ANNALS OF MATHEMATICS

Densities for rough differential equations under Hörmander's condition

By Thomas Cass and Peter Friz

SECOND SERIES, VOL. 171, NO. 3

May, 2010

ANMAAH

Annals of Mathematics, **171** (2010), 2115–2141

Densities for rough differential equations under Hörmander's condition

By THOMAS CASS and PETER FRIZ

Abstract

We consider stochastic differential equations $dY = V(Y)\,dX$ driven by a multidimensional Gaussian process X in the rough path sense [T. Lyons, *Rev. Mat. Iberoamericana* **14**, (1998), 215–310]. Using Malliavin Calculus we show that Y_t admits a density for $t \in (0, T]$ provided (i) the vector fields $V = (V_1, \dots, V_d)$ satisfy Hörmander's condition and (ii) the Gaussian driving signal X satisfies certain conditions. Examples of driving signals include fractional Brownian motion with Hurst parameter $H > 1/4$, the Brownian bridge returning to zero after time T and the Ornstein-Uhlenbeck process.

1. Introduction

In the theory of stochastic processes, Hörmander's theorem on hypoellipticity of degenerate partial differential equations has always been an important tool to see if a diffusion process with a given generator admits a density. This dependence on PDE theory was removed when P. Malliavin devised a purely probabilistic approach to Hörmander's theorem, which is perfectly adapted to prove existence and smoothness of densities for diffusions given as strong solution to an Itô stochastic differential equation driven by Brownian motion.

The key ingredients of Malliavin's machinery, better known as *Malliavin Calculus* or *stochastic calculus of variations* can be formulated in the setting of an abstract Wiener space $(W, \mathcal{H}, \mu)$. This concept is standard (e.g. [27] or any modern book on stochastic analysis) as is the notion of *weakly nondegenerate* $\mathbb{R}^e$-valued functional φ which has the desirable property that the image measure $\varphi_* \mu$ is absolutely continuous with respect to Lebesgue measure on $\mathbb{R}^e$. (Functionals which are *nondegenerate* have a smooth density.) Precise definitions are given later on in the text.

Given these abstract tools, we turn to the standard Wiener space $C([0, T], \mathbb{R}^d)$ equipped with Wiener measure i.e. the standard model for Brownian motion $B =$

$B(\omega)$. From Itô's theory, we know how to solve the stochastic differential equation

$$dY = \sum_{i=1}^{d} V_i(Y) \circ dB^i \equiv V(Y) \circ dB, \ \ Y(0) = y_0 \in \mathbb{R}^e.$$

The Itô-map $B \mapsto Y$ is notorious for its lack of strong regularity properties. On the positive side, it is smooth in a weak Sobolev type sense ("smooth in Malliavin's sense") and under Hörmander's condition at $y_0 \in \mathbb{R}^e$

$$(H): \ \ \text{Lie}\,[V_1, \dots, V_d]\,|_{y_0} = \mathcal{T}_{y_0}\mathbb{R}^e \cong \mathbb{R}^e \tag{1.1}$$

one can show (e.g. [27], [32], [2], [29]) that the solution map $B \mapsto Y_t$ is nondegenerate for all $t \in (0, T]$. This line of reasoning provides a direct probabilistic approach to the study of transition densities of Y and has found applications from stochastic fluid dynamics to interest rate theory, e.g. [11], [19]. The same range of applications[1] nowadays demand stochastic models of type

$$dY = V(Y)\, dX \tag{1.2}$$

where X is a Gaussian process, such as fractional Brownian motion (short: fBm) popularized by [28]. Differential equations of this type have also been used as simple examples for the study of ergodicity of non-Markovian systems, [18]. (Although one can add a drift term of form $V_0(Y)\, dt$ in (1.2) our present focus is the driftless case.)

Let us remark that in the example of fBM with Hurst parameter $H > 1/2$, Kolmogorov's criterion shows that B has nice sample paths (more precisely, Hölder continuous sample paths of exponent greater $1/2$) which has the great advantage that (1.2) can be understood as integral equation for fixed ω based on Young integrals (i.e. limits of Riemann-Stieltjes sums). In this setting of nice sample paths, existence of a density was established in [30] assuming ellipticity. Using deterministic estimates for the Jacobian of the flow this density was then shown to be smooth [20]; building on the same estimates the Hörmander case was obtained in [1]. For $H \leq 1/2$ the situation appears to be fundamentally different: first, in view of Brownian (and worse) sample path regularity one needs Itô or rough path ideas to make sense of (1.2) for $H \leq 1/2$. Secondly, the proof of [1] does not extend to the rough path setting[2] and relies somewhat delicately on specific properties of fractional Brownian motion.

---

[1] For instance, stochastic differential equations driven by fBM have applications to vortex filaments; applications to finance include (geometric) fractional Brownian motion as paradigm of a non-semimartingale which admits no arbitrage under transaction costs. The reader is referred to the books [29, §5.3 and 5.4], [8, §8.1.] and the references therein.

[2] The estimates of [20] can be generalized [17] to (sharp) deterministic estimates on the Jacobian of RDE solutions giving L^p-estimates on the flow of RDEs driven by fBM if and only if $H > 1/2$. In particular, one sees that L^p-estimates on the flow of Stratonovich SDEs ($H = 1/2$) are fundamentally probabilistic, i.e., rely on cancellations in stochastic integration. At present, the question of how to

In a previous paper [3] we linked rough paths and Malliavin calculus by showing that existence of a density for solutions of (1.2) holds true under ellipticity; i.e.,

$$(E): \quad \operatorname{Span}[V_1, \dots, V_d]\,|_{y_0} = \mathcal{T}_{y_0}\mathbb{R}^e \cong \mathbb{R}^e$$

and generic nondegeneracy conditions on the multi-dimensional Gaussian process X, the differential equation (1.2) being understood in the rough path sense [23], [26], a unified framework which covers *at once* Young and Stratonovich solutions (and goes well beyond). The aim of this paper is to prove the existence of densities under Hörmander's condition (H) in the following form:

THEOREM 1. *Let* $(X_t^1, \dots, X_t^d) = (X_t : t \in [0,T])$ *be a continuous, centered Gaussian process with independent components* $X^1, \dots, X^d$. *Assume* X *satisfies the conditions listed in Section* 4. (*In particular,* X *is assumed to lift to a geometric rough path so that* (1.2) *makes sense as random rough differential equation.*) *Let* $V = (V_1, \dots, V_d)$ *a collection of smooth bounded vector fields on* $\mathbb{R}^e$ *with bounded derivatives which satisfies Hörmander's condition* (*H*) *at* y_0. *Then the random* RDE *solution* $Y_t = Y_t(\omega) \in \mathbb{R}^e$ *to* (1.2) *started at* $Y_0 = y_0$ *admits a density with respect to Lebesgue measure on* $\mathbb{R}^e$ *for all times* $t \in (0, T]$.

One should note that X, the Gaussian driving signal of (1.2), is fully described by the covariance function of each component and, under the further assumption of IID components, by the covariance of a single component, i.e., $R(s,t) = \mathbb{E}(X_s^1 X_t^1)$. In principle all conditions on X are checkable from the covariance, in practice it is convenient to have conditions available which involve the *reproducing kernel Hilbert* or *Cameron-Martin space* associated to X as well as certain sample path properties. Leaving these technical details to Section 4 we emphasize that our conditions are readily checked in many cases including fractional Brownian motion with Hurst parameter $H > 1/4$, the Brownian Bridge returning to zero after time T and the Ornstein-Uhlenbeck process; details are found in Section 8.

It may be helpful to note that whenever X is a semi-martingale on $[0,T]$ then (1.2) can be understood as Stratonovich stochastic differential equation, i.e.,

$$dY = \sum_{i=1}^{d} V_i\,(Y) \circ dX^i.$$

In such cases, rough path theory appears as intermediate tool that is neither needed to understand the assumptions nor the conclusions of Theorem 1. There may be cases when X can be written in terms of Brownian motion so that ultimately the techniques of [4], [33] are applicable. But in general Theorem 1 covers new grounds.

---

obtain good integrability when $H < 1/2$ is open although one suspects that Gaussian isoperimetry will ultimately play a role.

The proof of Theorem 1 is based on the fact [3] that RDE solutions driven by Gaussian signals are "$\mathcal{H}$-differentiable," i.e., differentiable in Cameron-Martin directions. Existence of a density is then reduced to showing that the Malliavin covariance matrix is weakly nondegenerate. The standard proof of this (e.g. [27], [2] or [29, §2.3.2]) is based on Blumenthal's 0-1 law and the Doob-Meyer decomposition for semi-martingales. The main difficulty to overcome in the general Gaussian context of this paper is that the Doob-Meyer decomposition is not available and we manage to bypass its use by suitable *small time developments for* RDE*s*, obtained in [16], in conjunction with (*Stroock-Varadhan type*) *support description for certain Gaussian rough paths* (as conjectured by Ledoux et al. [22] and carried out independently in [10], [14] with some definite statements obtained in [12].)

The crucial induction step — which explains the appearance of higher brackets — requires us to assume a "nonstandard" Hörmander condition which involves only iterated Lie-brackets contracted against certain tensors arising from free nilpotent Lie groups. Equivalence to the usual Hörmander condition (H) is then established separately.

2. Preliminaries on ODE and RDEs

2.1. *Controlled ordinary differential equations.* Consider the controlled ordinary differential equations, driven by a smooth R^d-valued signal $f = f(t)$ along sufficiently smooth and bounded vector fields $V = (V_1, \dots, V_d)$,

$$dy = V(y)df \equiv \sum_{i=1}^{d} V_i\,(y)\, f'\,(t)\, dt,\ y\,(t_0) = y_0 \in \mathbb{R}^e. \tag{2.1}$$

We call $U^f_{t\leftarrow t_0}(y_0) \equiv y_t$ the associated flow. Let J denote the Jacobian of U. It satisfies the ODE obtain by formal differentiation w.r.t. y_0. More specifically,

$$a \mapsto \left\{ \frac{d}{d\varepsilon} U^f_{t\leftarrow t_0}\,(y_0 + \varepsilon a) \right\}_{\varepsilon=0}$$

is a linear map from $\mathbb{R}^e \to \mathbb{R}^e$ and we let $J^f_{t\leftarrow t_0}(y_0)$ denote the corresponding $e \times e$ matrix. It is immediate to see that

$$\frac{d}{dt} J^f_{t\leftarrow t_0}\,(y_0) = \left[\frac{d}{dt} M^f \left(U^f_{t\leftarrow t_0}\,(y_0)\,, t \right) \right] \cdot J^f_{t\leftarrow t_0}\,(y_0)$$

where $\cdot$ denotes matrix multiplication and

$$\frac{d}{dt} M^f\,(y,t) = \sum_{i=1}^{d} V_i'\,(y)\, \frac{d}{dt} f^i_t\,.$$

Note that $J^f_{t_2 \leftarrow t_0} = J^f_{t_2 \leftarrow t_1} \cdot J^f_{t_1 \leftarrow t_0}$. We can also consider Gateaux derivatives in the driving signal and define

$$D_h U^f_{t \leftarrow 0} = \left\{ \frac{d}{d\varepsilon} U^{f+\varepsilon h}_{t \leftarrow 0} \right\}_{\varepsilon = 0}.$$

One sees that $D_h U^f_{t \leftarrow 0}$ satisfies a linear ODE and the variation of constants formula leads to

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

Finally, given a smooth vector field W a straight-forward computation gives

$$dJ^f_{0 \leftarrow t} \left(W \left(U^f_{t \leftarrow 0} \right) \right) = \sum_{i=1}^d J^f_{0 \leftarrow t} \left([V_i, W] \left(U^f_{t \leftarrow 0} \right) \right) df^i_t. \tag{2.2}$$

2.2. *Rough differential equations.* Following [23], [26], and [12], a geometric p-rough path $\mathbf{x}$ over $\mathbb{R}^d$ is a continuous path on $[0, T]$ with values in $G^{[p]}(\mathbb{R}^d)$, the step-$[p]$ nilpotent group over $\mathbb{R}^d$, and of finite p-variation relative to the [12] Carnot-Carathéodory metric d on $G^{[p]}(\mathbb{R}^d)$, i.e.,

$$\sup_{n \in \mathbb{N}} \sup_{0 < t_1 < \cdots < t_n < T} \sum_i d\left(\mathbf{x}_{t_i}, \mathbf{x}_{t_{i+1}} \right)^p < \infty.$$

As in [5], [23], we view $G^{[p]}(\mathbb{R}^d)$ as embedded in its enveloping tensor algebra; i.e.,

$$G^{[p]}(\mathbb{R}^d) \subset T^{[p]}(\mathbb{R}^d) := \oplus_{k=0,\ldots,[p]} (\mathbb{R}^d)^{\otimes k}.$$

One can then think of $\mathbf{x}$ as a path $x : [0, T] \to \mathbb{R}^d$ enhanced with its iterated integrals although the later need not make classical sense[3]. The canonical projection to $(\mathbb{R}^d)^{\otimes k}$ is denoted $\pi_k(\mathbf{x})$ or $\mathbf{x}^k$. Lyons' theory of rough paths then gives deterministic meaning to the rough differential equation (short: RDE)

$$dy = V(y)\, d\mathbf{x}. \tag{2.3}$$

(One can think of RDE solutions as limit points of corresponding ODEs of form (2.1) in which the smooth driving signals *plus their iterated integrals up to order* $[p]$ converge to $\mathbf{x}$ in suitable p-variation distance.) The motivating example, e.g., [23], [26], is that *almost every* continuous joint realization of Brownian motion and Lévy's area process (equivalently: iterated Stratonovich integrals) gives rise to a geometric p-rough path for $p > 2$, known as Brownian rough path or Enhanced

[3] In fact, $G^N(\mathbb{R}^d)$ can realized as all points in the tensor algebra which arise from computing iterated integrals up to order N of smooth paths over a fixed time interval. The group product then corresponds to the concatenation of paths, the inverse corresponds to running a path backwards in time etc.

Brownian motion (cf. §8.1) which provides in particular a robust path-by-path view of Stratonovich SDEs.

Back to the deterministic RDE (2.3) and assuming smoothness of the vector fields $V = (V_1, \dots, V_d)$, the solution induces a flow $y_0 \mapsto U^{\mathbf{x}}_{t \leftarrow t_0}(y_0)$. Following [24], [25], the Jacobian $J^{\mathbf{x}}_{t \leftarrow t_0}$ of the flow exists and satisfies a linear RDE, as does the directional derivative

$$D_h U^{\mathbf{x}}_{t \leftarrow 0} = \left\{ \frac{d}{d\varepsilon} U^{T_{\varepsilon h}\mathbf{x}}_{t \leftarrow 0} \right\}_{\varepsilon = 0}$$

for a smooth path h. If $\mathbf{x}$ arises from a smooth path x together with its iterated integrals the *translated rough path* $T_h \mathbf{x}$ (cf. [24], [26]) is nothing but $x + h$ together with its iterated integrals. In the general case, we assume $h \in C^{q\text{-var}}$ with $1/p + 1/q > 1$, the translation $T_h \mathbf{x}$ can be written in terms of $\mathbf{x}$ and cross-integrals between $\pi_1(\mathbf{x}_{0,\cdot}) =: x$ and the perturbation h. (These integrals are well-defined Young-integrals.)

PROPOSITION 1. *Let* $\mathbf{x}$ *be a geometric* p*-rough path over* $\mathbb{R}^d$ *and* $h \in C^{q\text{-var}}\left([0,T], \mathbb{R}^d\right)$ *such that* $1/p + 1/q > 1$. *Then*

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

where the right-hand side is well-defined as Young integral.

Proof. $J^{\mathbf{x}}_{t \leftarrow 0}, D_h U^{\mathbf{x}}_{t \leftarrow 0}$ satisfy (at least jointly with $U^{\mathbf{x}}_{t \leftarrow 0}$) RDEs driven by $\mathbf{x}$ which allows, in essence, to use Lyons' limit theorem; this is discussed in detail in [24], [25]. A little care is needed since the resulting vector fields are not bounded anymore. However we can rule out explosion and then localize the problem: the needed remark is that $J^{\mathbf{X}}_{t \leftarrow 0}$ also satisfy a linear RDE of form

$$dJ^{\mathbf{x}}_{t \leftarrow 0} = dM^{\mathbf{x}}\left(U^{\mathbf{x}}_{t \leftarrow 0}(y_0), t \right) \cdot J^{\mathbf{x}}_{t \leftarrow 0}(y_0)$$

and explosion can be ruled out by direct iterative expansion and estimates of the Einstein sum as in [23]. □

3. **RDEs driven by Gaussian signals**

We consider a continuous, centered Gaussian process $X = \left(X^1, \dots, X^d\right)$ with independent components started at zero. This gives rise to an abstract Wiener space $(W, \mathcal{H}, \mu)$ where $W = \bar{\mathcal{H}} \subset C_0\left([0,T], \mathbb{R}^d\right)$. Note that $\mathcal{H} = \oplus_{i=1}^d \mathcal{H}^{(i)}$ and recall that element of $\mathcal{H}$ are of form $h_t = \mathbb{E}(X_t \xi(h) t)$ where $\xi(h)$ is a Gaussian random variable. The ("reproducing kernel") Hilbert-structure on $\mathcal{H}$ is given by $\langle h, h' \rangle_{\mathcal{H}} := \mathbb{E}(\xi(h)\xi(h'))$.

Existence of a Gaussian geometric p-rough path above X is tantamount to the existence of certain Lévy area integrals. The case of fractional Brownian motion is well understood and several construction have been carried out [6], [26], [10], [31]. In particular, one requires $H > 1/4$ for the existence of stochastic areas (which can be defined as $L^2(\mathbb{P})$-limits as in Itô's theory). As a result, one has to deal with geometric p-rough paths for $p < 4$. (When $p < 2$ there is enough sample path regularity to use Young integration and we avoid speaking of rough paths.)

CONDITION 1. *Assume X admits a natural*[4] *lift to a* (*random*) *geometric p-rough path $\mathbf{X}$ and $\exists q : 1/p + 1/q > 1$ such that*

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

The example to have in mind is Brownian motion for which the above condition is satisfied with $p = 2 + \varepsilon$ and $q = 1$. (We shall say more about other Gaussian examples in §4.)

If $\mathbf{X} = \mathbf{B}^H$ denotes the geometric p-rough path, $p \in (1/H, [1/H] + 1)$, associated to fractional Brownian motion then it satisfies a *Stroock-Varadhan support description in rough path topology*. This was first conjectured by Ledoux et al. [22] (who obtained it for the Brownian rough path) and carried out independently in [10], [14] for $H > 1/3$. The more difficult[5] case of $H > 1/4$ is covered by a general support theorem for Gaussian rough path [12] of form

$$\operatorname{supp}(\mathbb{P}_* \mathbf{X}) = \overline{\{S_{[p]}(\mathcal{H})\}} \tag{3.1}$$

where support and closure are relative to the homogeneous p-variation topology for geometric p-rough paths. We recall that $S_{[p]}$, for $[p] = 2, 3$ given by

$$S_2 : h \mapsto 1 + \int_0^t dh + \int_0^t \int_0^s dh \otimes dh$$
$$S_3 : h \mapsto 1 + \int_0^t dh + \int_0^t \int_0^s dh \otimes dh + \int_0^t \int_0^s \int_0^r dh \otimes dh \otimes dh$$

lifts $\mathbb{R}^d$-valued paths canonically to $G^{[p]}\big(\mathbb{R}^d\big)$-valued paths. In [12] it is seen that $\mathbf{X}$ exists provided the covariance has finite ρ-variation with $\rho < 2$ and it is also established that $\mathcal{H} \hookrightarrow C^{\rho\text{-var}}$ which guarantees that $S_{[p]}(\mathcal{H})$ is well-defined via Young integration. Such support description will be important in checking Condition 5.

[4] In the sense of [12]. This implies, for instance, that $\mathbf{X}$ is the (p-variation) limit (in probability) of (lifted) piecewise linear approximations.

[5] The case $H > 1/3$ only involves stochastic area and can be handled by martingale arguments; for $H \in (1/4, 1/3]$ one has to deal with third iterated integrals and additional arguments are needed; cf [12].

Definition 1 ([21], [29, §4.1.3], [34, §3.3]). Given an abstract Wiener space $(W, \mathcal{H}, \mu)$, a random variable (i.e. measurable map) $F : W \to \mathbb{R}$ is continuously $\mathcal{H}$-differentiable, in symbols $F \in C^1_{\mathcal{H}}$, if for μ-almost every ω, the map

$$h \in \mathcal{H} \mapsto F(\omega + h)$$

is continuously Fréchet differentiable. A vector-valued r.v. $F = (F^1, \dots, F^e) : W \to \mathbb{R}^e$ is continuously $\mathcal{H}$-differentiable if and only if each F^i is continuously $\mathcal{H}$-differentiable. In particular, μ-almost surely, $DF(\omega) = (DF^1(\omega), \dots, DF^e(\omega)$ is a linear bounded map from $\mathcal{H} \to \mathbb{R}^e$. One then defines the Malliavin covariance matrix as the random matrix

$$\sigma(\omega) := \left(\langle DF^i, DF^j \rangle_{\mathcal{H}}\right)_{i,j=1,\dots,e} \in \mathbb{R}^{e \times e}.$$

We call F weakly nondegenerate if $\det(\sigma) \neq 0$ almost surely.

PROPOSITION 2. *Assume Condition* 1. *Then, for fixed* $t \geq 0$, *the* $\mathbb{R}^e$*-valued random variable*

$$\omega \mapsto U^{\mathbf{X}(\omega)}_{t \leftarrow 0}(y_0)$$

is continuously $\mathcal{H}$*-differentiable.*

Proof. Let us recall for $h \in \mathcal{H} \subset C^{q\text{-var}}$, the translation $T_h \mathbf{X}(\omega)$ can be written (for ω fixed!) in terms of $\mathbf{X}(\omega)$ and cross-integrals between $\pi_1\left(\mathbf{X}_{0,\cdot}\right) =: X \in C^{p\text{-var}}$ and h. (These integrals are well-defined Young-integrals.) Thanks to the definition of $\mathbf{X}(\omega)$ as the limit in probability of piecewise linear approximations to X and its iterated integrals (cf. [12]) and continuity properties of the translation operator (e.g. [6], essentially inherited from continuity properties of Young integrals) we see that the event

$$\{\omega : \mathbf{X}(\omega + h) \equiv T_h \mathbf{X}(\omega) \text{ for all } h \in \mathcal{H}\} \tag{3.2}$$

has probability one. We show that $h \in \mathcal{H} \mapsto U^{\mathbf{X}(\omega+h)}_{t \leftarrow 0}(y_0)$ is continuously Fréchet differentiable for every ω in the above set of full measure. By basic facts of Fréchet theory, we must show (a) Gateaux differentiability and (b) continuity of the Gateaux differential.

Ad (a). Using $\mathbf{X}(\omega + g + h) \equiv T_g T_h \mathbf{X}(\omega)$ for $g, h \in \mathcal{H}$ it suffices to show Gateaux differentiability of $U^{\mathbf{X}(\omega+\cdot)}_{t \leftarrow 0}(y_0)$ at $0 \in \mathcal{H}$. For fixed t, define

$$Z_{i,s} \equiv J^{\mathbf{X}}_{t \leftarrow s}\left(V_i\left(U^{\mathbf{X}}_{s \leftarrow 0}\right)\right).$$

Note that $s \mapsto Z_{i,s}$ is of finite p-variation. We have, with implicit summation over i,

$$
\begin{aligned}
\left|D_h U_{t\leftarrow 0}^{\mathbf{X}}(y_0)\right| &= \left|\int_0^t J_{t\leftarrow s}^{\mathbf{X}}\left(V_i\left(U_{s\leftarrow 0}^{\mathbf{X}}\right)\right) dh_s^i\right| \\
&= \left|\int_0^t Z_i\, dh^i\right| \\
&\le c\left(|Z|_{p\text{-var}} + |Z(0)|\right) \times |h|_{\rho\text{-var}} \\
&\le c\left(|Z|_{p\text{-var}} + |Z(0)|\right) \times |h|_{\mathcal{H}},
\end{aligned}
$$

where c varies from line to line. Hence, the linear map

$$
DU_{t\leftarrow 0}^{\mathbf{X}}(y_0) : h \mapsto D_h U_{t\leftarrow 0}^{\mathbf{X}}(y_0) \in \mathbb{R}^e
$$

is bounded and each component is an element of $\mathcal{H}^*$. We just showed that

$$
h \mapsto \left\{\frac{d}{d\varepsilon} U_{t\leftarrow 0}^{T_{\varepsilon h}\mathbf{X}(\omega)}(y_0)\right\}_{\varepsilon=0} = \left\langle DU_{t\leftarrow 0}^{\mathbf{X}(\omega)}(y_0), h\right\rangle_{\mathcal{H}}
$$

and hence

$$
h \mapsto \left\{\frac{d}{d\varepsilon} U_{t\leftarrow 0}^{\mathbf{X}(\omega+\varepsilon h)}(y_0)\right\}_{\varepsilon=0} = \left\langle DU_{t\leftarrow 0}^{\mathbf{X}(\omega)}(y_0), h\right\rangle_{\mathcal{H}}
$$

emphasizing again that $\mathbf{X}(\omega + h) \equiv T_h\mathbf{X}(\omega)$ almost surely for all $h \in \mathcal{H}$ simultaneously. Repeating the argument with $T_g\mathbf{X}(\omega) = \mathbf{X}(\omega + g)$ shows that the Gateaux differential of $U_{t\leftarrow 0}^{\mathbf{X}(\omega+\cdot)}$ at $g \in \mathcal{H}$ is given by

$$
DU_{t\leftarrow 0}^{\mathbf{X}(\omega+g)} = DU_{t\leftarrow 0}^{T_g\mathbf{X}(\omega)}.
$$

(b) It remains to be seen that $g \in \mathcal{H} \mapsto DU_{t\leftarrow 0}^{T_g\mathbf{X}(\omega)} \in L(\mathcal{H}, \mathbb{R}^e)$, the space of linear bounded maps equipped with operator norm, is continuous. To this end, assume $g_n \to_{n\to\infty} g$ in $\mathcal{H}$ (and hence in $C^{\rho\text{-var}}$). Continuity properties of the Young integral imply continuity of the translation operator viewed as map $h \in C^{\rho\text{-var}} \mapsto T_h\mathbf{X}(\omega)$ as p-rough path (see [26]) and so

$$
T_{g_n}\mathbf{X}(\omega) \to T_g\mathbf{X}(\omega)
$$

in p-variation rough path metric. To point here is that

$$
\mathbf{x} \mapsto J_{t\leftarrow \cdot}^{\mathbf{x}} \text{ and } J_{t\leftarrow \cdot}^{\mathbf{x}}\left(V_i\left(U_{\cdot\leftarrow 0}^{\mathbf{x}}\right)\right) \in C^{p\text{-var}}
$$

depends continuously on $\mathbf{x}$ with respect to p-variation rough path metric: using the fact that $J_{t\leftarrow \cdot}^{\mathbf{x}}$ and $U_{\cdot\leftarrow 0}^{\mathbf{x}}$ both satisfy rough differential equations driven by $\mathbf{x}$ this is just a consequence of Lyons' limit theorem (the *universal limit theorem* of rough path theory). We apply this with $\mathbf{x} = \mathbf{X}(\omega)$ where ω remains a fixed element in (3.2). It follows that

$$
\left\| DU_{t\leftarrow 0}^{T_{g_n}\mathbf{X}(\omega)} - DU_{t\leftarrow 0}^{T_g\mathbf{X}(\omega)} \right\|_{\mathrm{op}} = \sup_{h : |h|_{\mathcal{H}}=1} \left| D_h U_{t\leftarrow 0}^{T_{g_n}\mathbf{X}(\omega)} - D_h U_{t\leftarrow 0}^{T_g\mathbf{X}(\omega)} \right|
$$

and defining $Z_i^g(s) \equiv J_{t\leftarrow s}^{T_g\mathbf{X}(\omega)}\left(V_i\left(U_{s\leftarrow 0}^{T_g\mathbf{X}(\omega)}\right)\right)$, and similarly $Z_i^{g_n}(s)$, the same reasoning as in part (a) leads to the estimate

$$\left\| DU_{t\leftarrow 0}^{T_{g_n}\mathbf{X}(\omega)} - DU_{t\leftarrow 0}^{T_g\mathbf{X}(\omega)} \right\|_{\mathrm{op}} \leq c\left(|Z^{g_n} - Z^g|_{p\text{-var}} + |Z^{g_n}(0) - Z^g(0)|\right).$$

From the explanations just given this tends to zero as $n \to \infty$ which establishes continuity of the Gateaux differential, as required, and the proof is finished. □

4. **Conditions on driving process**

We now give a complete list of assumptions on the (d-dimensional) Gaussian driving signal $(X_t : t \in [0,T])$. The first condition was already needed in the previous section to show $\mathcal{H}$-differentiability of RDE solutions driven by X; we repeat it for completeness and to give some additional examples.

CONDITION 2. *Assume that X admits a natural lift to a (random) geometric p-rough path $\mathbf{X}$ and $\exists q : 1/p + 1/q > 1$ such that*

$$\mathcal{H} \hookrightarrow C^{q\text{-var}}([0,T], \mathbb{R}^d). \tag{4.1}$$

At the price of a deterministic time-change (cf. following remark) we can and will assume that the p-variation of $\mathbf{X}$ is controlled by a (random) constant times $\omega(s,t) = t - s$, i.e.

$$\|\mathbf{X}_{s,t}\|^p = d(\mathbf{X}_s, \mathbf{X}_t)^p = (\textit{random const}) \times \omega(s,t). \tag{4.2}$$

Remark 1. The natural lift constructed in [12, Prop. 16 applied with $\rho = 1$] satisfies an estimate of form

$$|d(\mathbf{X}_s, \mathbf{X}_t)|_{L^q(\mathbb{P})} \lesssim q^{1/2} |R|^{1/2}_{\rho\text{-var};[s,t]}, \ \rho \in [1,2)$$

from which finite p-variation ($p > 2\rho$) is readily deduced. Using $|R|^{\rho}_{\rho\text{-var};[s,t]} \leq |R|^{\rho}_{\rho\text{-var};[0,t]} - |R|^{\rho}_{\rho\text{-var};[0,s]}$ we can define $\mathbf{X}'(\cdot)$ by requiring that

$$\mathbf{X}(t) \equiv \mathbf{X}'\left(|R|^{\rho}_{\rho\text{-var};[0,t]}\right)$$

so that $|d(\mathbf{X}'_s, \mathbf{X}'_t)|_{L^q(\mathbb{P})} \lesssim q^{1/2}|t-s|^{1/(2\rho)}$. This implies (4.2) for $\mathbf{X}'$.

In the Brownian motion case this holds, as already remarked earlier, with $p = 2 + \varepsilon$ and $q = 1$. The same is true for the Brownian bridge and the Ornstein-Uhlenbeck examples discussed in the introduction; although case-by-case verifications are not difficult, there is general criterion on the covariance which implies (4.1), see [12, Prop. 16 applied with $\rho = 1$], which also covers fBM. (In all these examples, (4.2) is satisfied without need of time-change.) Let us give a direct argument for case of fBM which covers any Hurst parameter $H > 1/4$. Writing

$\mathcal{H}^H$ for the Cameron-Martin space of fBM, the variation embedding in [15] gives

$$\mathcal{H}^H \hookrightarrow C^{q\text{-var}} \text{ for any } q > (H + 1/2)^{-1}.$$

At the same time [6], [26], [10], [31] fBM lifts to a geometric p-rough path for $p > 1/H$. By choosing p, q small enough $1/p + 1/q$ can be made arbitrarily close to $H + (H + 1/2) = 2H + 1/2 > 1$ and so (4.1) holds indeed for fBM with Hurst parameter $H > 1/4$.

CONDITION 3. *Fix $T > 0$. We assume nondegeneracy on $[0, T]$ in the sense that for any $f = (f_1, \dots, f_d) : [0, T] \to \mathbb{R}^d$ of finite p-variation*[6]

$$\left(\int_0^T f\,dh \equiv \sum_{j=1}^d \int_0^T f_j\,dh^j = 0 \forall h \in \mathcal{H}\right) \implies f \equiv 0 \text{ a.e.} \tag{4.3}$$

Fractional Brownian motion, for instance, satisfies this nondegeneracy condition simply because $C_0^1\left([0, T], \mathbb{R}^d\right) \subset \mathcal{H}^H$, cf. [14], which implies that any such f is orthogonal to a dense subset of $L^2[0, T]$ and hence 0 almost everywhere on $[0, T]$. Similar reasoning shows that an Ornstein-Uhlenbeck process, or a Brownian bridge which returns to zero *after* time T, satisfies Condition 3; while a Brownian bridge which returns to zero *at* time T is ruled out. Condition 3 already appeared in [3] which is also where the reader can find some further remarks and ramifications. Let us just note that (i) nondegeneracy on $[0, T]$ implies nondegeneracy on $[0, t]$ for any $t \in (0, T]$, (ii) for continuous f the conclusion reads $f \equiv 0$, and (iii) that the quantifies $\forall h \in \mathcal{H}$ can be relaxed to the quantifier "for all h in some orthonormal basis of $\mathcal{H}$", as is easily seen by continuity of the Young-integral $\int_0^T f\,dh$ with respect to $h \in \mathcal{H} \hookrightarrow C^{q\text{-var}}$.

CONDITION 4 ("0-1 law"). *The germ σ-algebra $\cap_{t>0}\sigma(X_s : s \in [0, t])$ contains only events of probability zero or one.*

When X is Brownian motion, this is the well-known Blumenthal zero-one law. More generally, it holds whenever X is an adapted functional of Brownian motion, including all examples (such as fBM) in which X has a Volterra presentation [7]

$$X_t = \int_0^t K(t, s)\,dB_s.$$

(Nothing is assumed on K other than having the above Wiener-Itô integral well-defined.) The 0-1 law also holds when X is the strong solution of an SDE driven by Brownian motion; this includes the Ornstein-Uhlenbeck and Brownian bridge examples. An example where the 0-1 law fails is given by the *random-ray* $X : t \mapsto$

[6]This guarantees, together with Condition 2, that the integral in (4.3) is well-defined in Young sense.

$tB_T(\omega)$ in which case the germ-event $\{\omega : dX_t(\omega)/dt|_{t=0+} \geq 0\}$ has probability $1/2$. (In fact, sample path differentiability at $0+$ implies nontriviality of the germ σ-algebra see [9] and references therein.) We observe that the random ray example is (a) already ruled out by Condition 3 and (b) should be ruled out anyway since it does not trigger to the bracket phenomenon needed for a Hörmander statement.

The next condition expresses some sort of scaled support statement at $t = 0+$ and is precisely what is needed in the last part (Step 4) in the proof of Theorem 1 below. We give examples and easier-to-check conditions below. To state it, we recall [23, Th. 2.2.1] that a geometric p-rough path $\mathbf{x}$ lifts uniquely lifts uniquely and continuously (with respect to homogeneous p-variation distances) to a path in the free step-N nilpotent group[7], say

$$S_N(\mathbf{x}) \in C_0^{p\text{-var}}\left([0,T], G^N\left(\mathbb{R}^d\right)\right) \text{ for } N \geq [p].$$

We also recall that $G^N\left(\mathbb{R}^d\right)$ carries a dilation operator δ which generalizes scalar multiplication on $\mathbb{R}^d$.

CONDITION 5. *Assume there exists* $H \in (0,1)$ *such that* (i) *for all fixed* $N \geq [p]$, *writing* $\tilde{\mathbf{X}} = S_N(\mathbf{X})$, *all* $g \in G^N\left(\mathbb{R}^d\right)$ *and for all* $\varepsilon > 0$,

$$\liminf_{n\to\infty} \mathbb{P}\left(d\left(\delta_{n^H}\tilde{\mathbf{X}}_{1/n}, g\right) < \varepsilon\right) > 0.$$

(ii) $Hp < 1 + \frac{1}{r}$ *where* $r \in \mathbb{N}$ *is such that in Hörmander's condition* (1.1), *full span is achieved with* $\{V_1, \ldots, V_d\}|_{y_0}$ *and bracket vector fields* $[V_{i_1},[V_{i_2},\ldots]]|_{y_0}$ *involving up to* r *brackets.*

PROPOSITION 3. *Let* B *denote* d*-dimensional fractional Brownian motion with fixed Hurst parameter* $H \in (1/4, 1)$ *and consider the lift to a* (*random*) *geometric* p*-rough path, denoted by* $\mathbf{X} = \mathbf{B}$, *with* $p < 4$. *Then it satisfies Condition* 5.

Remark 2. Brownian motion is covered with $H = 1/2$.

Proof. Write $\tilde{\mathbf{B}} = S_N(\mathbf{B})$. From Section 3, and the references therein, the support of the law of $\mathbf{B}$ w.r.t. homogeneous p-variation distance is

$$C_0^{0,p\text{-var}}\left([0,T], G^{[p]}\left(\mathbb{R}^d\right)\right),$$

that is, the closure of lifted smooth path started at 0 with respect to homogeneous p-variation distance [23], [12]. By continuity of S_N [23, Th. 2.2.1] followed by evaluation of the path at time 1 it follows that the support of the law of $\tilde{\mathbf{B}}_1$ is full, that is, equal to $G^N\left(\mathbb{R}^d\right)$. On the other hand, fractional scaling

$$\left(n^H B_{t/n} : t \geq 0\right) \overset{D}{=} \left(B_t : t \geq 0\right)$$

[7]The 0 in $C_0^{p\text{-var}}$ indicates that $\mathbf{X}_0$ is started at the unit element in the group.

implies $\delta_{n^H}\tilde{\mathbf{B}}_{1/n} \overset{D}{=} \tilde{\mathbf{B}}_1$ and so, thanks to full support of $\tilde{\mathbf{B}}_1$,

$$\liminf_{n\to\infty} \mathbb{P}\left(d\left(\delta_{n^H}\tilde{\mathbf{B}}_{1/n}, g\right) < \varepsilon\right) = \mathbb{P}\left(d\left(\tilde{\mathbf{B}}_1, g\right) < \varepsilon\right) > 0.$$

Part (ii) in Condition 5 is harmless to check: an immediate application of Kolmogorov's criterion shows that $1/p$, the Hölder-exponent of fractional Brownian sample paths, can be taken to be anything strictly less than H. In particular, this allows us to choose $p(> 1/H)$ small enough such that $1 < Hp \leq 1 + 1/r$. □

Although scaling was important in the previous proof, it is only used at times near $0+$. One thus suspects that every other Gaussian signal X which scales similarly (on the level of N^{th} iterated integrals!) also satisfies Condition 5. To make this precise we need

THEOREM 2 ([12]). *Let* $(X, Y) = \left(X^1, Y^1, \ldots, X^d, Y^d\right)$ *be a centered continuous Gaussian process on* $[0, 1]$ *such that* $\left(X^i, Y^i\right)$ *are independent for* $i = 1, \ldots, d$. *Let* $\rho \in [1, 2)$ *and assume the covariance of* (X, Y), *as function on* $[0, 1]^2$, *is of finite* ρ-*variation* (*in* 2D *sense*[8]). *Then, for every* $p > 2\rho$, X *and* Y *can be lifted to geometric* p-*rough paths denoted* $\mathbf{X}$ *and* $\mathbf{Y}$. *Moreover, there exists positive constants* $\theta = \theta(p, \rho)$ *and* $C = C(p, \rho, K)$ *with* $\left|R_{(X,Y)}\right|_{\rho\text{-var};[0,1]^2} \leq K$ *so that for all* $q \in [1, \infty)$,

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

(Note that $R_{X-Y}(s, t)$ is a diagonal matrix with entries depending on s, t.)

COROLLARY 1. *Let* (X, B) *satisfy the conditions of the previous theorem and assume that* B *is a* (d-*dimensional*) *fractional Brownian motion with fixed Hurst parameter* $H \in (1/4, 1)$. *Assume in addition that*

$$n^{2H}\,|R_{X-B}|_{\infty;[0,1/n]^2} \to 0. \tag{4.4}$$

Then Condition 5 *holds.*

Proof. With focus on one diagonal entry and with mild abuse of notation (writing X, B instead of $X^i, B^i \ldots$)

$$n^{2H}\,|R_{X-B}|_{\infty;[0,1/n]^2} = \sup_{s,t\in[0,1]} \mathbb{E}\left[n^H\left(X_{s/n} - B_{s/n}\right)n^H\left(X_{t/n} - B_{t/n}\right)\right]$$

which can be rewritten in terms of the rescaled process $X^{(n)} = n^H X_{\cdot/n}$, and similarly for B, as

$$\sup_{s,t\in[0,1]} \mathbb{E}\left[\left(X_s^{(n)} - B_s^{(n)}\right)\left(X_t^{(n)} - B_t^{(n)}\right)\right] = |R_{X^{(n)}-B^{(n)}}|_{\infty;[0,1]^2}.$$

[8]Given a function f from $[0, 1]^2$ into some normed space, its variation (in the 2D sense!) is an immediate generalization of the standard definition but based on "rectangular increments" over $[a, b) \times [c, d)$ of the form $f(b, d) + f(a, c) - f(a, d) - f(b, c)$.

By assumption and the previous theorem, this entails that

$$d_{p\text{-var}}\left(\mathbf{X}^{(n)}, \mathbf{B}^{(n)}\right) \to 0 \text{ in probability.}$$

By continuity of S_N, still writing $\tilde{\mathbf{X}}^{(n)} = S_N\left(\mathbf{X}^{(n)}\right)$ for fixed N, and similarly for $\mathbf{B}^{(n)}$, we have

$$d\left(\tilde{\mathbf{X}}_1^{(n)}, \tilde{\mathbf{B}}_1^{(n)}\right) \le d_{p\text{-var};[0,1]}\left(\tilde{\mathbf{X}}^{(n)}, \tilde{\mathbf{B}}^{(n)}\right) \to 0 \text{ in probability.}$$

But then

$$\begin{aligned}
\mathbb{P}\left(d\left(\delta_{n^H}\tilde{\mathbf{X}}_{1/n}, g\right) < \varepsilon\right) &= \mathbb{P}\left(d\left(\tilde{\mathbf{X}}_1^{(n)}, g\right) < \varepsilon\right) \\
&\ge \mathbb{P}\left(d\left(\tilde{\mathbf{X}}_1^{(n)}, \tilde{\mathbf{B}}_1^{(n)}\right) + d\left(\tilde{\mathbf{B}}_1^{(n)}, g\right) < \varepsilon\right) \\
&\ge \mathbb{P}\left(d\left(\tilde{\mathbf{B}}_1^{(n)}, g\right) < \varepsilon/2\right) - \mathbb{P}\left(d\left(\tilde{\mathbf{X}}_1^{(n)}, \tilde{\mathbf{B}}_1^{(n)}\right) > \varepsilon/2\right)
\end{aligned}$$

and so

$$\liminf_{n\to\infty} \mathbb{P}\left(d\left(\delta_{n^H}\tilde{\mathbf{X}}_{1/n}, g\right) < \varepsilon\right) \ge \liminf_{n\to\infty} \mathbb{P}\left(d\left(\tilde{\mathbf{B}}_1^{(n)}, g\right) < \varepsilon/2\right)$$

and this is positive by the example in which we discussed the case of $\mathbf{B}$ resp. $\tilde{\mathbf{B}}$. The proof is finished. $\square$

5. Taylor expansions for rough differential equations

Given a smooth vector field W and smooth driving signal $x(\cdot)$ for the ODE $dy = V(y)\,dx$, it follows from (2.2) that

$$J_{0\leftarrow t}^{x}\left(W\left(y_t^x\right)\right) = W(y_0) + \int_0^t J_{0\leftarrow s}^{x}\left([V_i, W]\left(y_s^x\right)\right) dx_s^i,$$

where Einstein's summation convention is used throughout. Iterated use of this leads to the Taylor expansion

$$\begin{aligned}
J_{0\leftarrow t}^{x}\left(W\left(y_t^x\right)\right) = W|_{y_0} &+ [V_i, W]\,|_{y_0}\mathbf{x}_{0,t}^{1;i} \\
&+ \left[V_i, \left[V_j, W\right]\right]|_{y_0}\mathbf{x}_{0,t}^{2;i,j} \\
&+ \dots \\
&+ [V_{i_1}, \dots [V_{i_N}, W]]\,|_{y_0}\mathbf{x}_{0,t}^{N;i_1,\dots,i_N} \\
&+ \dots\ ,
\end{aligned}$$

where $\mathbf{x}_{0,t}$ denotes the signature of $x(\cdot)\,|_{[0,t]}$ in $\mathbb{R}^d \oplus \left(\mathbb{R}^d\right)^{\otimes 2} \oplus \dots \oplus \left(\mathbb{R}^d\right)^{\otimes N}$ $\oplus \dots$ (Note that such an expansion makes immediate sense when x is replaced by a weak geometric p-rough path $\mathbf{x}$.)[9] It will be convenient to express $J^x_{0\leftarrow t}\left(W\left(y^x_t\right)\right)$ as solution of some ODE of form $dz = \hat{V}(z)dx$. This is accomplished by setting

$$z := \left(z^1, z^2, z^3\right) := \left(y^x, J^x_{0\leftarrow t}, J^x_{0\leftarrow t}\left(W\left(y^x_t\right)\right)\right) \in \mathbb{R}^e \oplus \mathbb{R}^{e\times e} \oplus \mathbb{R}^e.$$

Noting that $J^x_{0\leftarrow t}\left(W\left(y^x_t\right)\right)$ is given by $z^2 \cdot W\left(z^1\right)$ in terms of matrix multiplication we have

$$\begin{aligned}
dz^1 &= V_i\left(z^1\right)dx^i\\
dz^2 &= -z^2 \cdot DV_i\left(z^1\right)dx^i\\
dz^3 &= \left(dz^2\right)\cdot W\left(z^1\right) + z^2 \cdot d\left(W\left(z^1\right)\right)\\
&= z^2 \cdot \left(-DV_i\left(z^1\right)\cdot W\left(z^1\right) + DW\left(z^1\right)\cdot V_i\left(z^1\right)\right)dx^i\\
&= z^2 \cdot [V_i, W]\,|_{z^1}\, dx^i
\end{aligned}$$

started from $(y_0, I, W(y_0))$ where I denotes the identity matrix in $\mathbb{R}^{e\times e}$ and we see that $\hat{V}$ is given by

$$\hat{V}_i\left(z^1, z^2, z^3\right) = \begin{pmatrix} V_i\left(z^1\right) \\ -z^2 \cdot DV_i\left(z^1\right) \\ z^2 \cdot [V_i, W](z_1) \end{pmatrix}, \quad i = 1, \dots, d. \tag{5.1}$$

Let us now consider the corresponding rough differential equation, $dz = \hat{V}(z)d\mathbf{x}$ where x is weak geometric p-rough path.

LEMMA 1. *Assume $V_1, \dots, V_d, W$ are smooth vector fields, bounded with all derivatives bounded. Then $\hat{V} = \left(\hat{V}_1, \dots, \hat{V}_d\right)$ is a collection of smooth (possibly unbounded) vector fields but explosion does not occur. More precisely, there exists a unique* RDE *solution to $dz = \hat{V}(z)d\mathbf{x}$ on any compact time interval* $[0, T]$. *In fact, for some increasing function φ from $\mathbb{R}^+$ into itself*

$$|z|_{\infty;[0,t]} \le \varphi(M) \text{ when } \|\mathbf{x}\|_{p\text{-var};[0,t]} \le M.$$

Proof. Smoothness of $\hat{V}$ is obvious and so the RDE $dz = \hat{V}(z)d\mathbf{x}$ has a solution up to some possible explosion time. From the particular structure of $\hat{V}$ we now argue that explosion cannot occur in finite time: z^1 does not explode as it is a genuine RDE solution along bounded vector fields with bounded derivatives of all orders (in fact, $\mathrm{Lip}^{\gamma-1}$ in the sense of Stein, $\gamma > p$ would be sufficient for nonexplosion of z^1).

Secondly, z^2 does not explode as it satisfies a linear RDE (cf. [23]) driven by some rough path $M^{\mathbf{x}}$ as already remarked in the proof of Proposition 1). Clearly then, $z^3 = z^2 \cdot W\left(z^1\right)$ where W is a bounded vector fields cannot explode. More

[9] By definition, such a p-rough path takes values in the step-$[p]$ tensor algebra but recall that there is a unique lift to the step-N group for any $N > [p]$.

precisely, using the estimates for RDE solutions driven along Lip- respectively linear vector fields (see [16] and [23]) it is clear that z remains in a ball of radius only depending on M if $\|\mathbf{x}\|_{p\text{-var};[0,t]} \leq M$. □

Let us make the following definitions: given $(m-1)$-times differentiable vector fields $V = (V_1, \ldots, V_d)$ on $\mathbb{R}^e$, $\mathbf{g} \in \oplus_{k=0}^{m} (\mathbb{R}^d)^{\otimes k}$ and $y \in \mathbb{R}^e$ we write

$$\mathcal{E}_{(V)}(y, \mathbf{g}) := \sum_{k=1}^{m} \sum_{\substack{i_1,\ldots,i_k \\ \in \{1,\ldots,d\}}} \mathbf{g}^{k,i_1,\ldots,i_k} V_{i_1} \ldots V_{i_k} I(y).$$

(Here I denote the identity function on $\mathbb{R}^e$ and vector fields identified with first order differential operators.) In a similar spirit, given another sufficiently smooth vector field W we first write

$$[V_{i_1}, V_{i_2}, \ldots V_{i_k}, W] := [V_{i_1}, [V_{i_2}, \ldots [V_{i_k}, W] \ldots]]$$

and then

$$\mathbf{g}^k \cdot [V, \ldots, V, W]|_{y_0} := \sum_{\substack{i_1,\ldots,i_k \\ \in \{1,\ldots,d\}}} \mathbf{g}^{k,i_1,\ldots,i_k} [V_{i_1}, V_{i_2}, \ldots V_{i_k}, W] I(y_0) \tag{5.2}$$

with the convention that $\mathbf{g}^0 \cdot V_k = V_k$. The following result is an Euler-estimate for the rough differential equation $dy = V(y)\, d\mathbf{x}$; a special case of [16, Th. 19] and included for the reader's convenience.

THEOREM 3 (Euler estimate for RDEs). *Let $p \geq 1$ and fix* Lip$^\infty$*-vector fields $V_1, \ldots, V_d$ on $\mathbb{R}^e$. Let $\mathbf{x}$ be a weak geometric p-rough path whose p-variation is controlled by $\omega(s,t) = t - s$ so that*

$$\text{for all } 0 \leq s \leq t \leq 1 : \|\mathbf{x}_{s,t}\|^p \leq M\omega(s,t).$$

Then, for any integer $m > p - 1$ and all $0 \leq s \leq t \leq 1$,

$$\left| y_{s,t} - \mathcal{E}_{(V)}\left(y_s, S_m(\mathbf{x})_{s,t}\right)\right| \leq C\omega(s,t)^\theta \text{ with } \theta = \frac{m+1}{p} > 1$$

where $C = C(M)$, a constant which may also depend on N, p, y_0 and $V_1, \ldots, V_d$.

COROLLARY 2 (Localized Euler Estimates). *With the assumptions of the previous theorem, consider*

$$dz = \hat{V}(z) d\mathbf{x}$$

where $\hat{V} = (\hat{V}_1, \ldots, \hat{V}_d)$ is a collection of smooth (possibly unbounded) vector fields, nonexplosive in the sense that $|z|_{\infty;[0,1]} \leq \varphi(M)$ for some increasing function φ. Then for all $t \in [0,1]$

$$\left| z_{0,t} - \mathcal{E}_{(\hat{V})}\left(z_0, S_m(\mathbf{x})_{0,t}\right)\right| \leq C(M) \times t^{\frac{m+1}{p}}.$$

Proof. We can replace $\widehat{V}$ by (compactly supported) vector fields $\widetilde{V}$ such that $\widehat{V} \equiv \widetilde{V}$ on the ball $B(0, \varphi(M))$.

After this localization we apply the previous theorem. □

LEMMA 2. *Let f be a smooth function on $\mathbb{R}^e$ lifted to a smooth function on $\mathbb{R}^e \oplus \mathbb{R}^{e \times e} \oplus \mathbb{R}^e$ by*

$$\hat{f}(z^1, z^2, z^3) = f(z^3).$$

Viewing vector fields as first order differential operators, we have

$$\widehat{V}_{i_1} \dots \widehat{V}_{i_N}|_{z_0} \hat{f} = [V_{i_1}, \dots, V_{i_N}, W]|_{y_0} f.$$

As a consequence, for any element

$$S_m(\mathbf{x})_{0,t} = \left(\mathbf{x}^k_{0,t} : k \in \{0, 1, \dots, m\}\right) \in G^m(\mathbb{R}^d),$$

we have

$$\left| z^3_t - W|_{y_0} - \sum_{k=1}^{m} \mathbf{x}^k_{0,t} \cdot [V, \dots, V, W]|_{y_0} \right| \leq \left| z_{0,t} - \mathscr{E}_{(\widehat{V})}\left(z_0, S_m(\mathbf{x})_{0,t}\right) \right|.$$

Proof. Taylor expansion of the evolution equation of $z^3(t)$ shows that

$$\widehat{V}_{i_1} \dots \widehat{V}_{i_N}|_{z_0} f = [V_{i_1}, \dots, V_{i_N}, W]|_{y_0} f,$$

as required. □

COROLLARY 3. *Fix $a \in \mathcal{T}_{y_0}\mathbb{R}^e \cong \mathbb{R}^e$ with $|a| = 1$. Let $p > 1$ and $\mathbf{X}$ be a random geometric p-rough path whose p-variation is controlled by a (random) constant times $\omega(s,t) = t - s$ so that*

$$\|\mathbf{X}\|_{1/p\text{-Höl};[0,1]} \equiv \sup_{0 \leq s \leq t \leq 1} \|\mathbf{X}_{s,t}\| / |t - s|^{1/p} < \infty \text{ a.s.}$$

Assume that for some $H \in (0,1)$, we have $Hp < 1 + 1/m$. Then, writing y for the solution to the random RDE $dy = V(y)\,d\mathbf{X}$ *started at y_0, and J for the Jacobian of its flow, we have*

$$\mathbb{P}\left[\left| a^T J_{0 \leftarrow t}(W(y_t)) - \sum_{k=0}^{m} a^T \left(\mathbf{X}^k_{0,t} \cdot [V, \dots, V, W]|_{y_0}\right)\right|_{t=1/n} > \frac{\varepsilon}{2} n^{-mH}\right] \to 0$$

with $n \to \infty$.

Proof. From [14], Wiener's characterization applies to geometric rough paths, $p > 1$, and in particular

$$\mathbb{P}\left[\|\mathbf{X}\|_{1/p\text{-Höl};[0,1/n]} \geq 1\right] = o(1) \text{ i.e. } \to 0 \text{ as } n \to \infty.$$

We then bound

$$\mathbb{P}\left[\left|a^T J_{0\leftarrow t}\left(W\left(y_t\right)\right)-\sum_{k=0}^{m} a^T\left(\mathbf{X}_{0,t}^k \cdot [V,\ldots,V,W]|_{y_0}\right)\right|_{t=1/n} > \frac{\varepsilon}{2} n^{-mH}\right]$$

from above by

$$\begin{aligned}
&\mathbb{P}\Bigg[\left|a^T J_{0\leftarrow t}\left(W\left(y_t\right)\right)-\sum_{k=0}^{m} a^T\left(\mathbf{X}_{0,t}^k \cdot [V,\ldots,V,W]|_{y_0}\right)\right|_{t=1/n} \\
&\qquad > \frac{\varepsilon}{2} n^{-mH}; \|\mathbf{X}\|_{1/p\text{-Höl};[0,1/n]} \le 1\Bigg] \\
&\quad + \mathbb{P}\left[\|\mathbf{X}\|_{1/p\text{-Höl};[0,1/n]} \ge 1\right] \qquad \text{then, using } |a|=1 \text{ and the previous lemma,} \\
&\le \mathbb{P}\left[\left|z_{0,1/n}-\mathcal{E}_{(\hat{V})}\left(z_0, S_m(\mathbf{X})_{0,1/n}\right)\right| > \frac{\varepsilon}{2} n^{-mH}; \|\mathbf{X}\|_{1/p\text{-Höl};[0,1/n]} \le 1\right]+o(1) \\
&\le \mathbb{P}\left[C(1)\times\left(\frac{1}{n}\right)^{\frac{m+1}{p}} > \frac{\varepsilon}{2} n^{-mH}\right]+o(1) \text{ using the localized Euler estimates.}
\end{aligned}$$

The probability of the (deterministic) event

$$C(1)\left(\frac{1}{n}\right)^{\frac{m+1}{p}} > \frac{\varepsilon}{2}\left(\frac{1}{n}\right)^{mH}$$

will be zero for n large enough provided $\frac{m+1}{p} > mH$ which is what we assumed. □

6. On Hörmander's condition

Let $V=(V_1,\ldots,V_d)$ denote a collection of smooth vector fields defined in a neighborhood of $y_0 \in \mathbb{R}^e$. Given a multi-index $I=(i_1,\ldots,i_k) \in \{1,\ldots,d\}^k$, with length $|I|=k$, the vector field V_I is defined by iterated Lie brackets

$$V_I := [V_{i_1}, V_{i_2}, \ldots, V_{i_k}] \equiv [V_{i_1}, [V_{i_2}, \ldots, [V_{i_{k-1}}, V_{i_k}] \ldots]. \tag{6.1}$$

If W is another smooth vector field defined in a neighborhood of $y_0 \in \mathbb{R}^e$ we write[10]

$$\underbrace{a}_{\in\left(\mathbb{R}^d\right)^{\otimes(k-1)}} \cdot \underbrace{[V,\ldots,V,W]}_{\text{length } k} := \sum_{\substack{i_1,\ldots,i_{k-1} \\ \in\{1,\ldots,d\}}} a^{i_1,\ldots,i_{k-1}} [V_{i_1}, V_{i_2}, \ldots, V_{i_{k-1}}, W].$$

Recall that the step-r free nilpotent group with d generators, $G^r(\mathbb{R}^d)$, was realized as a submanifold of the tensor algebra

$$T^{(r)}\left(\mathbb{R}^d\right) \equiv \oplus_{k=0}^{r}\left(\mathbb{R}^d\right)^{\otimes k}.$$

[10] We introduced this notation already in the previous section; cf. (5.2).

Definition 2. Given $r \in \mathbb{N}$ we say that condition $(\mathrm{H})_r$ holds at $y_0 \in \mathbb{R}^e$ if

$$\operatorname{span}\left\{V_I|_{y_0} : |I| \le r\right\} = \mathcal{T}_{y_0}\mathbb{R}^e \cong \mathbb{R}^e. \tag{6.2}$$

Similarly, we say that $(\mathrm{HT})_r$ holds at y_0 if the span of

$$\left\{\pi_{k-1}(\mathbf{g}) \cdot \underbrace{[V, \dots, V, V_i]}_{\text{length } k}|_{y_0} : k = 1, \dots, r; i = 1, \dots, d,\ \mathbf{g} \in G^{r-1}(\mathbb{R}^d)\right\} \tag{6.3}$$

equals $\mathcal{T}_{y_0}\mathbb{R}^e \cong \mathbb{R}^e$. Hörmander's condition (H) is satisfied at y_0 if and only if $(\mathrm{H})_r$ holds for some $r \in \mathbb{N}$. Similarly, we say that the Hörmander-type condition (HT) is satisfied at y_0 if and only if $(\mathrm{HT})_r$ holds for some $r \in \mathbb{N}$. (When no confusion arises we omit reference to y_0.)

PROPOSITION 4. *For any fixed $r \in \mathbb{N}$, the span of* (6.2) *equals the span of* (6.3). *Consequently, Hörmander's condition* (*H*) *at y_0 is equivalent to the Hörmander-type condition* (HT) *at y_0.*

Proof. We first make the trivial observation that $(\mathrm{HT})_r$ implies $(\mathrm{H})_r$ for any $r \in \mathbb{N}$. For the converse, fixing a multi-index $I = (i_1, \dots, i_{k-1}, i_k)$ of length $k \le r$ and writing $e_1, \dots e_d$ for the canonical basis of $\mathbb{R}^d$ define

$$\begin{aligned}\mathbf{g} &= \mathbf{g}(t_1, \dots, t_{k-1}) \\ &= \exp(t_1 e_{i_1}) \otimes \cdots \otimes \exp\left(t_{k-1} e_{i_{k-1}}\right) \\ &\in G^{r-1}(\mathbb{R}^d) \subset T^{r-1}\left(\mathbb{R}^d\right).\end{aligned}$$

(Recall that $T^{r-1}\left(\mathbb{R}^d\right)$ is a tensor algebra with multiplication $\otimes$, exp is defined by the usual series and the CBH formula shows that the so-defined g is indeed in $G^{r-1}(\mathbb{R}^d)$ as claimed.) It follows that any

$$\pi_{k-1}(\mathbf{g}) \cdot \underbrace{\left[V, \dots, V, V_{i_k}\right]}_{\text{length } k}|_{y_0}$$

lies in the $(\mathrm{HT})_r$-span i.e. the linear span of (6.3). Now, the $(\mathrm{HT})_r$-span is a closed linear subspace of $\mathcal{T}_{y_0}\mathbb{R}^e \cong \mathbb{R}^e$ and so it is clear that any element of form

$$\pi_{k-1}(\partial_\alpha \mathbf{g}) \cdot \underbrace{\left[V, \dots, V, V_{i_k}\right]}_{\text{length } k}|_{y_0}$$

where ∂_α stands for any higher order partial derivative with respect to $t_1, \dots, t_{k-1}$, i.e.,

$$\partial_\alpha = \left(\frac{\partial}{\partial t_1}\right)^{\alpha_1} \cdots \left(\frac{\partial}{\partial t_{k-1}}\right)^{\alpha_{k-1}} \quad \text{with } \alpha \in (\mathbb{N} \cup \{0\})^{k-1}$$

is also in the $(\mathrm{HT})_r$-span for any $t_1, \ldots, t_{k-1}$ and, in particular, when evaluated at $t_1 = \cdots = t_{k-1} = 0$. For the particular choice $\alpha = (1, \ldots, 1)$ we have

$$\frac{\partial^{k-1}}{\partial t_1 \ldots \partial t_{k-1}} \mathbf{g}|_{t_1=0 \ldots t_{k-1}=0} = e_{i_1} \otimes \cdots \otimes e_{i_{k-1}} =: \mathbf{h},$$

where $\mathbf{h}$ is an element of $T^{r-1}(\mathbb{R}^d)$ with the only nonzero entry arising on the $(k-1)^{\text{th}}$ tensor level; i.e.,

$$\pi_{k-1}(\mathbf{h}) = e_{i_1} \otimes \cdots \otimes e_{i_{k-1}}.$$

Thus,

$$\pi_{k-1}(\mathbf{h}) \cdot \underbrace{[V, \ldots, V, V_{i_k}]}_{\text{length } k}|_{y_0} = [V_{i_1}, \ldots, V_{i_{k-1}}, V_{i_k}]|_{y_0}$$

is in our $(\mathrm{HT})_r$- span. But this says precisely that, for any multi-index I of length $k \leq r$, the bracket vector field evaluated at y_0, i.e., $V_I|_{y_0}$ is an element of our $(\mathrm{HT})_r$-span. □

7. **Proof of main result**

We are now in a position to give

Proof (*of Theorem* 1). We fix $t \in (0, T]$. As usual it suffices to show a.s. invertibility of

$$\sigma_t = \left(\left\langle DY_t^i, DY_t^j \right\rangle_{\mathcal{H}} \right)_{i,j=1,\ldots,e} \in \mathbb{R}^{e \times e}.$$

In terms of an ONB (h_n) of the Cameron-Martin space we can write

$$\begin{aligned} \sigma_t &= \sum_n \langle DY_t, h_n \rangle_{\mathcal{H}} \otimes \langle DY_t, h_n \rangle_{\mathcal{H}} \\ &= \sum_n \int_0^t J_{t \leftarrow s}^{\mathbf{X}} (V_k(Y_s))\, dh_{n,s}^k \otimes \int_0^t J_{t \leftarrow s}^{\mathbf{X}} (V_l(Y_s))\, dh_{n,s}^l. \end{aligned} \tag{7.1}$$

(Summation over up-down indices is from here on tacitly assumed.) Invertibility of σ is equivalent to invertibility of the reduced covariance matrix

$$C_t := \sum_n \int_0^t J_{0 \leftarrow s}^{\mathbf{X}} (V_k(Y_s))\, dh_{n,s}^k \otimes \int_0^t J_{0 \leftarrow s}^{\mathbf{X}} (V_l(Y_s))\, dh_{n,s}^l$$

which has the advantage of being adapted, i.e. being $\sigma(X_s : s \in [0, t])$-measurable. We now assume that

$$\mathbb{P}(\det C_t = 0) > 0$$

and will see that this leads to a contradiction with Hörmander's condition.

Step 1. Let K_s be the random subspace of $\mathcal{T}_{y_0}\mathbb{R}^e \cong \mathbb{R}^e$. spanned by

$$\left\{ J^{\mathbf{X}}_{0\leftarrow r}\left(V_k\left(Y_r\right)\right); r \in [0,s], k = 1,\dots,d\right\}.$$

The subspace $K_{0+} = \cap_{s>0} K_s$ is measurable with respect to the germ σ-algebra and by our "0-1 law" assumption, deterministic with probability one. A random time is defined by

$$\Theta = \inf\{s \in (0,t] : \dim K_s > \dim K_{0+}\} \wedge t,$$

and we note that $\Theta > 0$ a.s. For any vector $v \in \mathbb{R}^e$ we have

$$v^T C_t v = \sum_n \left| \int_0^t v^T J^{\mathbf{X}}_{0\leftarrow s}\Big(V_k(Y_s)) dh^k_{n,s}\right|^2.$$

Assuming $v^T C_t v = 0$ implies

$$\forall n : \int_0^t v^T J^{\mathbf{X}}_{0\leftarrow s}\left(V_k\left(Y_s\right)\right) dh^k_{n,s} = 0$$

and hence, by our nondegeneracy condition on the Gaussian process

$$v^T J^{\mathbf{X}}_{0\leftarrow s}\left(V_k\left(Y_s\right)\right) = 0$$

for any $s \in [0,t]$ and any $k = 1,\dots,d$ which implies that v is orthogonal to K_t. Therefore, $K_{0+} \neq \mathbb{R}^e$, otherwise $K_s = \mathbb{R}^e$ for every $s > 0$ so that v must be zero, which implies C_t is invertible a.s. in contradiction with our hypothesis.

Step 2. We saw that K_{0+} is a deterministic and linear subspace of $\mathbb{R}^e$ with strict inclusion $K_{0+} \subsetneqq \mathbb{R}^e$ In particular, there exists a deterministic vector $z \in \mathbb{R}^e \backslash \{0\}$ which is orthogonal to K_{0+}. We will show that z is orthogonal to all vector fields and (suitable) brackets evaluated at y_0, thereby contradicting the fact that our vector fields satisfy Hörmander's condition. By definition of Θ, $K_{0+} \equiv K_t$ for $0 \le t < \Theta$ and so for every $k = 1,\dots d$,

$$z^T J^{\mathbf{X}}_{0\leftarrow t}\left(V_k\left(Y_t\right)\right) = 0 \text{ for } t \le \Theta. \tag{7.2}$$

Observe that, by evaluation at $t = 0$, this implies $z \perp \operatorname{span}\{V_1,\dots,V_d\}|_{y_0}$.

Step 3. We call an element $\mathbf{g} \in \oplus_{k=0}^{\infty}\left(\mathbb{R}^d\right)^{\otimes k}$ group-like if and only if for any $N \in \mathbb{N}$,

$$\left(\pi_0\left(\mathbf{g}\right),\dots,\pi_N\left(\mathbf{g}\right)\right) \in G^N\left(\mathbb{R}^d\right) \subset \oplus_{k=0}^{N}\left(\mathbb{R}^d\right)^{\otimes k}.$$

We now keep k fixed and make *induction hypothesis* $I(m-1)$:

$$\text{for all } \mathbf{g} \text{ group-like, } j \le m-1 : \; z^T \pi_j\left(\mathbf{g}\right)\left[V,\dots,V,V_k\right]\big|_{y_0} = 0.$$

To this end, take the shortest path $\gamma^n : [0,1/n] \to \mathbb{R}^d$ such that $S_m\left(\gamma^n\right)$ equals $\pi_{1,\dots,m}\left(\mathbf{g}\right)$, the projection of $\mathbf{g}$ to the free step-m nilpotent group with d generators, denoted $G^m(\mathbb{R}^d)$. Then

$$|\gamma^n|_{1\text{-var};[0,1/n]} = \|\pi_{1,\dots,m}(\mathbf{g})\|_{G^m(\mathbb{R}^d)} < \infty$$

and the scaled path

$$h^n(t) = n^{-H}\gamma^n(t), \quad H \in (0,1)$$

has length (over the interval $[0,1/n]$) proportional to n^{-H} which tends to 0 as $n \to \infty$. Our plan is to show that

$$\text{(7.3)} \qquad \text{for all } \varepsilon > 0 : \liminf_{n\to\infty} \mathbb{P}\left(\left|z^T J^{h^n}_{0\leftarrow 1/n}\left(V_k\left(y^{h^n}_{1/n}\right)\right)\right| < \varepsilon/n^{mH}\right) > 0$$

which, since the event involved is deterministic, really says that

$$\left|n^{mH} z^T J^{h^n}_{0\leftarrow 1/n}\left(V_k\left(y^{h^n}_{1/n}\right)\right)\right| < \varepsilon$$

holds true for all $n \ge n_0(\varepsilon)$ large enough. Then, sending $n \to \infty$, a Taylor expansion and $I(m-1)$ shows that the left-hand side converges to

$$\left|z^T \underbrace{n^{mH}\pi_m\left(S_m\left(h^n\right)\right)}_{=\pi_m(\mathbf{g})} \cdot [V,\dots,V,V_k]\big|_{y_0}\right| < \varepsilon$$

and since $\varepsilon > 0$ is arbitrary we showed $I(m)$ which completes the induction step.

Step 4. The only thing left to show is (7.3), that is, positivity of lim inf of

$$\begin{aligned}
&\mathbb{P}\left(\left|z^T J^{h^n}_{0\leftarrow 1/n}\left(V_k\left(y^{h^n}_{1/n}\right)\right)\right| < \varepsilon/n^{mH}\right) \\
&\quad \ge \mathbb{P}\left(\left|z^T J^{\mathbf{X}}_{0\leftarrow \cdot}\left(V_k\left(y_\cdot\right)\right) - z^T J^{h^n}_{0\leftarrow \cdot}\left(V_k\left(y^{h^n}_\cdot\right)\right)\right|_{\cdot=1/n} < \varepsilon/n^{mH}\right) \\
&\qquad -\mathbb{P}\left(\Theta \le 1/n\right)
\end{aligned}$$

and since $\Theta > 0$ a.s. it is enough to show that

$$\liminf_{n\to\infty} \mathbb{P}\left(\left|z^T J^{\mathbf{X}}_{0\leftarrow \cdot}\left(V_k\left(y_\cdot\right)\right) - z^T J^{h^n}_{0\leftarrow \cdot}\left(V_k\left(y^{h^n}_\cdot\right)\right)\right|_{\cdot=1/n} < \varepsilon/n^{mH}\right) > 0.$$

Using $I(m-1)$ + stochastic Taylor expansion (more precisely, Corollary 3) this is equivalent to show positivity of lim inf of

$$\mathbb{P}\left(\left|z^T \mathbf{X}^m_{0,\cdot}[V,\dots,V,V_k] - z^T J^{h^n}_{0\leftarrow \cdot}\left(V_k\left(y^{h^n}_\cdot\right)\right)\right|_{\cdot=1/n} < \frac{\varepsilon}{2}/n^{mH}\right).$$

(Let us remark that the assumption $Hp < 1 + 1/m$ needed to apply Corollary 3 is satisfied thanks to Condition 5, part (ii), and the remark that our induction stops when m has reached r, the number of brackets needed in Hörmander's condition.)

Rewriting things, we need to show positivity of lim inf of

$$\mathbb{P}\Big(|n^{mH} z^T\, [V,\dots,V,V_k]\,\mathbf{X}^m_{0,1/n} - \underbrace{z^T n^{mH} J^{h^n}_{0\leftarrow 1/n}\left(V_k\left(y^{h^n}_{1/n}\right)\right)}_{\to z^T[V,\dots,V,V_k]\pi_m(\mathbf{g})}| < \frac{\varepsilon}{2}\Big)$$

or, equivalently, that

$$\lim\inf_{n\to\infty} \mathbb{P}\left(\left|z^T\, [V,\dots,V,V_k]|_{y_0}\left(n^{mH}\mathbf{X}^m_{0,1/n} - \pi_m(\mathbf{g})\right)\right| < \frac{\varepsilon}{2}\right) > 0.$$

But this is implied by Condition 5 and so the proof is finished. □

8. **Examples**

8.1. *Brownian motion.* For d-dimensional standard Brownian motion on $[0,T]$, each component has covariance $R(s,t) = \min(s,t)$. As is well understood [26], [13] one needs to add Lévy's area process to obtain a geometric p-rough path, any $p > 2$ (known in this context as Brownian rough path or Enhanced Brownian motion). A solution to (1.2) in the rough path sense then precisely solves the stochastic differential equation in Stratonovich form

$$dY = \sum_{i=1}^{d} V_i\,(Y) \circ dB^i.$$

Subject to Hörmander's condition (H), Theorem 1 then shows that $Y_t = Y_t(\omega)$ has a density for $t > 0$ which is of course well-known.

8.2. *Fractional Brownian motion.* The covariance of fractional Brownian motion with Hurst parameter $H \in (0,1)$ is given by

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

so that Brownian motion corresponds to $H = 1/2$. For $H > 1/4$ it admits a lift to a Gaussian geometric p-rough path[11], for any $p > 1/H$. Various constructions are possible and references were already given in Section 3, see also the discussion following Condition 2. As was detailed throughout Section 4, all conditions formulated therein are satisfied for fBM and so Theorem 1 tells us that $Y_t = Y_t\,(\omega)$, solution to the RDE driven by a multi-dimensional fBM with Hurst parameter $H > 1/4$, has a density for all positive times provided the vector fields satisfy Hörmander's condition.

The novelty is of course the degenerate regime $H < 1/2$ with sample path regularity worse than Brownian motion.

[11] As is well understood [26], for $H \le 1/4$ fractional Brownian increments decorrelate too slowly for stochastic area to exist and so there is no meaningful lift of fBM with $H \le 1/4$ to a geometric rough path.

8.3. *Ornstein-Uhlenbeck process.* Let B be a standard d-dimensional Brownian motion and define the centered Gaussian process X by Wiener-Itô integration,

$$X_t^i = \int_0^t e^{-(t-r)} dB_r^i \text{ with } i = 1, \ldots, d.$$

X satisfies the Itô differential equations, $dX_t = -X_t dt + dB_t$ and is also a semimartingale. The conditions of Section 4 are readily checked (in essence, one uses $X_t \sim B_t$ at $t \to 0+$ and the absence of Brownian bridge type degeneracy); only Condition 5 deserves a detailed discussion which we present below. The conclusion of Theorem 1 can then be stated by saying that the unique Stratonovich solution to $dY = \sum V_i (Y) \circ dX^i$ admits a density for all positive times provided the vector fields satisfy Hörmander's condition (H).

To see that X satisfies Condition 5 we first remark that (X, B) is easily seen to satisfies the assumptions of Theorem 2. (In fact, one sees $\rho = 1$ and we are dealing with geometric p-rough paths of Brownian regularity, i.e. $p = 2 + \varepsilon$.) Condition (4.4) then holds with $H = 1/2$: take $s, t \in [0, 1/n]$ and compute, with focus on one nondiagonal entry,

$$\begin{aligned} R_{X-B}(s,t) &\equiv \mathbb{E}[(X_s - B_s)(X_t - B_t)] \\ &= \int_0^t \left(e^{-(s-r)} - 1\right)\left(e^{-(t-r)} - 1\right) dr = O\left(n^{-3}\right). \end{aligned}$$

By Corollary 1 we see Condition 5 holds for the Ornstein-Uhlenbeck examples.

8.4. *Brownian bridge.* Let B be a d-dimensional standard Brownian motion. Define the Brownian bridge returning to zero at time T by

$$X_t^T := B_t - \frac{t}{T} B_T \text{ for } t \in [0, T].$$

Equivalently, one can define X^T via the covariance

$$R^T(s,t) = \min(s,t)\,(1 - \max(s,t)/T).$$

Clearly, $X_t^T|_{t=T} = 0$ and trivially (take $dY = dX$) the conclusion of Theorem 1 cannot hold; this behavior is indeed ruled out by Condition 3 in Section 4. On the other hand, we may consider $X^{T+\varepsilon}$ restricted to $[0, T]$ and in this case the conditions in Section 4 are readily verified. (In particular, Condition 5 is checked as in the Ornstein-Uhlenbeck example, by comparison of X_t with B_t for $t \to 0+$.) It is worth remarking that $Z := X^{T+\varepsilon}$ stopped at time T is also a semimartingale; for instance, by writing $\left(X_t^{T+\varepsilon} : t \leq T\right)$ as strong solution to an Itô differential equation with (well-behaved) drift (as long as $t \leq T$). The conclusion of Theorem 1 can then be stated by saying that the unique Stratonovich solution to $dY = \sum V_i (Y) \circ dZ^i$ admits a density for all times $t \in (0, T]$ provided the vector fields satisfy Hörmander's condition (H).

8.5. *Further examples.* Further examples (for instance, "fractional" versions of the Brownian bridge and Ornstein-Uhlenbeck process) are readily constructed. Generalizing Examples 8.2 and 8.3 one could consider Volterra processes [7], i.e., Gaussian process with representation $X_t = \int_0^t K(t,s)dB_s$ and derive sufficient conditions on the kernel K which imply those of Section 4. Existence of a rough path lift of X aside, one would need nondegeneracy of K and certain scaling properties as $t \to 0+$ but we shall not pursue this here. (In any case, there are non-Volterra examples, such as the Brownian bridge returning to zero at $(T+\varepsilon)$, to which Theorem 1 applies.)

Acknowledgment. The second author is partially supported by a Leverhulme Research Fellowship and EPSRC grant EP/E048609/1.

References

[1] F. BAUDOIN and M. HAIRER, A version of Hörmander's theorem for the fractional Brownian motion, *Probab. Theory Related Fields* **139** (2007), 373–395. MR 2008e:60195 Zbl 1123.60038

[2] D. R. BELL, *The Malliavin Calculus*, Dover Publ., Mineola, NY, 2006. MR 2007k:60003 Zbl 1099.60041

[3] T. CASS, P. FRIZ, and N. VICTOIR, Non-degeneracy of Wiener functionals arising from rough differential equations, *Trans. Amer. Math. Soc.* **361** (2009), 3359–3371. MR 2485431 Zbl 05558366

[4] P. CATTIAUX and L. MESNAGER, Hypoelliptic non-homogeneous diffusions, *Probab. Theory Related Fields* **123** (2002), 453–483. MR 2004d:60178 Zbl 1009.60058

[5] K.-T. CHEN, Integration of paths, geometric invariants and a generalized Baker-Hausdorff formula, *Ann. of Math.* **65** (1957), 163–178. MR 19,12a Zbl 0077.25301

[6] L. COUTIN and Z. QIAN, Stochastic analysis, rough path analysis and fractional Brownian motions, *Probab. Theory Related Fields* **122** (2002), 108–140. MR 2003c:60066 Zbl 1047.60029

[7] L. DECREUSEFOND, Stochastic integration with respect to Volterra processes, *Ann. Inst. H. Poincaré Probab. Statist.* **41** (2005), 123–149. MR 2005m:60117 Zbl 1071.60040

[8] F. DELBAEN and W. SCHACHERMAYER, *The Mathematics of Arbitrage*, *Springer Finance*, Springer-Verlag, New York, 2006. MR 2007a:91001 Zbl 1106.91031

[9] R. M. DUDLEY, Sample functions of the Gaussian process, *Ann. Probability* **1** (1973), 66–103. MR 49 #11605 Zbl 0261.60033

[10] D. FEYEL and A. DE LA PRADELLE, Curvilinear integrals along enriched paths, *Electron. J. Probab.* **11** (2006), no. 34, 860–892. MR 2007k:60112 Zbl 1110.60031

[11] D. FILIPOVIĆ and J. TEICHMANN, On the geometry of the term structure of interest rates, in *Stochastic Analysis with Applications to Mathematical Finance*, *Proc. Royal Soc. Lond. Ser. A Math. Phys. Eng. Sci.* **460**, 2004, pp. 129–167. MR 2005b:60145 Zbl 1048.60045

[12] P. FRIZ and N. VICTOIR, Differential equations driven by Gaussian signals, to appear in *Ann. Inst. H. Poincaré Probab. Statist.* Available at http://www.imstat.org/aihp/pdf/AIHP202.pdf

[13] ———, Approximations of the Brownian rough path with applications to stochastic analysis, *Ann. Inst. H. Poincaré Probab. Statist.* **41** (2005), 703–724. MR 2007e:60018 Zbl 1080.60021

[14] P. FRIZ and N. VICTOIR, A note on the notion of geometric rough paths, *Probab. Theory Related Fields* **136** (2006), 395–416. MR 2007k:60114 Zbl 1108.34052

[15] ———, A variation embedding theorem and applications, *J. Funct. Anal.* **239** (2006), 631–637. MR 2007k:60113 Zbl 1114.46022

[16] ———, Euler estimates for rough differential equations, *J. Differential Equations* **244** (2008), 388–412. MR 2009j:60151 Zbl 1140.60037

[17] ———, *Multidimensional Stochastic Processes as Rough Paths*: *Theory and Applications*, *Cambridge Stud. Adv. Math.* **120**, Cambridge Univ. Press, Cambridge, 2010. Zbl pre05677244

[18] M. HAIRER, Ergodicity of stochastic differential equations driven by fractional Brownian motion, *Ann. Probab.* **33** (2005), 703–758. MR 2005k:60178 Zbl 1071.60045

[19] M. HAIRER and J. C. MATTINGLY, Ergodicity of the 2D Navier-Stokes equations with degenerate stochastic forcing, *Ann. of Math.* **164** (2006), 993–1032. MR 2008a:37095 Zbl 1130.37038

[20] Y. HU and D. NUALART, Differential equations driven by Hölder continuous functions of order greater than 1/2, in *Stochastic Analysis and Applications*, *Abel Symp.* **2**, Springer-Verlag, New York, 2007, pp. 399–413. MR 2009d:60179 Zbl 1144.34038

[21] S. KUSUOKA, The nonlinear transformation of Gaussian measure on Banach space and absolute continuity. I, *J. Fac. Sci. Univ. Tokyo Sect. IA Math.* **29** (1982), 567–597. MR 84j:60015a Zbl 0525.60050

[22] M. LEDOUX, Z. QIAN, and T. ZHANG, Large deviations and support theorem for diffusion processes via rough paths, *Stochastic Process. Appl.* **102** (2002), 265–283. MR 2003m:60152 Zbl 1075.60510

[23] T. J. LYONS, Differential equations driven by rough signals, *Rev. Mat. Iberoamericana* **14** (1998), 215–310. MR 2000c:60089 Zbl 0923.34056

[24] T. J. LYONS and Z. QIAN, Calculus of variation for multiplicative functionals, in *New Trends in Stochastic Analysis* (*Charingworth*, 1994), World Sci. Publ., River Edge, NJ, 1997, pp. 348–374. MR 2000k:60112

[25] ———, Flow of diffeomorphisms induced by a geometric multiplicative functional, *Probab. Theory Related Fields* **112** (1998), 91–119. MR 99k:60153 Zbl 0918.60009

[26] ———, *System Control and Rough Paths*, *Oxford Mathematical Monographs*, Oxford University Press, Oxford, 2002. MR 2005f:93001 Zbl 1044.93009

[27] P. MALLIAVIN, *Stochastic Analysis*, *Grundl. Math. Wissen.* **313**, Springer-Verlag, New York, 1997. MR 99b:60073 Zbl 0878.60001

[28] B. B. MANDELBROT and J. W. VAN NESS, Fractional Brownian motions, fractional noises and applications, SIAM *Rev.* **10** (1968), 422–437. MR 39 #3572 Zbl 0179.47801

[29] D. NUALART, *The Malliavin Calculus and Related Topics*, Second ed., *Prob. Appl.* (N.Y.), Springer-Verlag, New York, 2006. MR 2006j:60004 Zbl 1099.60003

[30] D. NUALART and B. SAUSSEREAU, Malliavin calculus for stochastic differential equations driven by a fractional Brownian motion, *Stochastic Process. Appl.* **119** (2009), 391–409. MR : 2493996 Zbl 05527425

[31] J. PICARD, A tree approach to p-variation and to integration, *Ann. Probab.* **36** (2008), 2235–2279. MR 2478682 Zbl 1157.60055

[32] I. SHIGEKAWA, *Stochastic Analysis*, *Transl. Math. Monogr.* **224**, Amer. Math. Soc., Providence, RI, 2004. MR 2005k:60002 Zbl 1064.60003

[33] S. TANIGUCHI, Applications of Malliavin's calculus to time-dependent systems of heat equations, *Osaka J. Math.* **22** (1985), 307–320. MR 87a:60066 Zbl 0583.35055

[34] A. S. ÜSTÜNEL and M. ZAKAI, *Transformation of Measure on Wiener Space*, *Springer Monogr. Math.*, Springer-Verlag, New York, 2000. MR 2001g:60137 Zbl 0974.46044

(Received October 5, 2007)
(Revised May 30, 2008)

E-mail address: cass@maths.ox.ac.uk
STATISTICAL LABORATORY/DPMMS, CENTRE OF MATHEMATICAL SCIENCES, WILBERFORCE RD., CAMBRIDGE CB3 0WB ENGLAND
Current address: MATHEMATICAL INSTITUTE, 24-29 ST. GILES', OXFORD OX1 3LB, ENGLAND
http://www.maths.ox.ac.uk/ldapcontact/userdetails/Cass/0

E-mail address: friz@math.tu-berlin.de
STATISTICAL LABORATORY/DPMMS, CENTRE OF MATHEMATICAL SCIENCES, WILBERFORCE RD., CAMBRIDGE CB3 0WB ENGLAND
Current address: WEIERSTRASSE INSTITUTE FOR APPLIED ANALYSIS AND STOCHASTICS, MOHRENSTRASSE 39, 10117 BERLIN, GERMANY
Current address: TU BERLIN, FAKULTÄT II, INSTITUT FÜR MATHEMATIK, MA 7-2, STRASSE DES 17. JUNI 136, D-10623 BERLIN, GERMANY
http://www.math.tu-berlin.de/~friz