \oplus L(\mathbb{R}^e,\mathbb{R}^e),$$

and if $W=(W_1,\ldots,W_d)$ is the collection vector fields given by

$$W_i(y,J) = \left(V_i(y), \nabla V_i(y)\cdot J\right), \quad i=1,\ldots,d,$$

and

$$W_0(y,J) = \left(V_0(y), \nabla V_0(y)\cdot J\right)$$

then $\Phi^{\mathbf{x}}_{t\leftarrow 0}(y_0)$ is the solution[1] to the RDE

$$d\Phi^{\mathbf{x}}_{t\leftarrow 0} = W\left(\Phi^{\mathbf{x}}_{t\leftarrow 0}\right)d\mathbf{x}_t + W_0\left(\Phi^{\mathbf{x}}_{t\leftarrow 0}\right)dt, \Phi^{\mathbf{x}}_{t\leftarrow 0}|_{t=0} = (y_0, I).$$

In fact, the Jacobian is invertible as a linear map and the inverse, which we will denote $J^{\mathbf{x}}_{0\leftarrow t}(y_0)$, is also a solution to an RDE (again jointly with the base flow $U^{\mathbf{x}}_{t\leftarrow 0}(y_0)$). We also recall the relation

$$J^{\mathbf{x}}_{t\leftarrow s}(y) := \frac{d}{d\epsilon} U^{\mathbf{x}}_{t\leftarrow s}(y+\epsilon\cdot)\bigg|_{\epsilon=0} = J^{\mathbf{x}}_{t\leftarrow 0}(y)\cdot J^{\mathbf{x}}_{0\leftarrow s}(y).$$

Notation 3. *In what follows we will let*

$$M^{\mathbf{x}}_{\cdot\leftarrow 0}(y_0) \equiv \left(U^{\mathbf{x}}_{t\leftarrow 0}(y_0), J^{\mathbf{x}}_{t\leftarrow 0}(y_0), J^{\mathbf{x}}_{0\leftarrow t}(y_0)\right) \in \mathbb{R}^e \oplus \mathbb{R}^{e\times e} \oplus \mathbb{R}^{e\times e}. \tag{7.2}$$

For any path h in $C^{q-\mathrm{var}}\left([0,T],\mathbb{R}^d\right)$ with $1/q+1/p>1$ we can canonically define the translated rough path $T_h\mathbf{x}$ (see [12]). Hence, we have the directional derivative

$$D_h U^{\mathbf{x}}_{t\leftarrow 0}(y_0) \equiv \frac{d}{d\epsilon} U^{T_{\epsilon h}\mathbf{x}}_{t\leftarrow 0}(y_0)\bigg|_{\epsilon=0}.$$

It is not difficult to show that

$$D_h U^{\mathbf{x}}_{t\leftarrow 0}(y_0) = \sum_{i=1}^d \int_0^t J^{\mathbf{x}}_{t\leftarrow s}(y_0)\, V_i\left(U^{\mathbf{x}}_{s\leftarrow 0}(y_0)\right) dh^i_s,$$

which implies by Young's inequality that

$$\left|D_h U^{\mathbf{x}}_{t\leftarrow 0}(y_0)\right| \le C \left\|M^{\mathbf{x}}_{\cdot\leftarrow 0}(y_0)\right\|_{p\text{-var};[0,t]} |h|_{q\text{-var};[0,t]}. \tag{7.3}$$

[1] A little care is needed because the vector fields have linear growth (and hence are not Lip-γ. But one can exploit the 'triangular' dependence structure in the vector fields to rule out the possibility of explosion. See [12] for details.

In this section we will be interested in the form of the higher order directional derivatives

$$D_{h_1}\dots D_{h_n}U_{t\leftarrow 0}^{\mathbf{x}}(y_0) := \frac{\partial^n}{\partial\epsilon_1,\dots\partial\epsilon_n}U_{t\leftarrow 0}^{T_{\epsilon_n h_n}\dots T_{\epsilon_1 h_1}\mathbf{x}}(y_0)\bigg|_{\epsilon_1=\dots=\epsilon_n=0}.$$

Our aim will be to obtain bounds of the form (7.3). To do this in a systematic way is a challenging exercise. We rely on the treatment presented in [21]. For the reader's convenience when comparing the two accounts, we note that [21] uses the notation

$$\left(D_s U_{t\leftarrow 0}^{\mathbf{x}}(y_0)\right)_{s\in[0,T]} = \left(D_s^1 U_{t\leftarrow 0}^{\mathbf{x}}(y_0),\dots,D_s^d U_{t\leftarrow 0}^{\mathbf{x}}(y_0)\right)_{s\in[0,T]} \in \mathbb{R}^d$$

to identify the derivative. The relationship between $D_s U_{t\leftarrow 0}^{\mathbf{x}}(y_0)$ and $D_h U_{t\leftarrow 0}^{\mathbf{x}}(y_0)$ is simply that

$$D_h U_{t\leftarrow 0}^{\mathbf{x}}(y_0) = \sum_{i=1}^d \int_0^t D_s^i U_{t\leftarrow 0}^{\mathbf{x}}(y_0)\,dh_s^i.$$

Note, in particular $D_s U_{t\leftarrow 0}^{\mathbf{x}}(y_0) = 0$ if $t < s$.

Proposition 7.1. *Assume $\mathbf{x}$ is in $WG\Omega_p\left(\mathbb{R}^d\right)$ and let $V=(V_1,\dots,V_d)$ be a collection of smooth and bounded vector fields. Denote the solution flow to the RDE (7.1) by*

$$U_{t\leftarrow 0}^{\mathbf{x}}(y_0) = \left(U_{t\leftarrow 0}^{\mathbf{x}}(y_0)_1,\dots,U_{t\leftarrow 0}^{\mathbf{x}}(y_0)_e\right) \in \mathbb{R}^e$$

Suppose $q\geq 1$ and $n\in\mathbb{N}$ and let $\{h_1,\dots,h_n\}$ be any subset of $C^{q-var}\left([0,T],\mathbb{R}^d\right)$. Then the directional derivative $D_{h_1}\dots D_{h_n}U_{t\leftarrow 0}^{\mathbf{x}}(y_0)$ exists for any $t\in[0,T]$. Moreover, there exists a collection of finite indexing sets

$$\left\{\mathbf{K}_{(i_1,\dots,i_n)} : (i_1,\dots,i_n)\in\{1,\dots,d\}^n\right\},$$

such that for every $j\in\{1,..,e\}$ we have the identity

$$D_{h_1}\dots D_{h_n}U_{t\leftarrow 0}^{\mathbf{x}}(y_0)_j = \sum_{i_1,\dots,i_n=1}^d \sum_{k\in\mathbf{K}_{(i_1,\dots,i_n)}} \int_{0<t_1<\dots<t_n<t} f_1^k(t_1)\dots f_n^k(t_n)\,f_{n+1}^k(t)\,dh_{t_1}^{i_1}\dots dh_{t_n}^{i_n}, \tag{7.4}$$

for some functions f_l^k which are in $C^{p-var}([0,T],\mathbb{R})$ for every l and k, i.e.

$$\cup_{(i_1,\dots,i_n)\in\{1,\dots,d\}^n}\cup_{k\in\mathbf{K}_{(i_1,\dots,i_n)}}\left\{f_l^k : l=1,..,n+1\right\}\subset C^{p-var}([0,T],\mathbb{R}).$$

Furthermore, there exists a constant C, which depends only on n and T such that

$$\left|f_l^k\right|_{p\text{-}var;[0,T]} \leq C\left(1+\left\|M_{\cdot\leftarrow 0}^{\mathbf{x}}(y_0)\right\|_{p\text{-}var;[0,T]}\right)^p \tag{7.5}$$

for every $l=1,\dots,n+1$, every $k\in\mathbf{K}_{(i_1,\dots,i_n)}$ and every $(i_1,\dots,i_n)\in\{1,\dots,d\}^n$.

Proof. We observe that $D_{h_1}\dots D_{h_n}U_{t\leftarrow 0}^{\mathbf{x}}(y_0)_j$ equals

$$\sum_{i_1,\dots,i_n=1}^d \int_{0<t_1<\dots<t_n<t} D_{t_1\dots t_n}^{i_1\dots i_n}U_{t\leftarrow 0}^{\mathbf{x}}(y_0)_j\,dh_{t_1}^{i_1}\dots dh_{t_n}^{i_n}. \tag{7.6}$$

The representation for the integrand in (7.6) derived in Proposition 4.4 in [21] then allows us to deduce (7.4) and (7.5). □

7.1.1. *Malliavin differentiability.* We now switch back to the context of a continuous Gaussian process $(X_t)_{t\in[0,T]} = \left(X_t^1,\dots,X_t^d\right)_{t\in[0,T]}$ with i.i.d. components associated to the abstract Wiener space $(\mathcal{W},\mathcal{H},\mu)$. Under the assumption of finite 2d ρ-variation we have already remarked that, for any $p>2\rho$, X has a unique natural lift to a geometric p-rough path $\mathbf{X}$. But the assumption of finite ρ-variation on the covariance also gives rise to the embedding

$$\mathcal{H}\hookrightarrow C^{q-\text{var}}\left([0,T],\mathbb{R}^d\right) \tag{7.7}$$

for the Cameron-Martin space, for any $1/p+1/q>1$ [7, Prop. 2]. The significance of this result it twofold. First, it is proved in [7, Prop. 3] that it implies the existence of a (measurable) subset $\mathcal{V}\subset\mathcal{W}$ with $\mu(\mathcal{V})=1$ on which

$$T_h\mathbf{X}(\omega)\equiv\mathbf{X}(\omega+h),$$

for all $h\in\mathcal{H}$ simultaneously. It follows that the Malliavin derivative $\mathcal{D}U_{t\leftarrow0}^{\mathbf{X}(\omega)}(y_0):\mathcal{H}\to\mathbb{R}^e$

$$\mathcal{D}U_{t\leftarrow0}^{\mathbf{X}(\omega)}(y_0):h\mapsto\mathcal{D}_hU_{t\leftarrow0}^{\mathbf{X}(\omega)}(y_0):=\left.\frac{d}{d\epsilon}U_{t\leftarrow0}^{\mathbf{X}(\omega+\epsilon h)}(y_0)\right|_{\epsilon=0}, \tag{7.8}$$

coincides with the directional derivative of the previous section, i.e.

$$\left.\frac{d}{d\epsilon}U_{t\leftarrow0}^{\mathbf{X}(\omega+\epsilon h)}(y_0)\right|_{\epsilon=0}=\left.\frac{d}{d\epsilon}U_{t\leftarrow0}^{T_{\epsilon h}\mathbf{X}}(y_0)\right|_{\epsilon=0}. \tag{7.9}$$

The second important consequence results from combining (7.7), (7.9) and (7.3), namely that

$$\left\|\mathcal{D}U_{t\leftarrow0}^{\mathbf{X}(\omega)}(y_0)\right\|_{op}\le C\left\|M_{\cdot\leftarrow0}^{\mathbf{X}(\omega)}(y_0)\right\|_{p\text{-var};[0,t]}. \tag{7.10}$$

If we can show that the right hand side of (7.10) has finite positive moments of all order, then these observations lead to the conclusion that

$$Y_t=U_{t\leftarrow0}^{\mathbf{X}}(y_0)\in\cap_{p>1}\mathbb{D}^{1,p}\left(\mathbb{R}^e\right),$$

where $\mathbb{D}^{k,p}$ is the Shigekawa-Sobolev space (see Nualart). The purpose of Proposition 7.1 is to extend this argument to the higher order derivatives. We will make this more precisely shortly, but first we remark that the outline just given is what motivates the assumption

$$\mathcal{H}\hookrightarrow C^{q-\text{var}}\left([0,T],\mathbb{R}^d\right)$$

detailed in Condition 1[2].

The following theorem follows from the recent paper [5]. It asserts the sufficiency of Condition 1 to show the existence of finite moments for the p-variation of the Jacobian of the flow (and its inverse).

Theorem 7.2 (Cass-Litterer-Lyons (CLL))**.** *Let $(X_t)_{t\in[0,T]}$ be a continuous, centred Gaussian process in $\mathbb{R}^d$ with i.i.d. components. Let X satisfy Condition 1, so that for some $p\ge1$, X admits a natural lift to a geometric p-rough path $\mathbf{X}$. Assume $V=(V_0,V_1,\dots,V_d)$ is any collection of smooth bounded vector fields on $\mathbb{R}^e$ and let $U_{t\leftarrow0}^{\mathbf{X}}(\cdot)$ denote the solution flow to the RDE*

$$\begin{aligned}&dU_{t\leftarrow0}^{\mathbf{X}}(y_0)=V\left(U_{t\leftarrow0}^{\mathbf{X}}(y_0)\right)d\mathbf{X}_t+V_0\left(U_{t\leftarrow0}^{\mathbf{X}}(y_0)\right)dt,\\&U_{0\leftarrow0}^{\mathbf{X}}(y_0)=y_0.\end{aligned}$$

[2]The requirement of complementary regularity in the Condition 1 then amounts to $\rho\in[1,3/2)$. This covers BM, the OU process and the Brownian bridge (all with $\rho=1$) and fBm for $H>1/3$ (taking $\rho=1/2H$). For the special case of fBm one can actually improve on this general embedding statement via Remark 2.6 . The requirement of complementary then leads to the looser restriction $H>1/4$.

Then the map $U^{\mathbf{X}}_{t\leftarrow 0}(\cdot)$ *is differentiable with derivative* $J^{\mathbf{X}}_{t\leftarrow 0}(y_0)\in\mathbb{R}^{e\times e}$*;* $J^{\mathbf{X}}_{t\leftarrow 0}(y_0)$ *is invertible as a linear map with inverse denoted by* $J^{\mathbf{X}}_{0\leftarrow t}(y_0)$*. Furthermore, if we define*

$$M^{\mathbf{X}}_{\cdot\leftarrow 0}(y_0)\equiv\left(U^{\mathbf{X}}_{t\leftarrow 0}(y_0),J^{\mathbf{X}}_{t\leftarrow 0}(y_0),J^{\mathbf{X}}_{0\leftarrow t}(y_0)\right)\in\mathbb{R}^{e}\oplus\mathbb{R}^{e\times e}\oplus\mathbb{R}^{e\times e},$$

and assume X *satisfies Condition 1, we have that*

$$\left\|M^{\mathbf{X}}_{\cdot\leftarrow 0}(y_0)\right\|_{p\text{-}var;[0,T]}\in\bigcap_{q\geq 1}L^{q}(\mu).$$

Proof. This follows from by repeating the steps of [5] generalized to incorporate a drift term. □

Remark 7.3. *Under the additional assumption that the covariance* R *has finite Hölder-controlled* ρ*-variation, it is possible to prove a version of this theorem showing that:*

$$\left\|M^{\mathbf{X}}_{\cdot\leftarrow 0}(y_0)\right\|_{1/p}\in\bigcap_{q\geq 1}L^{q}(\mu).$$

7.2. **Proof that** $U^{\mathbf{X}(\cdot)}_{t\leftarrow 0}(y_0)\in\mathbb{D}^{\infty}(\mathbb{R}^{e})$**.** We have already seen that appropriate assumptions on the covariance lead to the observation that for all $h\in\mathcal{H}$,

$$D_hU^{\mathbf{X}(\omega)}_{t\leftarrow 0}(y_0)\equiv\left.\frac{d}{d\epsilon}U^{T_h\mathbf{X}(\omega)}_{t\leftarrow 0}(y_0)\right|_{\epsilon=0}$$

for all ω in a set of μ-full measure. We will show that the Wiener functional $\omega\mapsto U^{\mathbf{X}(\omega)}_{t\leftarrow 0}(y_0)$ belongs to the Sobolev space $\mathbb{D}^{\infty}(\mathbb{R}^{e})$. Recall that

$$\mathbb{D}^{\infty}(\mathbb{R}^{e}):=\cap_{p>1}\cap_{k=1}^{\infty}\mathbb{D}^{k,p}(\mathbb{R}^{e}),$$

where $\mathbb{D}^{k,p}$ is the usual Shigekawa-Sobolev space, which is defined as the completion of the smooth random variables with respect to a Sobolev-type norm (see Nualart [30]). There is an equivalent characterisation of the spaces $\mathbb{D}^{k,p}$ (originally due to Kusuoka and Stroock), which is easier to use in the present context. We briefly recall the main features of this characterisation starting with the following definitions. Suppose E is a given Banach space and $F:\mathcal{W}\to E$ is a measurable function. Recall (see Sugita [31]) that F is called ray absolutely continuous (RAC) if for every $h\in\mathcal{H}$, there exists a measurable map $\tilde{F}_h:\mathcal{W}\to E$ satisfying:

$$F(\cdot)=\tilde{F}_h(\cdot),\ \mu-\text{a.e.},$$

and for every $\omega\in\mathcal{W}$

$$t\mapsto\tilde{F}_h(\omega+th)\ \text{ is absolutely continuous in } t\in\mathbb{R}.$$

And furthermore, F is called stochastically Gateaux differentiable (SGD) if there exists a measurable $G:\mathcal{W}\to L(\mathcal{H},E)$, such that for any $h\in\mathcal{H}$

$$\frac{1}{t}\left[F(\cdot+th)-F(\cdot)\right]\xrightarrow{\mu}G(\omega)(h)\ \text{ as } t\to 0,$$

where $\xrightarrow{\mu}$ indicates convergence in μ-measure.

If F is SGD, then its derivative G is unique μ-a.s. and we denote it by $\mathcal{D}F$. Higher order derivatives are defined inductively in the obvious way. Hence $\mathcal{D}^nF(\omega)$ (if it exists) is a multilinear map (in n variables) from $\mathcal{H}$ to E.

We now define the spaces $\tilde{\mathbb{D}}^{k,p}(\mathbb{R}^{e})$ for $1<p<\infty$ by

$$\tilde{\mathbb{D}}^{1,p}(\mathbb{R}^{e}):=\left\{F\in L^{p}(\mathbb{R}^{e}):F\text{ is RAC and SGD},\mathcal{D}F\in L^{p}(L(\mathcal{H},\mathbb{R}^{e}))\right\},$$

and for $k = 2, 3, \ldots$.

$$\tilde{\mathbb{D}}^{k,p}\left(\mathbb{R}^{e}\right) := \left\{F \in \tilde{\mathbb{D}}^{k-1,p}\left(\mathbb{R}^{e}\right) : \mathcal{D}F \in \tilde{\mathbb{D}}^{k-1,p}\left(L\left(\mathcal{H},\mathbb{R}^{e}\right)\right)\right\}.$$

Theorem 7.4 (Sugita [31])**.** *For $1 < p < \infty$ and $k \in \mathbb{N}$ we have $\tilde{\mathbb{D}}^{k,p}\left(\mathbb{R}^{e}\right) = \mathbb{D}^{k,p}\left(\mathbb{R}^{e}\right)$.*

It follows immediately from this result that we have

$$\mathbb{D}^{\infty}\left(\mathbb{R}^{e}\right) = \cap_{p>1} \cap_{k=1}^{\infty} \tilde{\mathbb{D}}^{k,p}\left(\mathbb{R}^{e}\right).$$

With these preliminaries out the way, we can prove the following.

Proposition 7.5. *Suppose $(X_t)_{t\in[0,T]}$ is an $\mathbb{R}^d$-valued, zero-mean Gaussian process with i.i.d components associated with the abstract Wiener space $(\mathcal{W},\mathcal{H},\mu)$.Assume that for some $p \geq 1$, X lifts to a geometric p-rough path $\mathbf{X}$. Let $V = (V_0, V_1, \ldots, V_d)$ be a collection of C^∞-bounded vector fields on $\mathbb{R}^e$, and let $U_{t\leftarrow 0}^{\mathbf{X}(\omega)}(y_0)$ denote the solution flow of the RDE*

$$dY_t = V\left(Y_t\right) d\mathbf{X}_t\left(\omega\right) + V_0\left(Y_t\right) dt,\ Y\left(0\right) = y_0.$$

Then, under the assumption that X satisfies Condition 1, we have that the Wiener functional

$$U_{t\leftarrow 0}^{\mathbf{X}(\cdot)}\left(y_0\right) : \omega \mapsto U_{t\leftarrow 0}^{\mathbf{X}(\omega)}\left(y_0\right)$$

is in $\mathbb{D}^\infty\left(\mathbb{R}^e\right)$ for every $t \in [0,T]$.

Proof. We have already remarked that Condition 1 implies that on a set of μ-full measure

$$T_h\mathbf{X}\left(\omega\right) \equiv \mathbf{X}\left(\omega + h\right) \tag{7.11}$$

for all $h \in \mathcal{H}$. It easily follows that $U_{t\leftarrow 0}^{\mathbf{X}(\cdot)}(y_0)$ is RAC. Furthermore, its stochastic Gateaux derivative is precisely the map $\mathcal{D}U_{t\leftarrow 0}^{\mathbf{X}(\omega)}(y_0)$ defined in (7.8). The relation (7.11) implies that the directional and Malliavin derivatives coincide (on a set of μ-full measure) hence $\mathcal{D}U_{t\leftarrow 0}^{\mathbf{X}(\omega)}(y_0) \in L\left(\mathcal{H},\mathbb{R}^e\right)$ is the map

$$\mathcal{D}U_{t\leftarrow 0}^{\mathbf{X}(\omega)}\left(y_0\right) : h \mapsto D_h U_{t\leftarrow 0}^{\mathbf{X}(\omega)}\left(y_0\right).$$

We have shown in (7.10) that

$$\left\|\mathcal{D}U_{t\leftarrow 0}^{\mathbf{X}(\omega)}\left(y_0\right)\right\|_{op} \leq C\left\|M_{\cdot\leftarrow 0}^{\mathbf{X}}\left(y_0\right)\right\|_{p\text{-var};[0,T]}, \tag{7.12}$$

where

$$M_{\cdot\leftarrow 0}^{\mathbf{X}}\left(y_0\right) \equiv \left(U_{t\leftarrow 0}^{\mathbf{X}}\left(y_0\right), J_{t\leftarrow 0}^{\mathbf{X}}\left(y_0\right), J_{0\leftarrow t}^{\mathbf{X}}\left(y_0\right)\right). \tag{7.13}$$

It follows from Theorem 7.2 that

$$\left\|M_{\cdot\leftarrow 0}^{\mathbf{X}}\left(y_0\right)\right\|_{p\text{-var};[0,T]} \in \cap_{p\geq 1} L^p\left(\mu\right).$$

Using this together with (7.12) proves that $U_{t\leftarrow 0}^{\mathbf{X}(\cdot)}(y_0)$ is in $\cap_{p>1}\tilde{\mathbb{D}}^{1,p}\left(\mathbb{R}^e\right)$ which equals $\cap_{p>1}\mathbb{D}^{1,p}\left(\mathbb{R}^e\right)$ by Theorem 7.4

We prove that $U_{t\leftarrow 0}^{\mathbf{X}(\cdot)}(y_0)$ is in $\cap_{p>1}\tilde{\mathbb{D}}^{k,p}\left(\mathbb{R}^e\right)$ for all $k \in \mathbb{N}$ by induction. If $U_{t\leftarrow 0}^{\mathbf{X}(\cdot)}(y_0) \in \tilde{\mathbb{D}}^{k-1,p}\left(\mathbb{R}^e\right)$ then, by the uniqueness of the stochastic Gateaux derivative, we must have

$$\mathcal{D}^{k-1}U_{t\leftarrow 0}^{\mathbf{X}(\omega)}\left(y_0\right)\left(h_1,\ldots,h_{k-1}\right) = D_{h_1}\ldots D_{h_k}U_{t\leftarrow 0}^{\mathbf{X}(\omega)}\left(y_0\right).$$

It is then easy to see that $\mathcal{D}^{k-1}U_{t\leftarrow 0}^{\mathbf{X}(\omega)}(y_0)$ is RAC and SGD. Moreover, the stochastic Gateaux derivative is

$$\mathcal{D}^{k}U_{t\leftarrow 0}^{\mathbf{X}(\omega)}\left(y_0\right) : \left(h_1,\ldots,h_k\right) = D_{h_1}\ldots D_{h_k}U_{t\leftarrow 0}^{\mathbf{X}(\omega)}\left(y_0\right).$$

It follows from Proposition 7.1 together with Condition 1 that we can bound the operator norm of $\mathcal{D}^k U_{t\leftarrow 0}^{\mathbf{X}(\omega)}(y_0)$ in the following way:

$$\|\mathcal{D}^k U_{t\leftarrow 0}^{\mathbf{X}(\omega)}(y_0)\|_{op} \leq C\left(1+\left\|M_{\cdot\leftarrow 0}^{\mathbf{X}(\omega)}(y_0)\right\|_{p\text{-var};[0,T]}\right)^{(k+1)p}$$

for some non-random constants $C>0$. The conclusion that $U_{t\leftarrow 0}^{\mathbf{X}(\cdot)}(y_0) \in \cap_{p>1}\mathbb{D}^{k,p}(\mathbb{R}^e)$ follows at once from Theorems 7.2 and 7.4. □

8. Smoothness of the density: the proof of the main theorem

This section is devoted to the proof of our Hörmander type theorem 3.5. As mentioned in the introduction, apart from rather standard considerations concerning probabilistic proofs of Hörmander's theorem (see e.g. [21]), this boils down to the following steps:

(1) Let W be a smooth and bounded vector field in $\mathbb{R}^e$. Following [21], denote by $\left(Z_t^W\right)_{t\in[0,T]}$ the process

$$Z_t^W = J_{0\leftarrow t}^{\mathbf{X}} W\left(U_{t\leftarrow 0}^{\mathbf{X}}(y_0)\right). \tag{8.1}$$

Then assuming Conditions 2 and 3 we get a bound on $|Z^W|_\infty$ in terms of the Malliavin matrix C_T defined at (6.1). This will be the content of Proposition 8.4.

(2) We invoke iteratively our Norris lemma (Theorem 5.6) to processes like Z^W in order to generate enough upper bounds on Lie brackets of our driving vector fields at the origin.

In order to perform this second step, we first have to verify the assumptions of Theorem 5.6 for the process $M_{\cdot\leftarrow 0}^{\mathbf{x}}(y_0)$ defined by (7.13). Namely, we shall see that $M_{\cdot\leftarrow 0}^{\mathbf{x}}(y_0)$ is a process controlled by $\mathbf{X}$ in the sense of Definition 5.1 and relation (5.6).

Proposition 8.1. *Suppose $(X_t)_{t\in[0,T]}$ satisfies the condition of Theorem 7.2. In particular, X has a lift to $\mathbf{X}$, a geometric-p rough path for some $p>1$ which is in $C^{0,\gamma}([0,T];G^{\lfloor p\rfloor}(\mathbb{R}^d))$ for $\gamma = 1/p$. Then $M_{\cdot\leftarrow 0}^{\mathbf{x}}(y_0)$ is a process controlled by $\mathbf{X}$ in the sense of Definition 5.1 and*

$$\left\|M_{\cdot\leftarrow 0}^{\mathbf{X}}(y_0)\right\|_{\mathcal{Q}_{\mathbf{X}}^{\gamma}} \in \bigcap_{p\geq 1} L^p(\Omega).$$

Proof. For notational sake, the process $M_{\cdot\leftarrow 0}^{\mathbf{X}}(y_0)$ will be denoted by M only. It is readily checked that M is solution to a rough differential equation driven by $\mathbf{X}$, associated to the vector fields given by

$$F_i(y,J,K) = \left(V_i(y),\, \nabla V_i(y)\cdot J,\, -K\cdot\nabla V_i(y)\right), \quad i=1,\dots,d. \tag{8.2}$$

This equation can be solved either by genuine rough paths methods or within the landmark of algebraic integration. As mentioned in Proposition 5.5, both notions of solution coincide thanks to approximation procedures. This finishes the proof of our claim $M\in\mathcal{Q}_{\mathbf{X}}^{\gamma}$.

In order to prove integrability of M as an element of $\mathcal{Q}_{\mathbf{X}}^{\gamma}$, let us write the equation governing the dynamics of M under the form

$$dM_t = \sum_{i=1}^d F_i(M_t)\, d\mathbf{X}_t^i,$$

where $\mathbf{X}$ is our Gaussian rough path of order at most $N=3$. The expansion of M as a controlled process is simply given by the Euler scheme introduced in [12, Proposition 10.3]. More specifically, M admits a decomposition 5.4 of the form:

$$M_{s,t}^j = M_s^{j,i_1}\,\mathbf{X}_{s,t}^{1,i_1} + M_s^{j,i_1,i_2}\,\mathbf{X}_{s,t}^{2,i_1,i_2} + R_{s,t}^{j,M},$$

with

$$M_s^{j,i_1} = F_{i_1}^j(Z_s), \quad M_s^{j,i_1,i_2} = F_{i_2}F_{i_1}^j(Z_s), \quad |R_{s,t}^{j,M}| \le c_M|t-s|^{3\gamma}.$$

With the particular form (8.2) of the coefficient F and our assumptions on the vector fields V, it is thus readily checked that

$$\|M\|_{\mathcal{Q}_{\mathbf{X}}^\gamma} \le c_V\left(1 + \|J\|_\infty^2 + \|J^{-1}\|_\infty^2 + \|J\|_\gamma + \|U\|_\gamma\right),$$

and the right hand side of the latter relation admits moments of all order thanks to Theorem 7.2 and the remark which follows it. □

Define $\mathcal{L}_{\mathbf{x}}(y_0,\theta,T)$ to be the quantity

$$\mathcal{L}_{\mathbf{x}}(y_0,\theta,T) := 1 + L_\theta(x)^{-1} + |y_0| + \|M_{\cdot\leftarrow 0}^{\mathbf{x}}(y_0)\|_{\mathcal{Q}_{\mathbf{x}}^\gamma} + \mathcal{N}_{\mathbf{x},\gamma}$$

Corollary 8.2. *Under the assumptions of Proposition 8.1 we have*

$$\mathcal{L}_{\mathbf{x}}(y_0,\theta,T) \in \bigcap_{p\ge 1} L^p(\Omega).$$

Proof. We recall that the standing assumptions imply that $\|\mathbf{X}\|_{\gamma;[0,T]}$ has a Gaussian tail (see (2.8) from Section 2). It is easily deduce from this that

$$\mathcal{N}_{\mathbf{X},\gamma} \in \bigcap_{p\ge 1} L^p(\Omega).$$

Similarly we see from Corollary 5.8 and Proposition 8.1 that $L_\theta(x)^{-1}$ and $\|M_{\cdot\leftarrow 0}^{\mathbf{x}}(y_0)\|_{\mathcal{Q}_{\mathbf{x}}^\gamma}$ have moments of all orders and the claim follows. □

Definition 8.3. *We define the sets of vector fields $\mathcal{V}_k$ for $k \in \mathbb{N}$ inductively by*

$$\mathcal{V}_1 = \{V_i : i = 1,\dots,d\},$$

and then

$$\mathcal{V}_{n+1} = \{[V_i, W] : i = 0,1,\dots,d, W \in \mathcal{V}_n\}.$$

Proposition 8.4. *Let $(X_t)_{t\in[0,T]} = \left(X_t^1,\dots,X_t^d\right)_{t\in[0,T]}$ be a continuous Gaussian process, with i.i.d. components associated to the abstract Wiener space $(\mathcal{W},\mathcal{H},\mu)$. Assume that X satisfies the assumptions of Theorem 3.5. Then there exist real numbers p and θ satisfying $2/p > \theta > \alpha/2$ such that: (i) X is θ-Hölder rough and (ii) X has a natural lift to a geometric p rough path $\mathbf{X}$ in $C^{0,1/p}([0,T];G^{\lfloor p\rfloor}(\mathbb{R}^d))$. For $t \in (0,T]$ let*

$$C_t = \sum_{i=1}^d \int_{[0,t]^2} J_{t\leftarrow s}^{\mathbf{X}}(y_0)\,V_i(Y_s) \otimes J_{t\leftarrow s'}^{\mathbf{X}}(y_0)\,V_i(Y_{s'})\,dR(s,s'),$$

and suppose $k \in \mathbb{N}\cup\{0\}$. Then there exist constants $\mu = \mu(k) > 0$ and $C = C(t,k) > 0$ such that for all $W \in \mathcal{V}_k$ and all $v \in \mathbb{R}^e$ with $|v| = 1$, we have

$$\left|\langle v, Z_\cdot^W\rangle\right|_{\infty;[0,t]} \le C\mathcal{L}_{\mathbf{X}}(y_0,\theta,t)^\mu \left(v^T C_t v\right)^\mu. \tag{8.3}$$

Proof. Let us prove the first assertion. To do this, we note that the constraint on ρ implies that X lifts to a geometric p-rough path for any $p > 2\rho$. Because the ρ-variation is assumed to be Hölder-controlled, it follows that $\mathbf{X}$ is in $C^{0,1/p}([0,T];G^{\lfloor p\rfloor}(\mathbb{R}^d))$ by using (**??**). By assumption $\alpha < 2/\rho$, therefore we may always choose p close enough to 2ρ in order that

$$\frac{2}{p} > \frac{\alpha}{2}.$$

On the other hand X is θ-Hölder rough for any $\theta > \alpha/2$ by Corollary 5.8. Hence there always exist p and θ with the stated properties.

We have that

$$v^T C_t v = \sum_{i=1}^{d} \Lambda_t^i, \quad \text{with} \quad \Lambda_t^i \equiv \int_{[0,t]^2} f^i(s) f^i(s') \, dR(s,s'), \tag{8.4}$$

where we have set $f^i(s) := \left\langle v, J_{t\leftarrow s}^{\mathbf{X}}(y_0) V_i(y_s) \right\rangle = \left\langle v, Z_s^{V_i} \right\rangle$. Furthermore, because the hypotheses of Theorem 6.9 and Corollary 6.10 are satisfied, we can deduce that

$$\left|f^i\right|_{\infty;[0,t]} \leq 2 \max \left[\frac{\left|\Lambda_t^i\right|^{1/2}}{E\left[X_t^2\right]^{1/2}}, \frac{1}{\sqrt{c}} \left|\Lambda_t^i\right|^{\gamma/(2\gamma+\alpha)} \left|f^i\right|_{\gamma;[0,t]}^{\alpha/(2\gamma+\alpha)} \right], \tag{8.5}$$

for $i = 1, \ldots, d$. On the other hand Young's inequality for 2D integrals (see [14]) gives

$$\left|\Lambda_t^i\right| \lesssim \left[\left|f^i\right|_{\gamma;[0,t]} + \left|f^i(0)\right| \right]^2 V_\rho\left(R; [0,t]^2\right). \tag{8.6}$$

From (8.6), (8.5) and the relation $v^T C_t v = \sum_{i=1}^d \Lambda_t^i$ it follows that there exists some $C_1 > 0$, depending on t and c, such that we have

$$\left|f^i\right|_{\infty;[0,t]} \leq C_1 \left(v^T C_t v\right)^{\gamma/(2\gamma+\alpha)} \max_{i=1,\ldots,d} \left[\left|f^i(0)\right| + \left|f^i\right|_{\gamma;[0,t]} \right]^{\alpha/(2\gamma+\alpha)}.$$

Using the fact that for some $\nu > 0$

$$\left|f^i(0)\right| + \left|f^i\right|_{\gamma;[0,t]} \leq C_2 \mathcal{L}_{\mathbf{X}}(y_0, \theta, t)^\nu \text{ for } i = 1, .., d,$$

it is easy to deduce that (8.3) holds whenever $W \in \mathcal{V}_1$.

The proof of (8.3) for arbitrary $k \in \mathbb{N}$ now follows by induction. The key relation comes from observing that

$$\left\langle v, Z_u^W \right\rangle = \left\langle v, W(y_0) \right\rangle + \sum_{i=1}^{d} \int_0^u \left\langle v, Z_r^{[V_i,W]} \right\rangle dX_r^i,$$

in the sense of Proposition 5.5. Hence, assuming the induction hypothesis, we can use Theorem 5.6 to obtain a bound of the form (8.3) on

$$\left| \left\langle v, Z_\cdot^{[V_i,W]} \right\rangle \right|_{\infty;[0,t]}$$

for all $W \in \mathcal{V}_k$. Since $\mathcal{V}_{k+1} = \{[V_i, W] : i = 0, 1, \ldots, d, W \in \mathcal{V}_n\}$, the result is then established. □

We are now in a position to prove our main theorem. Since the structure of the argument is the classical one, we will minimise the amount of detail where possible.

Proof of Theorem 3.5. This involves assembling together the pieces we have developed in the paper. First let $2/p > \theta > \alpha/2$ be chosen such that X is θ-Hölder rough and X has a natural lift to a geometric p rough path $\mathbf{X}$ in $C^{0,1/p}([0,1]; G^{\lfloor p \rfloor}(\mathbb{R}^d))$. This is always possible by the first part of Proposition 8.4. Let $0 < t \leq T$ and note that we have shown in Proposition 7.5 that $U_{t\leftarrow 0}^{\mathbf{X}(\omega)}(y_0)$ is in $\mathbb{D}^\infty(\mathbb{R}^e)$. The result will therefore follow by showing that for every $q > 0$, there exists $c_1 = c_1(q)$ such that

$$P\left(\inf_{|v|=1} \langle v, C_t v \rangle < \epsilon\right) \leq c_1 \epsilon^q,$$

for all $\epsilon \in (0,1)$. It is classical that proving $(\det C_t)^{-1}$ has finite moments of all order is sufficient for $U_{t\leftarrow 0}^{\mathbf{X}(\omega)}(y_0)$ to have a smooth density (see, e.g., [30]).

Step 1: From Hörmander's condition there exists $N \in \mathbb{N}$ with the property that

$$span\left\{W(y_0) : W \in \cup_{i=1}^N \mathcal{V}_i\right\} = \mathbb{R}^e.$$

Consequently, we can deduce that

$$a := \inf_{|v|=1} \sum_{W \in \cup_{i=1}^N \mathcal{V}_i} |\langle v, W(y_0)\rangle| > 0. \tag{8.7}$$

For every $W \in \cup_{i=1}^N \mathcal{V}_i$ we have

$$\left|\left\langle v, Z_\cdot^W \right\rangle\right|_{\infty;[0,t]} \geq |\langle v, W(y_0)\rangle|, \tag{8.8}$$

and hence using (8.7), (8.8) and Proposition 8.4 we end up with

$$a \leq \inf_{|v|=1} \sup_{W \in \cup_{i=1}^N \mathcal{V}_i} \left|\left\langle v, Z_\cdot^W \right\rangle\right|_{\infty;[0,t]} \leq c_1 \mathcal{L}_{\mathbf{X}}(y_0, \theta, t)^\mu \inf_{|v|=1} \left|v^T C_t v\right|^\pi, \tag{8.9}$$

for some positive constants c_1, $\mu = \mu_N$ and $\pi = \pi_N$.

Step 2: From (8.9) can deduce that for any $\epsilon \in (0,1)$

$$P\left(\inf_{|v|=1} \left|v^T C_t v\right| < \epsilon\right) \leq P\left(\mathcal{L}_{\mathbf{X}}(y_0, \theta, t)^\mu > c_2 \epsilon^{-k}\right)$$

for some constants $c_2 > 0$ and $k > 0$ which do not depend on ϵ. It follows from Corollary 8.2 that for every $q > 0$ we have

$$P\left(\inf_{|v|=1} \left|v^T C_t v\right| < \epsilon\right) \leq c_3 \epsilon^{kq},$$

where $c_3 = c_3(q) > 0$ does not depend on ϵ. □

9. Appendix

Proof of Lemma 6.5. We prove the result for $S = T$, the modifications for $S < T$ are straight forward. Consider three nested sequences $(A_m)_{m=1}^\infty$, $(B_m)_{m=1}^\infty$ and $(C_m)_{m=1}^\infty$ consisting of partitions of $[0,s]$, $[s,t]$ and $[t,T]$ respectively, and suppose that the mesh of each sequence tends to zero as m tends to infinity. For each m_1 and m_2 in $\mathbb{N}$ let D^{m_1,m_2} denote the partition of $[0,T]$ defined by

$$D^{m_1,m_2} = A_{m_1} \cup B_{m_2} \cup C_{m_1}.$$

We now construct an increasing function $r : \mathbb{N} \to \mathbb{N}$ such that

$$(D_m)_{m=1}^\infty = \left(D^{r(m),m}\right)_{m=1}^\infty$$

together form a nested sequence of partitions of $[0,T]$ having the needed properties.

We do this inductively. First let $m = 1$, then for every two consecutive points $u < v$ in the partition B_m Lemma 6.4 implies that

$$\mathrm{Cov}\left(Z_{s,t} Z_{u,v} | \mathcal{F}^{A_n} \vee \mathcal{F}^{C_n}\right) \to \mathrm{Cov}\left(Z_{s,t} Z_{u,v} | \mathcal{F}_{s,t} \vee \mathcal{F}_{t,T}\right).$$

as $n \to \infty$. Z has positive conditional covariance, therefore the right hand side of the last expression is positive. This means we can choose $r(1)$ to ensure that

$$\mathrm{Cov}\left(Z_{s,t} Z_{u,v} | \mathcal{F}^{A_{r(1)}} \vee \mathcal{F}^{C_{r(1)}}\right) \geq 0, \tag{9.1}$$

for every two consecutive points u and v in B_m (the total number of such pairs does not depend on $r(1)$). We then let $D_1 = D^{r(1),1}$, both properties 2 and 3 in the statement are easy to check; the latter follows from (9.1), when we interpret the Schur complement as the covariance matrix of Z_2^1 conditional on Z_1^1 (see also the proof of Proposition 6.6). Having specified $r(1) < \ldots < r(k-1)$ we need only repeat the treatment outlined above by choosing some natural number $r(k) > r(k-1)$ to ensure that

$$\operatorname{Cov}\left(Z_{s,t}Z_{u,v}|\mathcal{F}^{A_{r(k)}} \vee \mathcal{F}^{C_{r(k)}}\right) \geq 0,$$

for each pair of consecutive points $u < v$ in B_k. It is easy to verify that $(D_m)_{m=1}^{\infty}$ constructed in this way has the properties we need. □

References

[1] Baudoin, F.; Hairer, M. A version of Hörmander's theorem for the fractional Brownian motion. Probab. Theory Related Fields 139 (2007), no. 3-4, 373–395.

[2] Baudoin, F., Ouyang C., Tindel ,S.: Upper bounds for the density of solutions of stochastic differential equations driven by fractional Brownian motions, arXiv:1104.3884, to appear in Ann. Inst. Henri Poincaré Probab. Stat.

[3] Berman, S. M.: Local nondeterminism and local times of Gaussian processes. Indiana Univ. Math. J. 23 (1973/74), 69–94

[4] Boyd, S.P, Vandenberghe, L. Convex optimization, Cambridge University Press, 2004.

[5] Cass, T., Litterer, C., Lyons, T.: Integrability Estimates for Gaussian Rough Differential Equations, arXiv:1104.1813

[6] Cass, T.; Friz, P.: Densities for rough differential equations under Hörmander's condition, Ann. of Math. (2), 171(3):2115–2141,201

[7] Cass, T.; Friz, P.; Victoir, N.: Non-degeneracy of Wiener functionals arising from rough differential equations. Trans. Amer. Math. Soc., 361(6):3359–3371, 2009

[8] Coutin, L., Qian, Z. : Stochastic analysis, rough path analysis and fractional Brownian motion Probab. Theory Related Fields 122 (2002), 108–140

[9] Cuzick, J.: Continuity of Gaussian local times. Ann. Probab. 10 (1982), no. 3, 818–823

[10] Cuzick, J. and DuPreez, J.P. : Joint Continuity of Gaussian Local Times, Ann. of Probab. , Vol. 10, No. 3 (1982), 810-817

[11] Friz, P., Oberhauser, H.: A generalized Fernique theorem and applications, Proceedings of the American Mathematical Society. 138, no. 10, 3679-3688, (2010)

[12] Friz, P., Victoir, N.: Multidimensional stochastic processes as rough paths. Theory and applications. Cambridge Studies in Advanced Mathematics, 120. Cambridge University Press, Cambridge, 2010

[13] Friz, P., Victoir, N.: A Variation Embedding Theorem and Applications, Journal of Functional Analysis, Vol 239 Nr 2 (2006) 631-637

[14] Friz, P., Victoir, N.: Differential equations driven by Gaussian signals. Ann. Inst. Henri Poincaré Probab. Stat. 46 (2010), no. 2, 369–41

[15] Friz, P., Victoir, N.: A note on higher dimensional p-variation arXiv:1102.4587v1

[16] Gubinelli M.: Controlling Rough Paths, Journal of Functional Analysis, 216 (2004), 86-140

[17] Gubinelli, M.: Ramification of rough paths, J. Differential Equations 248 (2010), 693-721.

[18] Hairer, M.: Malliavin's proof of Hörmander's theorem. Bull. Math. Sci. 135 (2011), no 6-7, pp. 650–666

[19] Hairer, M., Mattingly, J.: A theory of hypoellipticity and unique ergodicity for semilinear stochastic PDEs.

[20] Hairer, M., Pillai N.S.: Ergodicity of hypoelliptic SDEs driven by fractional Brownian motion, Ann. Inst. H. Poincaré Probab. Statist. Volume 47, Number 2 (2011), 601-628.

[21] Hairer, M., Pillai, N.: Regularity of Laws and Ergodicity of Hypoelliptic SDEs Driven by Rough Paths. Ann. Probab., to appear

[22] Hörmander, L.: Hypoelliptic second order differential equations. Acta Math. 119 (1967) 147–171.

[23] Hu, Y., Nualart, D.: Differential equations driven by Hölder continuous functions of order greater than 1/2., The Abel Symposium on Stochastic Analysis, 399-423. Springer, 2007.

[24] Hu, Y., Tindel, S.: Smooth density for some nilpotent rough differential equations, ArXiv 1104.1972, to appear in J. Theoretical Probab.

[25] Lyons, T.: Differential equations driven by rough signals, Rev. Mat. Iberoamericana 14 (1998), no. 2, 215–310.

[26] Lyons, T.; Caruana, M.; Levy, T.: Differential Equations Driven by Rough Paths, Differential Equations Driven by Rough Paths, Ecole d'Eté de Probabilités de Saint-Flour XXXIV-2004, Lecture Notes in Mathematics Vol. 1908, Springer, 2007
[27] Malliavin, P: Stochastic calculus of variation and hypoelliptic operators. Proceedings of the International Symposium on Stochastic Differential Equations (Res. Inst. Math. Sci., Kyoto Univ., Kyoto, 1976), pp. 195–263, Wiley, New York-Chichester-Brisbane, 1978
[28] Monrad, D. and Rootzen, H.: Small values of Gaussian processes and functional laws of the iterated logarithm, Probab. Theory Relat. Fields, 101,173-192 (1995)
[29] Norris, J.: Simplified Malliavin calculus. Séminaire de Probabilités, XX, 1984/85, 101–130, Lecture Notes in Math., 1204, Springer, Berlin, 1986
[30] Nualart, D: The Malliavin Calculus and Related Topics, Springer, 2006
[31] Sugita, H.: Sobolev spaces of Wiener functionals and Malliavin's calculus, J. Math. Kyoto Univ. Volume 25, Number 1 (1985), 31-48.
[32] Zhang, Fuzhen (Ed.): The Schur Complement and Its Applications, Series: Numerical Methods and Algorithms, Vol. 4, Springer 2005

Thomas Cass, Department of Mathematics, Imperial College London, The Huxley Building, 180 Queensgate, London.
E-mail address: thomas.cass@imperial.ac.uk

Martin Hairer, Mathematics Institute, University of Warwick, Coventry, CV4 7AL.
E-mail address: M.Hairer@Warwick.ac.uk

Christian Litterer, Department of Mathematics, Imperial College London, The Huxley Building, 180 Queensgate, London.

Samy Tindel, Institut Élie Cartan Nancy, Université de Lorraine, B.P. 239, 54506 Vandœuvre-lès-Nancy, France.
E-mail address: samy.tindel@univ-lorraine.fr