The Annals of Probability
2013, Vol. 41, No. 4, 3026–3050
DOI: 10.1214/12-AOP821

INTEGRABILITY AND TAIL ESTIMATES FOR GAUSSIAN ROUGH DIFFERENTIAL EQUATIONS

BY THOMAS CASS[1], CHRISTIAN LITTERER[2] AND TERRY LYONS[3]

Imperial College London, Imperial College London and University of Oxford

We derive explicit tail-estimates for the Jacobian of the solution flow for stochastic differential equations driven by Gaussian rough paths. In particular, we deduce that the Jacobian has finite moments of all order for a wide class of Gaussian process including fractional Brownian motion with Hurst parameter $H > 1/4$. We remark on the relevance of such estimates to a number of significant open problems.

1. Introduction.

Gaussian processes that are not necessarily semimartingales arise in modeling a large variety of natural phenomena. The range of their applications reaches from fluid dynamics (e.g., randomly forced Navier–Stokes systems [16]), the modeling of financial markets under transaction costs [14], to the study of internet traffic through queuing models based on fractional Brownian motion (fBm) [13]. These applications motivate the study of stochastic differential equations of the form

$$dY_t = V(Y_t)\,dX_t, \qquad Y(0) = y_0, \tag{1.1}$$

driven by a Gaussian process X. Over the past decade extensive progress has been made understanding the behavior of solutions to such equations. In particular, for the case of fBm with Hurst parameter $H > 1/4$, the work of Cass and Friz [2] shows the existence of the density for (1.1) under Hörmander's condition; Hairer et al. [1, 18] have shown the smoothness of this density and established ergodicity under the regime $H > 1/2$.

Various recent works (Coutin and Qian [5], Ledoux, Qian and Zhang [23], Friz and Victoir [12], Lyons and Hambly [19]) have explored the use of rough paths to understand differential equations driven by nonsemimartingale noise processes. Within this framework we can make sense of the solutions to (1.1) driven by a

Received July 2011; revised July 2012.

[1]Supported in part by EPSRC grant EP/F029578/1 and Christ Church, Oxford. Part of this research was carried out during a stay of the first two authors at the Hausdorff Institute, Bonn.

[2]Supported by a grant of the European Research Council (ERC Grant nr. 258237) and EPSRC Grant EP/H0005500/1.

[3]Supported by EPSRC Grants EP/F029578/1, EP/H000100/1 and by the Oxford-Man Institute for Quantitative Finance.

MSC2010 subject classifications. Primary 60H10; secondary 60G15.

Key words and phrases. Rough path analysis, Gaussian processes.

broader class of Gaussian noises than classical analysis based on Young integration would allow. This class includes fBm with $H > 1/4$. Thus, if we consider the flow $U^{\mathbf{X}}_{t\leftarrow 0}(y_0) \equiv Y_t$ of the RDE (1.1), then under sufficient regularity on V, the map $U^{\mathbf{X}}_{t\leftarrow 0}(\cdot)$ is a differentiable function (see, e.g., [12]), and its derivative ("the Jacobian")

$$J^{\mathbf{X}}_{t\leftarrow 0}(y_0) \equiv DU^{\mathbf{X}}_{t\leftarrow 0}(\cdot)|_{\cdot=y_0}$$

satisfies path-by-path an RDE of linear growth driven by $\mathbf{X}$.

A careful reading of the diverse applications in [1, 18] reveals a surprisingly generic common obstacle to the extensions of such results to the rough path regime. This obstacle eventually boils down to the need for sharp estimates on the integrability of the Jacobian of the flow $J^{\mathbf{X}}_{t\leftarrow 0}(y_0)$. Cass, Lyons [4] and Inahama [21] establish such integrability for the Brownian rough path, but only by using the independence of the increments; for more general Gaussian processes a more careful analysis is needed. To understand the difficulty of this problem, we note from [12] that the standard deterministic estimate on $J^{\mathbf{X}}_{t\leftarrow 0}(y_0)$ gives

$$|J^{\mathbf{x}}_{t\leftarrow 0}(y_0)| \le C \exp(C \|\mathbf{x}\|^p_{p\text{-var};[0,T]}). \tag{1.2}$$

The case where $\mathbf{X}$ is a Gaussian rough path and $p > 2$ (i.e., Brownian-type paths or rougher) the Fernique-type estimates of [7] unfortunately only give that $\|\mathbf{X}\|_{p\text{-var};[0,T]}$ has a Gaussian tail. The right-hand side of (1.2) is hence not integrable in general. Worse still, the work Oberhauser and Friz [6] shows that the inequality (1.2) can actually be saturated for a (deterministic) choice of differential equation and driving rough path. The essential contribution of this paper is that for random processes having enough structure (in particular for Gaussian processes) only a set of small (or zero) measure comes close to equality in (1.2). What is therefore needed (and what we provide!) is a deterministic estimate which respects the fine structure of path, and which allows us to more strongly interrogate its probabilistic structure.

Our results will allow us to deduce the existence of moments of all orders for $J^{\mathbf{x}}_{t\leftarrow 0}(y_0)$ for RDEs driven by a class of Gaussian processes (including, but not restricted to, fBm with Hurst index $H > 1/4$). In fact, our main estimate shows much more than simple moment estimates. Namely, that the logarithm of the Jacobian has a tail that decays faster than an exponential. To be a little more precise, we will show that

$$P(\log[|J^{\mathbf{X}}_{\cdot\leftarrow 0}(y_0)|_{p\text{-var};[0,T]}] > x) \lesssim \exp(-x^r) \tag{1.3}$$

for any $r < r_0 \in (1, 2]$. The constant r_0 will be described in terms of the regularity properties of the Gaussian path.

The results are relevant to a number of important problems. First, they are necessary if one wants to extend the work of [16] and [18] on the ergodicity of non-Markovian systems. Second, they are an important ingredient in a Malliavin calculus proof on the smoothness of the density for RDEs driven by rough Gaussian noise in the elliptic setting. Furthermore, they allow one to achieve an analogue of Hörmander's theorem for Gaussian RDEs in conjunction with a suitable version of Norris's lemma; see [27, 28]. In this context, we remark that Hu and Tindel [20] have recently obtained a Norris lemma for fBm with $H > 1/3$ and proved smoothness-of-density results for a class of nilpotent RDEs. Hairer and Pillai [17] have also proved Hörmander-type theorems for a general class of RDEs; their results are predicated on the assumption that the Jacobian has finite moments of all order. Hence, one application of this paper is to use the tail estimate (1.3) together with the results in [20] or [17] to conclude that for $t > 0$ the law of Y_t [the solution to (1.1)] will, under Hörmander's condition, have a smooth density w.r.t. Lebesgue measure on $\mathbb{R}^e$, for a rich classes of Gaussian processes X which includes fBm $H > 1/3$. All of these problems (and many more besides) require the existence of high-order moments of the Malliavin covariance matrix of $Y_t(\omega)$, which is itself expressed in terms of the Jacobian.

The techniques developed in this paper are relevant to the study of more general RDEs, and not just the one solved by the Jacobian. Our estimates can be applied to any random variable that can be controlled in terms of $N_{\alpha,p,I}(\cdot)$, which is a "greedy" approximation of the local p-variation we will introduce later. Similar deterministic estimates we derive can also be obtained in the following cases (cf. Friz, Victoir [12]):

(1) RDEs driven along linear vector fields of the form $V_i(z) = A_i z + b_i$ for $e \times e$ matrices A_i and b_i in $\mathbb{R}^e$;

(2) higher order derivatives of the flow (subject to suitably enhance regularity on the vector fields defining the flow);

(3) the inverse of the Jacobian of the flow;

(4) situations where one wants to control the distance between two RDE solutions in the (inhomogeneous) rough path metric (e.g., in fixed point theorems).

Recent work [15] has extended the class of linear-growth RDEs for which we have nonexplosion and there may be scope to extend our results to this setting. In this paper we focus only on the Jacobian because of its central role in the wide range of problems we have outlined and obtain explicit bounds for the tails of the distribution of the Jacobian.

We now outline the structure of the paper. In Section 2 we introduce some important notation and concepts on the theory of rough paths. Because this is now standard and there are many references available (e.g., [12, 24–26]), we keep the detail to a minimum. In Section 4 we derive a quantitative bound on the growth of $J^{\mathbf{X}}_{t\leftarrow 0}$; the estimates we derive here are based very closely on [12]. We end up with a control on $J^{\mathbf{X}}_{t\leftarrow 0}$ in terms of a function on the space on (rough) path space which we (suggestively) name the accumulated α local p-variation [denoted by $M_{\alpha,I,p}(\cdot)$]. When $\mathbf{X}$ is taken to be a Gaussian rough path the integrability properties of $M_{\alpha,I,p}(\mathbf{X})$ are not immediately obvious or easy to study. We therefore spend time in Section 4 deriving a relationship between $M_{\alpha,I,p}(\cdot)$ and another function on path space, which we denote $N_{\alpha,p,I}(\cdot)$. The analysis at this stage remains entirely deterministic. Section 5 records some facts about Gaussian rough paths, including the crucial embedding theorems for Cameron–Martin spaces that have been derived in [12]. We then present the main tail estimate on $N_{\alpha,p,I}(\mathbf{X})$. Our analysis is based on Gaussian isoperimetry and more specifically Borell's inequality, which we recall. Once this is achieved we can use the relationship between $J^{\mathbf{X}}_{t\leftarrow 0}$ and $N_{\alpha,p,I}(\mathbf{X})$ to exhibit the stated tail behavior of $J^{\mathbf{X}}_{t\leftarrow 0}$. This estimate then constitutes our main result.

2. Rough path concepts and notation.

There are now many articles and texts providing an overview on rough path theory (e.g., [26] and [12], to name just two). We will focus on establishing the notation we need for the current application. We will study continuous $\mathbb{R}^d$-valued paths x parameterized by time on a compact interval I (sometimes I will be taken to be $[0,T]$), and we denote the space of such functions by $C(I,\mathbb{R}^d)$. We write $x_{s,t} := x_t - x_s$ as a shorthand for the increments of a path when x in $C(I,\mathbb{R}^d)$. For $p \geq 1$ we will use

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

and we refer to these quantities both symbolically and by name (they are, resp., the uniform norm and the p-variation semi-norm). We denote by $C^{p\text{-var}}(I,\mathbb{R}^d)$ the linear subspace of $C(I,\mathbb{R}^d)$ consisting of path of finite p-variation. In the case where x is in $C^{p\text{-var}}(I,\mathbb{R}^d)$ and p is in $[1,2)$, the iterated integrals of x are canonically defined by Young integration. The collection of all these iterated integrals together then gives the signature: for $s<t$ in I

$$S(x)_{s,t} := 1 + \sum_{k=1}^{\infty}\int_{s<t_1<t_2<\cdots<t_k<t} dx_{t_1}\otimes dx_{t_2}\otimes\cdots\otimes dx_{t_k} \in T(\mathbb{R}^d).$$

By writing $S(x)_{\inf I,\cdot}$ we can regard the signature as a path (on I) with values in the tensor algebra. In a similar way, the truncated signature

$$S_N(x)_{s,t} := 1 + \sum_{k=1}^{N}\int_{s<t_1<t_2<\cdots<t_k<t} dx_{t_1}\otimes dx_{t_2}\otimes\cdots\otimes dx_{t_k} \in T^N(\mathbb{R}^d)$$

is a path in the truncated tensor algebra, $T^N(\mathbb{R}^d)$. It is a well-known fact that the path $S_N(x)_{\inf I,\cdot}$ takes values in the step-N free nilpotent group with d generators, which we denote $G^N(\mathbb{R}^d)$. More generally, if $p\geq 1$ we can consider the set of such group-valued paths

$$\mathbf{x}_t = (1,\mathbf{x}^1_t,\ldots,\mathbf{x}^{\lfloor p\rfloor}_t) \in G^{\lfloor p\rfloor}(\mathbb{R}^d).$$

The advantage this offers is that the group structure provides a natural notion of increment, namely $\mathbf{x}_{s,t} := \mathbf{x}_s^{-1}\otimes\mathbf{x}_t$. We can describe the set of "norms" on $G^{\lfloor p\rfloor}(\mathbb{R}^d)$

which are homogeneous with respect to the natural scaling operation on the tensor algebra; see [12] for definitions and details. The subset of these so-called homogeneous norms which are symmetric and sub-additive [12] gives rise to genuine metrics on $G^{\lfloor p\rfloor}(\mathbb{R}^d)$, which in turn gives rise to a notion of homogenous p-variation metrics $d_{p\text{-var}}$ on the $G^{\lfloor p\rfloor}(\mathbb{R}^d)$-valued paths. Let

$$\|\mathbf{x}\|_{p\text{-var};[0,T]} = \left(\sum_{i=1}^{\lfloor p\rfloor} \sup_{D=(t_j)} \sum_{j:t_j\in D} \left|\mathbf{x}^i_{t_j,t_{j+1}}\right|^{p/i}_{(\mathbb{R}^d)^{\otimes i}}\right)^{1/p}, \tag{2.1}$$

and note that if (2.1) is finite, then $\omega(s,t) := \|\mathbf{x}\|^p_{p\text{-var};[s,t]}$ is a control (i.e., it is a continuous, nonnegative, super-additive function on the simplex $\Delta_T = \{(s,t): 0 \le s \le t \le T\}$ which vanishes on the diagonal.)

The space of weakly geometric p-rough paths [denoted $WG\Omega_p(\mathbb{R}^d)$] is the set of paths parameterised over I although this is often implicit with values in $G^{\lfloor p\rfloor}(\mathbb{R}^d)$ such that (2.1) is finite. A refinement of this notion is the space of geometric p-rough paths, denoted $G\Omega_p(\mathbb{R}^d)$, which is the closure of

$$\left\{S_{\lfloor p\rfloor}(x)_{\inf I,\cdot} : x \in C^{1\text{-var}}(I, \mathbb{R}^d)\right\}$$

with respect to the rough path metric $d_{p\text{-var}}$.

We will often end up considering an RDE driven by a path $\mathbf{x}$ in $WG\Omega_p(\mathbb{R}^d)$ along a collection of vector fields $V = (V^1, \dots, V^d)$ on $\mathbb{R}^e$. And from the point of view of existence and uniqueness results, the appropriate way to measure the regularity of the V_is results turns out to be the notion of Lipschitz-γ (short: Lip-γ) in the sense of Stein[4]. This notion provides a norm on the space of such vector fields (the Lip-γ norm), which we denote $|\cdot|_{\text{Lip-}\gamma}$, and for the collection of vector fields V 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.

3. Translated rough paths. Suppose $\mathbf{x} = (1, \mathbf{x}^1, \dots, \mathbf{x}^{\lfloor p\rfloor})$ is a weakly geometric p-rough path. If h is in $C^{q\text{-var}}(I, \mathbb{R}^d)$ and $1/p + 1/q > 1$, then the cross-iterated integrals between h and $\mathbf{x}$ exists canonically using Young integration. This gives rise to the so-called translated rough path $T_h\mathbf{x}$. The definition, which is standard, can be found, for example, in [12] or [24]. A key technical estimate used in the paper will involve this object. Before we state and prove this estimate, we recall the specific structure of $T_h\mathbf{x}$ at the first two nontrivial tensor levels. For levels one and two we have

$$(T_h\mathbf{x})^1 = \mathbf{x}^1 + h,$$
$$(T_h\mathbf{x})^2 = \mathbf{x}^2 + \int h \otimes d\mathbf{x}^1 + \int \mathbf{x}^1 \otimes dh + \int h \otimes dh.$$

[4]See [12] and [26], and note the contrast with classical Lipschitzness.

The higher order terms become increasingly tiresome to write down. We will not go beyond the levelt $(\mathbb{R}^d)^{\otimes 3}$, so we simply record for reference that this can be written as

$$
\begin{aligned}
(T_h\mathbf{x})^3_{s,t} &= \mathbf{x}^3_{s,t} + \int_s^t \int_s^v h_{s,u} \otimes dh_u \otimes dh_v \\
&\quad + \int_s^t \mathbf{x}^2{}_{s,u} \otimes dh_u + \int_s^t \int_s^v \mathbf{x}^1_{s,u} \otimes dh_u \otimes d\mathbf{x}^1_v - \int_s^t h_{s,u} \otimes d\mathbf{x}^2{}_{u,t} \\
&\quad + \int_s^t \int_s^v h_{s,u} \otimes d\mathbf{x}^1_u \otimes dh_v + \int_s^t \int_s^v \mathbf{x}^1{}_{s,u} \otimes dh_u \otimes dh_v \\
&\quad + \int_s^t \int_s^v h_{s,u} \otimes dh_u \otimes d\mathbf{x}^1_v.
\end{aligned} \tag{3.1}
$$

The proof of the following result will occupy the remainder of this section. The lemma is important. It explains how we can control the p-variation of the translated rough path by the sum of the p-variation of the (untranslated) rough path and the q-variation of the path by which we translate.

LEMMA 3.1. *Let $1 \leq p < 4$. Suppose that $\mathbf{x}$ is a weakly geometric p-rough path parametrised over a compact interval I. Let h be a path in $C^{q\text{-var}}(I, \mathbb{R}^d)$ where $1/q + 1/p > 1$. If $T_h\mathbf{x}$ denotes the translated rough path, then for any $[s,t] \subseteq I$ we have the estimate*

$$
\|T_h\mathbf{x}\|^p_{p\text{-var};[s,t]} \leq C_{p,q}\big[\|\mathbf{x}\|^p_{p\text{-var};[s,t]} + |h|^p_{q\text{-var};[s,t]}\big].
$$

The constant $C_{p,q}$ is given explicitly by

$$
C_{p,q} = 2^{p-1}\big[1 + c^{p/2}_{p,q} + c^{p/3}_{p/2,q} + c^{2p/3}_{p,q}\big],
$$

where $c_{l,m} = 2 \cdot 4^{1/l+1/m}\zeta(\frac{1}{l} + \frac{1}{m})$, and ζ is the classical Riemann zeta function.

PROOF. We will only prove the lemma for the most difficult case $p \in [3,4)$. By definition we have that

$$
\|T_h\mathbf{x}\|^p_{p\text{-var};[s,t]} = \big|(T_h\mathbf{x})^1\big|^p_{p\text{-var};[s,t]} + \big|(T_h\mathbf{x})^2\big|^{p/2}_{p/2\text{-var};[s,t]} + \big|(T_h\mathbf{x})^3\big|^{p/3}_{p/3\text{-var};[s,t]},
$$

where for $i = 1,2,3$ we have

$$
\big|(T_h\mathbf{x})^i\big|^{p/i}_{p/i\text{-var};[s,t]} = \sup_{D[s,t]=(t_i)} \sum_{i:t_i \in D[s,t]} \big|(T_h\mathbf{x})^i_{t_i,t_{i+1}}\big|^{p/i}_{(\mathbb{R}^d)^{\otimes i}}.
$$

Note that the formula for the translated rough path gives at level one of the tensor algebra

$$
\begin{aligned}
\big|(T_h\mathbf{x})^1\big|^p_{p\text{-var};[s,t]} &\leq \big[\big|\mathbf{x}^1\big|_{p\text{-var};[s,t]} + |h|_{q\text{-var};[s,t]}\big]^p \\
&\leq 2^{p-1}\big[\|\mathbf{x}\|^p_{p\text{-var};[s,t]} + |h|^p_{q\text{-var};[s,t]}\big] \\
&=: C_1(p,q)\big[\|\mathbf{x}\|^p_{p\text{-var};[s,t]} + |h|^p_{q\text{-var};[s,t]}\big].
\end{aligned} \tag{3.2}
$$

At level two we need to analyze

$$\sum_{i:t_i\in D[s,t]}\Bigg|\mathbf{x}^2_{t_i,t_{i+1}}+\underbrace{\int_{t_i}^{t_{i+1}}h_{t_i,u}\otimes d\mathbf{x}^1_u}_{=:A^1_{t_i,t_{i+1}}}+\underbrace{\int_{t_i}^{t_{i+1}}\mathbf{x}^1_{t_i,u}\otimes dh_u}_{=:A^2_{t_i,t_{i+1}}} \tag{3.3}$$
$$+\int_{t_i}^{t_{i+1}}h_{t_i,u}\otimes dh_u\Bigg|^{p/2}_{(\mathbb{R}^d)^{\otimes 2}}.$$

Using Young's inequality we have for $j=1,2$ that

$$\begin{aligned}|A^j_{t_i,t_{i+1}}|^{p/2}_{(\mathbb{R}^d)^{\otimes 2}}&\le c^{p/2}_{p,q}|\mathbf{x}^1|^{p/2}_{p\text{-}\mathrm{var};[t_i,t_{i+1}]}|h|^{p/2}_{q\text{-}\mathrm{var};[t_i,t_{i+1}]}\\&\le\frac{c^{p/2}_{p,q}}{2}\big(|\mathbf{x}^1|^p_{p\text{-}\mathrm{var};[t_i,t_{i+1}]}+|h|^p_{q\text{-}\mathrm{var};[t_i,t_{i+1}]}\big).\end{aligned}$$

And also

$$\left|\int_{t_i}^{t_{i+1}}h_{t_i,u}\otimes dh_u\right|_{(\mathbb{R}^d)^{\otimes 2}}\le c_{p,q}|h|^2_{q\text{-}\mathrm{var};[t_i,t_{i+1}]}.$$

Hence, we can deduce that

$$\begin{aligned}&\sup_{D[s,t]=(t_i)}\sum_{i:t_i\in D[s,t]}|A^j_{t_i,t_{i+1}}|^{p/2}_{(\mathbb{R}^d)^{\otimes 2}}\\&\quad\le\frac{c^{p/2}_{p,q}}{2}\sup_{D[s,t]=(t_i)}\sum_{i:t_i\in D[s,t]}\big(|\mathbf{x}^1|^p_{p\text{-}\mathrm{var};[t_i,t_{i+1}]}+|h|^p_{q\text{-}\mathrm{var};[t_i,t_{i+1}]}\big)\\&\quad\le\frac{c^{p/2}_{p,q}}{2}\big(|\mathbf{x}^1|^p_{p\text{-}\mathrm{var};[s,t]}+|h|^p_{q\text{-}\mathrm{var};[s,t]}\big),\end{aligned}\tag{3.4}$$

and similarly

$$\sup_{D[s,t]=(t_i)}\sum_{i:t_i\in D[s,t]}\left|\int_{t_i}^{t_{i+1}}h_{t_i,u}\otimes dh_u\right|^{p/2}_{(\mathbb{R}^d)^{\otimes 2}}\le c^{p/2}_{p,q}|h|^p_{q\text{-}\mathrm{var};[t_i,t_{i+1}]}.\tag{3.5}$$

From (3.3), (3.4) and (3.5) we easily obtain that

$$\begin{aligned}&|(T_h\mathbf{x})^2|^{p/2}_{p/2\text{-}\mathrm{var};[s,t]}\\&\quad\le 4^{p/2-1}\big[|\mathbf{x}^2|^{p/2}_{p/2\text{-}\mathrm{var};[s,t]}+c^{p/2}_{p,q}\big(|\mathbf{x}^1|^p_{p\text{-}\mathrm{var};[s,t]}+|h|^p_{q\text{-}\mathrm{var};[s,t]}\big)\\&\qquad\qquad+c^{p/2}_{p,q}|h|^p_{q\text{-}\mathrm{var};[s,t]}\big]\\&\quad\le 4^{(p-1)/2}c^{p/2}_{p,q}\big[\|\mathbf{x}\|^p_{p\text{-}\mathrm{var};[s,t]}+|h|^p_{q\text{-}\mathrm{var};[s,t]}\big]\\&\quad=:C_2(p,q)\big[\|\mathbf{x}\|^p_{p\text{-}\mathrm{var};[s,t]}+|h|^p_{q\text{-}\mathrm{var};[s,t]}\big].\end{aligned}\tag{3.6}$$

We finish the proof by performing a similar analysis on the third level. We need to bound $|(T_h\mathbf{x})^3|^{p/3}_{p/3\text{-var};[s,t]}$. Recall that

$$
\begin{aligned}
(T_h\mathbf{x})^3_{s,t} &= \mathbf{x}^3_{s,t} + \int_s^t \int_s^v h_{s,u}\otimes dh_u \otimes dh_v \\
&\quad + \underbrace{\int_s^t \mathbf{x}^2{}_{s,u}\otimes dh_u}_{=:B^1_{s,t}} + \underbrace{\int_s^t\int_s^v \mathbf{x}^1_{s,u}\otimes dh_u\otimes d\mathbf{x}^1_v}_{=:B^2_{s,t}} - \underbrace{\int_s^t h_{s,u}\otimes d\mathbf{x}^2{}_{u,t}}_{=:B^3_{s,t}} \\
&\quad + \underbrace{\int_s^t\int_s^v h_{s,u}\otimes d\mathbf{x}^1_u\otimes dh_v}_{=:C^1_{s,t}} + \underbrace{\int_s^t\int_s^v \mathbf{x}^1{}_{s,u}\otimes dh_u\otimes dh_v}_{=:C^2_{s,t}} \\
&\quad + \underbrace{\int_s^t\int_s^v h_{s,u}\otimes dh_u\otimes d\mathbf{x}^1_v}_{=:C^3_{s,t}}.
\end{aligned}
\tag{3.7}
$$

We can split this up by first looking at the "pure" terms

$$
\sum_{i:t_i\in D[s,t]} |\mathbf{x}^3_{t_i,t_{i+1}}|^{p/3}_{(\mathbb{R}^d)^{\otimes 3}} \le \|\mathbf{x}\|^p_{p\text{-var};[s,t]},
\tag{3.8}
$$

$$
\begin{aligned}
&\sum_{i:t_i\in D[s,t]} \left|\int_{t_i}^{t_{i+1}}\int_{t_i}^{v} h_{t_i,u}\otimes dh_u\otimes dh_v\right|^{p/3}_{(\mathbb{R}^d)^{\otimes 3}} \\
&\qquad \le c^{2p/3}_{p,q} \sum_{i:t_i\in D[s,t]} |h|^p_{q\text{-var};[t_t,t_{i+1}]} \\
&\qquad \le c^{2p/3}_{p,q} |h|^p_{q\text{-var};[s,t]}.
\end{aligned}
\tag{3.9}
$$

Second, we analyze the mixed terms in (3.7). The strategy here as before is to use Young's inequality. For $j=1$ or $j=3$ we have

$$
\begin{aligned}
&\sum_{i:t_i\in D[s,t]} |B^j_{t_i,t_{i+1}}|^{p/3}_{(\mathbb{R}^d)^{\otimes 3}} \\
&\quad \le c^{p/3}_{p/2,q} \sum_{i:t_i\in D[s,t]} \|\mathbf{x}\|^{2p/3}_{p\text{-var};[t_i,t_{i+1}]} |h|^{p/3}_{q\text{-var};[t_i,t_{i+1}]} \\
&\quad \le c^{p/3}_{p/2,q} \sum_{i:t_i\in D[s,t]} \left[\frac{2}{3}\|\mathbf{x}\|^p_{p\text{-var};[t_i,t_{i+1}]} + \frac{1}{3}|h|^p_{q\text{-var};[t_i,t_{i+1}]}\right] \\
&\quad \le \frac{2}{3} c^{p/3}_{p/2,q} \big[\|\mathbf{x}\|^p_{p\text{-var};[s,t]} + |h|^p_{q\text{-var};[s,t]}\big].
\end{aligned}
\tag{3.10}
$$

A similar calculation yields

$$\sum_{i:t_i\in D[s,t]} \left|B^2_{t_i,t_{i+1}}\right|^{p/3}_{(\mathbb{R}^d)^{\otimes 3}} \le c^{2p/3}_{p,q}\left[\frac{2}{3}\|\mathbf{x}\|^p_{p\text{-var};[s,t]} + \frac{1}{3}|h|^p_{q\text{-var};[s,t]}\right]. \tag{3.11}$$

Finally we have for $j = 1, 2, 3$

$$\begin{aligned}
&\sum_{i:t_i\in D[s,t]} \left|C^j_{t_i,t_{i+1}}\right|^{p/3}_{(\mathbb{R}^d)^{\otimes 3}} \\
&\qquad \le c^{2p/3}_{p,q} \sum_{i:t_i\in D[s,t]} \left|\mathbf{x}^1\right|^{p/3}_{p\text{-var};[t_i,t_{i+1}]}|h|^{2p/3}_{q\text{-var};[t_i,t_{i+1}]} \\
&\qquad \le c^{2p/3}_{p,q} \sum_{i:t_i\in D[s,t]} \left[\frac{1}{3}\|\mathbf{x}\|^p_{p\text{-var};[t_i,t_{i+1}]} + \frac{2}{3}|h|^p_{q\text{-var};[t_i,t_{i+1}]}\right] \\
&\qquad \le c^{2p/3}_{p,q}\left[\frac{1}{3}\|\mathbf{x}\|^p_{p\text{-var};[s,t]} + \frac{2}{3}|h|^p_{q\text{-var};[s,t]}\right].
\end{aligned} \tag{3.12}$$

Using the fact that the estimates (3.8), (3.9), (3.10), (3.11) and (3.12) are uniform over all partitions, we derive the the elementary bound

$$\left|(T_h\mathbf{x})^3\right|^{p/3}_{p/3\text{-var};[s,t]} \le 8^{p/3-1}\left[(3.8) + (3.9) + (3.10) + (3.11) + (3.12)\right].$$

This then yields

$$\begin{aligned}
&\left|(T_h\mathbf{x})^3\right|^{p/3}_{p/3\text{-var};[s,t]} \\
&\qquad \le 8^{p/3-1}\left[\frac{4}{3}c^{p/3}_{p/2,q} + \frac{10}{3}c^{2p/3}_{p,q}\right]\left[\|\mathbf{x}\|^p_{p\text{-var};[s,t]} + |h|^p_{q\text{-var};[s,t]}\right] \\
&\qquad \le 8^{(p-1)/3}\left[c^{p/3}_{p/2,q} + c^{2p/3}_{p,q}\right]\left[\|\mathbf{x}\|^p_{p\text{-var};[s,t]} + |h|^p_{q\text{-var};[s,t]}\right] \\
&\qquad =: C_3(p,q)\left[\|\mathbf{x}\|^p_{p\text{-var};[s,t]} + |h|^p_{q\text{-var};[s,t]}\right].
\end{aligned} \tag{3.13}$$

Putting together (3.2), (3.6) and (3.13) we establish that

$$\|T_h\mathbf{x}\|^p_{p\text{-var};[s,t]} \le \sum_{i=1,2,3}\{C_i(p,q)\}\left[\|\mathbf{x}\|^p_{p\text{-var};[s,t]} + |h|^p_{q\text{-var};[s,t]}\right],$$

and since $\sum_{i=1,2,3}\{C_i(p,q)\} = 2^{p-1} + 4^{(p-1)/2}c^{p/2}_{p,q} + 8^{(p-1)/3}[c^{p/3}_{p/2,q} + c^{2p/3}_{p,q}]$ the estimate follows. $\square$

4. Deterministic estimates for solutions to RDEs. In this section we will develop the pathwise estimate obtained in the previous section. To assist with the clarity of the presentation it will be important to first introduce some definitions of the main objects featuring in our discussion.

NOTATION 1. If $\mathbf{x}$ is a weakly geometric p-rough path, then we will let $\omega_{\mathbf{x},p}$ denote the control which is induced by $\mathbf{x}$ in the sense that

$$\omega_{\mathbf{x},p}(s,t) \equiv \|\mathbf{x}\|_{p\text{-var};[s,t]}^{p}.$$

DEFINITION 4.1. Let $\alpha > 0$ and $I \subseteq \mathbb{R}$ be a compact interval. Suppose that $\omega : I \times I \to \mathbb{R}^{+}$ is a control. We define the accumulated α-local ω-variation by

$$M_{\alpha,I}(\omega) = \sup_{\substack{D(I)=(t_i) \\ \omega(t_i,t_{i+1}) \leq \alpha}} \sum_{i:t_i \in D(I)} \omega(t_i, t_{i+1}).$$

REMARK 4.2. Note that we have the scaling property $\beta M_{\alpha/\beta,I}(\omega) = M_{\alpha,I}(\beta\omega)$ for any $\beta > 0$.

Of special interest is the case when the control is induced (in the sense of the above notation) by a (weakly) geometric p-rough path.

DEFINITION 4.3. Let $\alpha > 0$ and $I \subseteq \mathbb{R}$ be a compact interval. We define the accumulated α-local p-variation to be the nonnegative function $M_{\alpha,I,p}$ which acts on weakly geometric p-rough paths (parameterized over I) by

$$M_{\alpha,I,p}(\mathbf{x}) \equiv M_{\alpha,I}(\omega_{\mathbf{x},p}). \tag{4.1}$$

REMARK 4.4. The function $M_{\alpha,I,p}$ is well-defined because the super-additivity of the control $\omega_{\mathbf{x},p}$ ensures that

$$M_{\alpha,I,p}(\mathbf{x}) \leq \|\mathbf{x}\|_{p\text{-var};I}^{p} < \infty$$

for any weakly geometric rough path $\mathbf{x}$ (again, parameterized over I). $M_{\alpha,I,p}(\mathbf{x})$ is continuous and increasing in α, and it equals $\|\mathbf{x}\|_{p\text{-var};I}^{p}$ whenever $\alpha \geq \|\mathbf{x}\|_{p\text{-var};I}^{p}$.

The following lemma shows how the α-local ω-variation can be used to derive Lipschitz bounds on solutions to RDEs.[5]

LEMMA 4.5. *Assume* $\lfloor p \rfloor + 1 \geq \gamma > p \geq 1$. *Suppose that* $\mathbf{x}$ *is a weakly geometric* p*-rough path parameterized on* $[0,T]$, *and* $V = (V^1, \ldots, V^d)$ *is a collection of* Lip-γ *vector fields on* $\mathbb{R}^e$. *Let* $(U_{t\leftarrow 0}^{\mathbf{x}}(\cdot))_{t\in[0,T]}$ *denote the flow induced by the RDE*

$$dy_t = V(y_t)\, d\mathbf{x}_t, \qquad y(0) = y_0,$$

[5] It should be compared with Theorem 10.26 of [12], on which the proof is based.

so that $U^{\mathbf{x}}_{\cdot\leftarrow 0}(y_0) \equiv y$. *If* ω *is the control* $\omega(u,v) \equiv |V|^p_{\text{Lip-}\gamma}\omega_{\mathbf{x},p}(u,v)$, *then for any* y^1_0 *and* y^2_0 *in* $\mathbb{R}^e$, *any* $\alpha > 0$ *and any* $[s,t] \subseteq [0,T]$, *we have*

$$\begin{aligned}&|U^{\mathbf{x}}_{\cdot\leftarrow 0}(y^1_0) - U^{\mathbf{x}}_{\cdot\leftarrow 0}(y^2_0)|_{p\text{-var};[s,t]}\\&\quad \le C|V|_{\text{Lip-}\gamma}\|\mathbf{x}\|_{p\text{-var};[s,t]}|y^1_0 - y^2_0|\exp[C\max(1,\alpha^{-1})M_{\alpha,[0,T]}(\omega)],\end{aligned} \tag{4.2}$$

where C *is a constant depending only on* p.

PROOF. Let $y^i_t \equiv U^{\mathbf{x}}_{t\leftarrow 0}(y^i_0)$ for $i = 1, 2$. We follow through the details in the proof of Theorem 10.26 of [12], with the exception that we enhance each application of their Lemma 10.63 by instead using Remark 10.64 of the same reference. The contents of this remark can be improved so that we use an arbitrary truncation parameter α rather than setting $\alpha = 1$; the details are easily checked and we omit them. These calculations result in the following estimate:

$$|y^1_{s,t} - y^2_{s,t}| \le C|y^1_0 - y^2_0|\omega(s,t)^{1/p}\exp[C\max(1,\alpha^{-1})M_{\alpha,[0,T]}(\omega)]. \tag{4.3}$$

The estimate (4.2) then follows by an elementary computation. □

Using these Lipschitz estimates on the flow, it is a relatively simple matter to derive growth bounds on the Jacobian. This is the content of the following corollary.

COROLLARY 4.6. *Assume* $\lfloor p\rfloor + 1 > \gamma > p \ge 1$. *Suppose that* $\mathbf{x}$ *is a weakly geometric* p-*rough path parameterized on* $[0,T]$, *and* $V = (V^1,\dots,V^d)$ *is a collection of* Lip-γ *vector fields on* $\mathbb{R}^e$. *Let* $(U^{\mathbf{x}}_{t\leftarrow 0}(\cdot))_{t\in[0,T]}$ *denote the flow induced by the RDE*

$$dy_t = V(y_t)\,d\mathbf{x}_t, \qquad y(0) = y_0.$$

Then the derivative $J^{\mathbf{x}}_{t\leftarrow 0}(y_0)$ *of* $U^{\mathbf{x}}_{t\leftarrow 0}(y_0)$ *exists and satisfies the growth-bound*

$$\begin{aligned}&|J^{\mathbf{x}}_{\cdot\leftarrow 0}(y_0)|_{p\text{-var};[0,T]}\\&\quad \le C|V|_{\text{Lip-}\gamma}\|\mathbf{x}\|_{p\text{-var};[0,T]}\exp[C\max(|V|^p_{\text{Lip-}\gamma},\alpha^{-1})M_{\alpha,[0,T],p}(\mathbf{x})].\end{aligned} \tag{4.4}$$

PROOF. It is well known [12] under these hypotheses that $U^{\mathbf{x}}_{\cdot\leftarrow 0}(y_0)$ is differentiable. Fix $\alpha > 0$ and define

$$\delta := \alpha|V|^p_{\text{Lip-}\gamma}.$$

Let h be in $\mathbb{R}^e$, and for a real number ε let $y^1_0 = y_0 + \varepsilon h$ and $y^2_0 = y_0$. Take $U^{\mathbf{x}}_{\cdot\leftarrow 0}(y^i_0) \equiv y^i$ for $i = 1, 2$. Applying the previous lemma we obtain that for any $[s,t] \subseteq [0,T]$

$$|y^1_{s,t} - y^2_{s,t}|^p \le C^p|V|^p_{\text{Lip-}\gamma}\|\mathbf{x}\|^p_{p\text{-var};[s,t]}\varepsilon^p|h|^p\exp[Cp\max(1,\delta^{-1})M_{\delta,[0,T]}(\omega)].$$

Dividing by ε^p, taking the limit as $\varepsilon \downarrow 0$ and then taking the supremum over all $|h| = 1$ this estimate becomes

$$\begin{aligned}\tag{4.5} &\big|J^{\mathbf{x}}_{t\leftarrow 0}(y_0) - J^{\mathbf{x}}_{s\leftarrow 0}(y_0)\big|^p \\ &\quad \le C^p |V|^p_{\text{Lip-}\gamma} \|\mathbf{x}\|^p_{p\text{-var};[s,t]} \exp\big[Cp \max\big(1, \delta^{-1}\big) M_{\delta,[0,T]}(\omega)\big].\end{aligned}$$

Fix an arbitrary partition D of $[0, T]$. Then by summing the terms in (4.5) and using the super-additivity of $\omega_{\mathbf{x},p}$ it follows that

$$\begin{aligned}&\bigg(\sum_{i:t_i\in D} \big|J^{\mathbf{x}}_{t_{i+1}\leftarrow 0}(y_0) - J^{\mathbf{x}}_{t_i\leftarrow 0}(y_0)\big|^p\bigg)^{1/p} \\ &\quad \le C|V|_{\text{Lip-}\gamma} \|\mathbf{x}\|t_{p\text{-var};[0,T]} \exp\big[C \max\big(1, \delta^{-1}\big) M_{\delta,[0,T]}(\omega)\big].\end{aligned}$$

To finish the proof we first optimize over all partitions D to give an estimate on the p-variation. We then use the scaling property in Remark 4.2 and the definition of $M_{\alpha,I,p}(\mathbf{x})$ to obtain that

$$M_{\delta,[0,T]}(\omega) = \frac{\delta}{\alpha} M_{\alpha,[0,T]}\bigg(\frac{\alpha}{\delta}\omega\bigg) = |V|^p_{\text{Lip-}\gamma} M_{\alpha,I,p}(\mathbf{x}).$$

Putting everything together gives (4.4). $\square$

We have succeeded in showing how the derivative of the flow can be controlled by using the function $M_{\alpha,I,p}(\cdot)$. But it is still not obvious how to get a handle on the tail behavior of $M_{\alpha,I,p}(\cdot)$ when we evaluate it at a Gaussian p-rough path. To expose the structure further, we will now consider another function $N_{\alpha,I,p}(\cdot)$ on $WG\Omega_p(\mathbb{R}^d)$, which is closely related to $M_{\alpha,I,p}(\cdot)$. The following sequence will play an important role in enabling us to achieve this.

DEFINITION 4.7 (The greedy sequence). Assume $\mathbf{x} \in WG\Omega_p(\mathbb{R}^d)$ is parameterized over a compact interval I. If $\alpha > 0$ we define a nondecreasing sequence $(\tau_i(\alpha, p, \mathbf{x}))_{i=0}^{\infty} = (\tau_i(\alpha))_{i=0}^{\infty}$ in I in by

$$\begin{aligned}\tag{4.6} \tau_0(\alpha) &= \inf I, \\ \tau_{i+1}(\alpha) &= \inf\big\{t : \|\mathbf{x}\|^p_{p\text{-var};[\tau_i,t]} \ge \alpha, \tau_i(\alpha) < t \le \sup I\big\} \wedge \sup I,\end{aligned}$$

with the convention that $\inf \varnothing = +\infty$. We call this sequence the greedy sequence.

REMARK 4.8. Note that for $\tau_i(\alpha) < \sup I$ and $\|\mathbf{x}\|^p_{p\text{-var};[\tau_i(\alpha),\sup I]} \ge \alpha$, $\tau_{i+1}(\alpha)$ is intuitively the first time $\|\mathbf{x}\|^p_{p\text{-var};[\tau_i(\alpha),\cdot]}$ reaches α (recall that the p-variation is a continuous function).

We want to show that the greedy sequence is actually a partition of I; in other words it has only a finite number of distinct terms which include the endpoints. With this objective in mind we introduce the function $N_{\alpha,I,p} : WG\Omega_p(\mathbb{R}^d) \to \mathbb{R}_+$ given by

$$N_{\alpha,I,p}(\mathbf{x}) := \sup\{n \in \mathbb{N} \cup \{0\} : \tau_n(\alpha) < \sup I\}. \tag{4.7}$$

We note that $N_{\alpha,I,p}$ describes the size of the nontrivial part of the sequence $(\tau_i(\alpha))_{i=0}^{\infty}$. More precisely, the number of distinct terms in the sequence $(\tau_i(\alpha))_{i=0}^{\infty}$ equals $N_{\alpha,I,p}(\mathbf{x}) + 1$. The partition of the interval given by

$$\{\tau_i(\alpha) : i = 0, 1, \ldots, N_{\alpha,I,p}(\mathbf{x}) + 1\}$$

can now heuristically be thought of as a "greedy" approximation to the supremum in identity (4.1), the definition of the accumulated α-local p-variation.

LEMMA 4.9. *For any $\alpha > 0$, $p \geq 1$ and any compact interval I the function $N_{\alpha,I,p} : WG\Omega_p(\mathbb{R}^d) \to \mathbb{R}_+$ is well defined; that is, $N_{\alpha,I,p}(\mathbf{x}) < \infty$ whenever $\mathbf{x}$ is in $WG\Omega_p(\mathbb{R}^d)$.*

PROOF. From the continuity of $\|\mathbf{x}\|_{p\text{-var};[s,\cdot]}$ we can deduce that

$$\|\mathbf{x}\|^p_{p\text{-var};[\tau_{i-1}(\alpha),\tau_i(\alpha)]} = \alpha \qquad \text{for } i = 1, 2, \ldots, N_{\alpha,I,p}(\mathbf{x}).$$

Thus, the super-additivity of $\omega_{\mathbf{x},p}$ implies that if $\mathbf{x}$ is in $WG\Omega_p(\mathbb{R}^d)$, then

$$\begin{aligned}\alpha N_{\alpha,I,p}(\mathbf{x}) &= \sum_{i=1}^{N_{\alpha,I,p}(\mathbf{x})} \omega_{\mathbf{x},p}(\tau_{i-1}(\alpha), \tau_i(\alpha)) \leq \omega_{\mathbf{x},p}(0, \tau_{N_{\alpha,I,p}(\mathbf{x})}(\alpha)) \\ &\leq \|\mathbf{x}\|^p_{p\text{-var};[0,T]} < \infty. \qquad \square\end{aligned}$$

COROLLARY 4.10. *Let $\mathbf{x}$ be a path in $WG\Omega_p(\mathbb{R}^d)$ and suppose $\alpha > 0$. Define the sequence $(\tau_i(\alpha))_{i=0}^{\infty}$ by (4.6), and let $N_{\alpha,I,p}(\mathbf{x})$ be given by (4.7). Then the set*

$$D_\tau = \{\tau_i(\alpha) : i = 0, 1, \ldots, N_{\alpha,I,p}(\mathbf{x}) + 1\}$$

is a partition of I.

PROOF. This now follows immediately from the definition of $(\tau_i(\alpha))_{i=0}^{\infty}$ and the fact that $N_{\alpha,I,p}(\mathbf{x})$ is finite. $\square$

The following proposition shows how we can use $N_{\alpha,I,p}(\mathbf{x})$ to bound on the α-local p-variation.

PROPOSITION 4.11. *Let $p \geq 1$ and suppose $\mathbf{x}$ is a path in $WG\Omega_p(\mathbb{R}^d)$ parameterized over the compact interval I, and then for every $\alpha > 0$*

$$M_{\alpha,I,p}(\mathbf{x}) \leq (2N_{\alpha,I,p}(\mathbf{x}) + 1)\alpha.$$

PROOF. First note that the case $N_{\alpha,I,p}(\mathbf{x}) = 0$ can be dealt with trivially. We therefore can (and will) assume in the following that $N_{\alpha,I,p}(\mathbf{x}) \geq 1$. Let $D = \{t_i : i = 0, 1, \ldots, n\}$ be any partition of I with the property that

$$\omega_{\mathbf{x},p}(t_{i-1}, t_i) \leq \alpha \qquad \text{for all } i = 1, \ldots, n. \tag{4.8}$$

Corollary 4.10 ensures that D_τ is a partition of I. We relabel the points in D with reference to the partition D_τ by writing $t_i = t_j^l$ for $i = 1, 2, \ldots, n$, where l indicates which of disjoint subintervals $\{(\tau_i(\alpha), \tau_{i+1}(\alpha)] : i = 0, 1, \ldots, N_{\alpha,I,p}(\mathbf{x})\}$ contains t_i, and j orders the t_is within each of these subintervals. More precisely, $l \in \{0, 1, \ldots, N_{\alpha,I,p}(\mathbf{x})\}$ is the unique natural number such that

$$\tau_l(\alpha) < t_i \leq \tau_{l+1}(\alpha);$$

and then $j \geq 1$ is well defined by

$$j = i - \max_{t_r \leq \tau_l(\alpha)} r.$$

For each $l \in \{0, 1, \ldots, N_{\alpha,I,p}(\mathbf{x})\}$ let n_l denote the number of elements of D in $(\tau_l(\alpha), \tau_{l+1}(\alpha)]$. Suppose now for a contradiction that $n_l = 0$. In this case, $t_{n_{l-1}}^{l-1}$ and t_1^{l+1} are two consecutive points of D with $t_{n_{l-1}}^{l-1} \leq \tau_l(\alpha) < \tau_{l+1}(\alpha) < t_1^{l+1}$, and since the $(\tau_i(\alpha))_{i=0}^{\infty}$ are defined to be maximal [recall (4.6)] we have

$$\omega_{\mathbf{x},p}\big(t_{n_{l-1}}^{l-1}, t_1^{l+1}\big) > \omega_{\mathbf{x},p}\big(\tau_l(\alpha), \tau_{l+1}(\alpha)\big) = \alpha.$$

This contradicts the assumptions on D (4.8). We deduce that $n_l \geq 1$.

We observe that if $n_l \geq 2$, then the super-additivity of $\omega_{\mathbf{x},p}$ results in

$$\sum_{j=1}^{n_l-1} \omega_{\mathbf{x},p}\big(t_j^l, t_{j+1}^l\big) \leq \omega_{\mathbf{x},p}\big(t_1^l, t_{n_l}^l\big) \qquad \text{for } l = 0, 1, \ldots, N_{\alpha,I,p}(\mathbf{x});$$

thus, by a simple calculation we have

$$\begin{aligned}
&\sum_{j=1}^{n} \omega_{\mathbf{x},p}(t_{j-1}, t_j) \\
&\quad \leq \sum_{l=0}^{N_{\alpha,I,p}(\mathbf{x})-1} \Big\{\big[\omega_{\mathbf{x},p}\big(t_{n_l}^l, t_1^{l+1}\big) + \omega_{\mathbf{x},p}\big(t_1^{l+1}, t_{n_{l+1}}^{l+1}\big)\big] 1_{\{n_{l+1} \geq 2\}} \\
&\qquad\qquad\qquad + \omega_{\mathbf{x},p}\big(t_{n_l}^l, t_{n_{l+1}}^{l+1}\big) 1_{\{n_{l+1} = 1\}}\Big\} \\
&\qquad + \omega_{\mathbf{x},p}\big(0, t_{n_0}^0\big).
\end{aligned} \tag{4.9}$$

To complete the proof we note that $\omega_{\mathbf{x},p}(t_1^{l+1}, t_{n_{l+1}}^{l+1}) \leq \alpha$ and $\omega_{\mathbf{x},p}(0, t_{n_0}^0) \leq \alpha$ by the definition of the sequence (t_j^l). Furthermore we have $\omega_{\mathbf{x},p}(t_{n_l}^l, t_1^{l+1}) \leq \alpha$

because $t^l_{n_l}$ and t^{l+1}_1 are two consecutive points in D. Hence, we may deduce from (4.9) that

$$\sum_{j=1}^{n} \omega_{\mathbf{x},p}(t_{j-1},t_j) \le \left(2N_{\alpha,I,p}(\mathbf{x})+1\right)\alpha.$$

Because the right-hand side of the last inequality does not depend on D, optimizing over all such partitions gives the stated result. □

As a direct consequence of Proposition 4.11 and Corollary 4.6 we have the estimate

$$\begin{aligned}\left|J^{\mathbf{x}}_{\cdot\leftarrow 0}(y_0)\right|_{p\text{-var};[0,T]} &\le C|V|_{\text{Lip-}\gamma}\|\mathbf{x}\|_{p\text{-var};[0,T]} \\ &\quad\times \exp\left[C\max\left(1,\alpha|V|^p_{\text{Lip-}\gamma}\right)\left(2N_{\alpha,[0,T],p}(\mathbf{x})+1\right)\right].\end{aligned} \tag{4.10}$$

If we take $\mathbf{x}=\mathbf{X}$ to be Gaussian rough path, then the tail of the Jacobian can be studied via the tail of $N_{\alpha,I,p}(\mathbf{X})$. This will be the objective of the remainder of the paper.

REMARK 4.12. There are several ways to obtain bounds for the Jacobian in terms of $N_{\alpha,I,p}(\mathbf{x})$. An alternative approach suggested by the anonymous referee uses the Gronwall estimate

$$\left|J^{\mathbf{x}}_{\cdot\leftarrow 0}(y_0)\right|_{\infty;[0,T]} \le C\exp\left(C\|X\|^p_{p\text{-var};[0,T]}\right).$$

Using the cocycle property

$$J^{\mathbf{x}}_{t\leftarrow 0}(y_0) = J^{\mathbf{x}}_{t\leftarrow s}\left(U^{\mathbf{x}}_{s\leftarrow 0}(y_0)\right)J^{\mathbf{x}}_{s\leftarrow 0}(y_0)$$

a simple induction argument gives the following bound on the infinity norm of the Jacobian:

$$\left|J^{\mathbf{x}}_{\cdot\leftarrow 0}(y_0)\right|_{\infty;[0,T]} \le C\exp\left(C\alpha^p N_{\alpha,I,p}(\mathbf{x})\right).$$

This argument may be generalized to cover the p-variation of the Jacobian; see, for example, [8] where the authors implement a variant of this idea based on a previous version of this paper.

5. Gaussian rough paths. The previous section developed the pathwise estimates on $J^{\mathbf{x}}_{t\leftarrow 0}(y_0)$ we need. We learned that the p-variation of $J^{\mathbf{x}}_{t\leftarrow 0}(y_0)$ can be bounded explicitly in terms of $N_{\alpha,[0,T],p}(\mathbf{x})$. The importance of controlling $J^{\mathbf{x}}_{t\leftarrow 0}(y_0)$ using $N_{\alpha,[0,T],p}(\mathbf{x})$, as opposed to simpler alternatives [see, e.g., identity (1.2)], is best appreciated when the driving rough path is taken to be random. Henceforth, we will distinguish situations where the path is random by writing it in upper-case: $\mathbf{X}$. Of special interest is when $\mathbf{X}$ is the lift[6] of some continuous $\mathbb{R}^d$-valued Gaussian process $(X_t)_{t\in I}$. A theory of such Gaussian rough paths has been developed by a succession of authors [3, 5, 7, 10], and we will mostly work within their framework.

To be more precise, we will assume that $X_t = (X_t^1, \dots, X_t^d)$ is a continuous, centered (i.e., mean zero) Gaussian process with independent and identically distributed components. Let $R : I \times I \to \mathbb{R}$ denote the covariance function of any component, that is,

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

Throughout we will assume that this process is realized on the abstract Wiener space $(\mathcal{W}, \mathcal{H}, \mu)$ where $\mathcal{W} = C_0(I, \mathbb{R}^d)$, the space of continuous $\mathbb{R}^d$-valued functions on I. More precisely we mean that X is the canonical process on $\mathcal{W}$; that is, $X_t(\omega) = \omega(t)$, and $(X_t)_{t\in I}$ has the required Gaussian distribution under μ. We recall the notion of the "rectangular increments of R" from [11]; these are defined by

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

The existence of a lift for X is guaranteed by insisting on a sufficient rate of decay on 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.

CONDITION 1. *There exists of* $1 \le \rho < 2$ *such that* R *has finite* ρ*-variation in the sense*

$$V_\rho(R; I \times I) := \left(\sup_{\substack{D=(t_i)\in\mathcal{D}(I) \\ D'=(t'_j)\in\mathcal{D}(I)}} \sum_{i,j} \left| R\begin{pmatrix} t_i, t_{i+1} \\ t'_j, t'_{j+1} \end{pmatrix} \right|^\rho \right)^{1/\rho} < \infty. \tag{5.1}$$

REMARK 5.1. Under Theorem 35, Condition 1 of [10], $(X_t)_{t\in[0,T]}$ lifts to a geometric p-rough path for any $p > 2\rho$. Moreover, there is a unique *natural lift* which is the limit (in the $d_{p\text{-var}}$-induced rough path topology) of the canonical lift of piecewise linear approximations to X.

The following theorem appears in [10] as Proposition 17; cf. also the recent note [11]. It shows how the assumption $V_\rho(R; [0,T]^2) < \infty$ allows us to embed $\mathcal{H}$ in the space of continuous paths with finite ρ variation. The result, as it appears in [10], applies to one-dimensional Gaussian processes. The generalization to arbitrary finite dimensions is straightforward, and we will not elaborate on the proof.

[6]Recall that by $\mathbf{X}$ being a lift of X, we mean that the projection of $\mathbf{X}$ to the first tensor level is exactly X.

THEOREM 5.2 ([10]). *Let $(X_t)_{t\in I} = (X_t^1, \dots, X_t^d)_{t\in I}$ be a continuous, 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 $C^{\rho\text{-}\mathrm{var}}(I, \mathbb{R}^d)$; in fact,*

$$|h|_{\mathcal{H}} \geq \frac{|h|_{\rho\text{-}\mathrm{var};I}}{\sqrt{V_\rho(R; I \times I)}}. \tag{5.2}$$

REMARK 5.3 ([9]). Writing $\mathcal{H}^H$ for the Cameron–Martin space of fBm for H in $(1/4, 1/2)$, the variation embedding in [9] gives the stronger result that

$$\mathcal{H}^H \hookrightarrow C^{q\text{-}\mathrm{var}}(I, \mathbb{R}^d) \qquad \text{for any } q > (H + 1/2)^{-1}.$$

Once we have established a lift $\mathbf{X}$ of X we will often want to make sense of $\mathbf{X}(\omega + h)$. The main technique used for achieving this is to relate it to the translated rough path $T_h\mathbf{x}$; recall Section 3. The the following result appeared in [3] and demonstrates that, under certain conditions, $\mathbf{X}(\omega + h)$ and $T_h\mathbf{X}(\omega)$ are equal for all h in $\mathcal{H}$ on a set of μ-full measure.

LEMMA 5.4. *Let $(X_t)_{t\in I} = (X_t^1, \dots, X_t^d)_{t\in I}$ be a mean-zero Gaussian process with i.i.d. components. Assume that X has a natural lift to a geometric p-rough path. Assume further that for some $q \geq 1$ such that $1/p + 1/q > 1$, we have $\mathcal{H} \hookrightarrow C^{q\text{-}\mathrm{var}}(I, \mathbb{R}^d)$. Then there exists a measurable subset $E \subseteq \mathcal{W}$ with $\mu(E) = 1$, such that for all ω in E, we have*

$$T_h\mathbf{X}(\omega) \equiv \mathbf{X}(\omega + h) \qquad \textit{for all } h \textit{ in } \mathcal{H}.$$

From the different choices of p and q with the properties that X lifts path in $G\Omega_p(\mathbb{R}^d)$ and $\mathcal{H}$ continuously embeds in $C^{q\text{-}\mathrm{var}}(I, \mathbb{R}^d)$, it will often prove useful to work with a particular choice that satisfies certain constraints. The purpose of the next lemma is to show that these constraints can always be satisfied for some p and q, for the examples of Gaussian processes that will interest us most.

COROLLARY 5.5. *Let $(X_t)_{t\in I} = (X_t^1, \dots, X_t^d)_{t\in I}$ be a continuous, mean-zero Gaussian process with i.i.d. components on $(\mathcal{W}, \mathcal{H}, \mu)$. Suppose that at least one of the following holds*:

(1) *For some ρ in $[1, \frac{3}{2})$ the covariance function of X has finite ρ-variation, in the sense of Condition 1*;

(2) *X is a fractional Brownian motion for H in $(1/4, 1/2)$.*

Then there exist real numbers p, q such that the following statements are true simultaneously:

(1) *X has a natural lift to a geometric p-rough path*;
(2) $\mathcal{H} \hookrightarrow C^{q\text{-var}}(I, \mathbb{R}^d)$ *where* $1/p + 1/q > 1$.

PROOF. If Condition 1 is satisfied with $\rho \in [1, 3/2)$, then (taking $\frac{1}{0} := \infty$)

$$2\rho < 3 < \frac{\rho}{\rho - 1}.$$

If we therefore set $q = \rho$ and choose p in $(2q, 3)$, Remark 5.1 guarantees the existence of a natural lift for X. Furthermore, Theorem 5.2 ensures that $\mathcal{H} \hookrightarrow C^{q\text{-var}}(I, \mathbb{R}^d)$.

In the case where X is fBm let $4 > p > \frac{1}{H}$, and then Remark 5.1 guarantees that X lifts to a geometric p-rough path. Let $q = (\frac{1}{p} + \frac{1}{2})^{-1}$. Then we have

$$\left(H + \frac{1}{2}\right)^{-1} < q$$

and

$$\frac{1}{p} + \frac{1}{q} = \frac{2}{p} + \frac{1}{2} > 1.$$

The fact that $\mathcal{H} \hookrightarrow C^{q\text{-var}}(I, \mathbb{R}^d)$ now follows by Remark 5.3. $\square$

6. The tail behavior of $N_{\alpha,I,p}(\mathbf{X}(\cdot))$ via Gaussian isoperimetry. We continue to work in the setting of an abstract Wiener space $(\mathcal{W}, \mathcal{H}, \mu)$. If $\mathcal{K}$ denotes the unit ball in $\mathcal{H}$, then for any $A \subseteq \mathcal{W}$ we can consider the Minkowski sum

$$A + r\mathcal{K} := \{x + ry : x \in A, y \in \mathcal{K}\}.$$

We then recall the following isoperimetric inequality of C. Borell; cf. Theorem 4.3 of [22].

THEOREM 6.1 (Borell). *Let $(\mathcal{W}, \mathcal{H}, \mu)$ be an abstract Wiener space and $\mathcal{K}$ denote the unit ball in $\mathcal{H}$. Suppose A is a Borel subset of $\mathcal{W}$ such that $\mu(A) \geq \Phi(a)$ for some real number a. Then for every $r \geq 0$,*

$$\mu_*(A + r\mathcal{K}) \geq \Phi(a + r),$$

where μ_* is the inner measure of μ, and Φ denotes the standard normal cumulative distribution function.

The next proposition is crucial. It will allow us to apply Borell's inequality to control the tail of the random variable $N_{\alpha,I,p}(\mathbf{X}(\omega))$.

PROPOSITION 6.2. *Let $(X_t)_{t\in I} = (X_t^1, \dots, X_t^d)_{t\in I}$ be a continuous, mean-zero Gaussian process, parameterized over a compact interval I on the abstract Wiener space $(\mathcal{W}, \mathcal{H}, \mu)$. Suppose that p and q are real numbers such that $1 \leq p < 4$ and $1/p + 1/q > 1$. Assume further that*:

(1) *X has a natural lift to a geometric p-rough path $\mathbf{X}$*;
(2) $\mathcal{H} \hookrightarrow C^{q\text{-var}}(I, \mathbb{R}^d)$.

Then there exists a set $E \subseteq \mathcal{W}$, of μ-full measure, with the following property: for all ω in E, h in $\mathcal{H}$ and $\alpha > 0$, if

$$\|\mathbf{X}(\omega - h)\|_{p\text{-var};I} \leq \alpha,$$

then

$$|h|_{q\text{-var};I} \geq \alpha N_{\tilde{\alpha}^p,I,p}(\mathbf{X}(\omega))^{1/q},$$

where $C_{p,q}$ is the constant in Lemma 3.1 and $\tilde{\alpha} = (2C_{p,q})^{1/p}\alpha$.

PROOF. Fix $\alpha > 0$. We first note that the case $N_{\tilde{\alpha}^p,I,p}(\mathbf{X}(\omega)) = 0$ is trivial. Hence we will assume in the following that $N_{\tilde{\alpha}^p,I,p}(\mathbf{X}(\omega)) \geq 1$. From the definition of the sequence $(\tau_i(\tilde{\alpha}^p))_{i=0}^{\infty}$ and the integer $N_{\tilde{\alpha}^p,I,p}(\mathbf{X}(\omega))$ we have for $i = 0, 1, 2, \ldots, N_{\tilde{\alpha}^p,I,p}(\mathbf{X}(\omega)) - 1$

$$\|\mathbf{X}(\omega)\|_{p\text{-var};[\tau_i(\tilde{\alpha}^p),\tau_{i+1}(\tilde{\alpha}^p)]} = \tilde{\alpha}. \tag{6.1}$$

Consider the (measurable) subset of $\mathcal{W}$ defined by

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

and recall from Lemma 5.4 that $\mu(E) = 1$. For every ω in E define a subset $F_{\alpha,\omega}$ of $\mathcal{H}$ by

$$F_{\alpha,\omega} := \{h \in \mathcal{H} : \|\mathbf{X}(\omega - h)\|_{p\text{-var};I} \leq \alpha\}.$$

Using the estimate in Lemma 3.1 we have for any ω in E

$$\begin{aligned}\|\mathbf{X}(\omega)\|^p_{p\text{-var};[\tau_i(\tilde{\alpha}^p),\tau_{i+1}(\tilde{\alpha}^p)]} &= \|T_h\mathbf{X}(\omega - h)\|^p_{p\text{-var};[\tau_i(\tilde{\alpha}^p),\tau_{i+1}(\tilde{\alpha}^p)]} \\ &\leq C_{p,q}\big(\|T_h\mathbf{X}(\omega)\|^p_{p\text{-var};I} + |h|^p_{q\text{-var};[\tau_i(\tilde{\alpha}^p),\tau_{i+1}(\tilde{\alpha}^p)]}\big).\end{aligned}$$

Hence, for any ω in E, h in $F_{\alpha,\omega}$ we have

$$\tilde{\alpha}^p \leq C_{p,q}\big(\alpha^p + |h|^p_{q\text{-var};[\tau_i(\tilde{\alpha}^p),\tau_{i+1}(\tilde{\alpha}^p)]}\big). \tag{6.2}$$

Substituting $(2C_{p,q})^{1/p}\alpha$ for $\tilde{\alpha}$, estimate (6.2) becomes

$$|h|^q_{q\text{-var};[\tau_i(\tilde{\alpha}^p),\tau_{i+1}(\tilde{\alpha}^p)]} \geq \alpha^q.$$

Summing over $i = 0, 1, \ldots, N_{\tilde{\alpha},I,p}(\mathbf{X}(\omega)) - 1$ then gives

$$|h|^q_{q\text{-var};I} \geq \sum_{i=0}^{N_{\tilde{\alpha}^p,I,p}(\mathbf{X}(\omega))-1} |h|^q_{q\text{-var};[\tau_i(\tilde{\alpha}^p),\tau_{i+1}(\tilde{\alpha}^p)]} \geq \alpha^q N_{\tilde{\alpha}^p,I,p}(\mathbf{X}(\omega)),$$

which yields the desired estimate. □

By using these estimates in concert with Borell's inequality we are lead directly to the following theorem which describes the needed tail-estimate on the random variable $N_{\tilde{\alpha}^p,I,p}(\mathbf{X}(\cdot))$.

THEOREM 6.3. *Let $(X_t)_{t\in I} = (X_t^1, \dots, X_t^d)_{t\in I}$ be a continuous, mean-zero Gaussian process, parameterized over a compact interval I, on the abstract Wiener space $(\mathcal{W}, \mathcal{H}, \mu)$. Suppose that p and q are real numbers satisfying $1 \le p < 4$ and $1/p + 1/q > 1$. Assume that:*

(1) *X has a natural lift to a geometric p-rough path $\mathbf{X}$;*
(2) *$\mathcal{H} \hookrightarrow C^{q\text{-var}}(I, \mathbb{R}^d)$, so that there exists C_{emb} in $(0,\infty)$ with $|h|_{q\text{-var};I} \le C_{\mathrm{emb}}|h|_{\mathcal{H}}$ for all h in $\mathcal{H}$.*

Let $C_{p,q}$ be the constant in Lemma 3.1. Then for all $\alpha > 0$ the natural lift $\mathbf{X}$ of X to a geometric p-rough path satisfies

$$\mu\{\omega : N_{\tilde{\alpha}^p, I, p}(\mathbf{X}(\omega)) > n\} \le C_1 \exp\left[\frac{-\alpha^2 n^{2/q}}{2C_{\mathrm{emb}}^2}\right] \tag{6.3}$$

for all $n \ge 1$, where $\tilde{\alpha} = (2C_{p,q})^{1/p}\alpha$. The constant C_1 is given explicitly by

$$C_1 = \exp[2\Phi^{-1}(\mu(A_\alpha))^2], \tag{6.4}$$

where Φ^{-1} is the inverse of the standard normal cumulative distribution function and

$$A_\alpha := \{\omega \in \mathcal{W} : \|\mathbf{X}(\omega)\|_{p\text{-var};I} \le \alpha\}.$$

PROOF. By applying Proposition 6.2 together with hypothesis 2, we can deduce that

$$\{\omega : N_{\tilde{\alpha}^p, I, p}(\mathbf{X}(\omega)) > n\} \cap E \subset \mathcal{W} \backslash (A_\alpha + r_n \mathcal{K}), \tag{6.5}$$

where $E \subseteq \mathcal{W}$ with $\mu(E) = 1$ and

$$r_n := \frac{\alpha n^{1/q}}{C_{\mathrm{emb}}}.$$

Noticing that $\mu(A_\alpha) =: \Phi(a_\alpha)$ is in $(0, 1)$ [i.e., a_α is in $(-\infty, \infty)$] an application of Borell's inequality then gives that

$$\mu\{\omega : N_{\tilde{\alpha}^p, I, p}(\mathbf{X}(\omega)) > n\} \le 1 - \Phi(a_\alpha + r_n) \le \exp\left[-\frac{(a_\alpha + r_n)^2}{2}\right]. \tag{6.6}$$

If $a_\alpha > -r_n/2$, then (6.6) implies

$$\mu\{\omega : N_{\tilde{\alpha}^p, I, p}(\mathbf{X}(\omega)) > n\} \le \exp\left(-\frac{r_n^2}{8}\right).$$

Alternatively if $a_\alpha \le -r_n/2$ then $r_n^2 \le 4a_\alpha^2$, and it is easy to see that

$$\begin{aligned}\mu\{\omega : N_{\tilde{\alpha}^p, I, p}(\mathbf{X}(\omega)) > n\} &\le \exp\left(-\frac{a_\alpha^2 + 2a_\alpha r_n}{2}\right)\exp\left(-\frac{r_n^2}{2}\right)\\ &\le \exp(2a_\alpha^2)\exp\left(-\frac{r_n^2}{2}\right).\end{aligned}$$

Since $a_\alpha = \Phi^{-1}(\mu(A_\alpha))$ we have shown the required estimate (6.3). □

REMARK 6.4. Suppose that for some ρ in $[1, \frac{3}{2})$ the covariance function of X has finite ρ-variation (in the sense of Condition 1). In this case we deduce from Corollary 5.5 and Theorem 5.2 that $q = \rho$ and $p \in (2\rho, 3)$ satisfy the hypothesis of Theorem 6.3 with the embedding constant given explicitly by

$$C_{\mathrm{emb}} = \sqrt{V_\rho(R; I \times I)}.$$

Hence, the tail estimates just proved lead to moment estimates on $N_{\alpha,I,p}(\mathbf{X}(\omega))$ in the usual way. This leads to the conclusion that for any $\alpha > 0$, and η satisfying

$$\eta < \frac{\alpha^2}{2V_\rho(R; I \times I)},$$

we have

$$\int_{\mathcal{W}} \exp\bigl[\eta N_{\tilde{\alpha}^p, I, p}(\mathbf{X}(\omega))^{2/\rho}\bigr]\mu(d\omega) < \infty. \tag{6.7}$$

For the Brownian rough path (i.e., $\rho = 1$) this shows that $N_{\alpha,I,p}(\mathbf{X}(\omega))$ has a Gaussian tail since in this case we have $\log|J_{t\leftarrow 0}^{\mathbf{X}(\omega)}(y_0)| \lesssim N_{\alpha,I,p}(\mathbf{X}(\omega))$. Rudimentary Itô or Stratonovich calculus tells us that we cannot expect the tail of $N_{\alpha,I,p}(\mathbf{X}(\omega))$ to decay any faster than Gaussian, suggesting a degree of sharpness to our approach. By a similar argument, we can show that for any $r < 2/\rho$

$$\exp\bigl[N_{\alpha,I,p}(\mathbf{X}(\cdot))^r\bigr] \qquad \text{is in } \bigcap_{q>0} L^q(\mu);$$

and similar calculations can be performed in the fractional Brownian setting too.

THEOREM 6.5 (Moment estimates on the Jacobian). *Let $(X_t)_{t\in[0,T]} = (X_t^1, \dots, X_t^d)_{t\in[0,T]}$ be a continuous, mean-zero Gaussian process with i.i.d. components associated to the abstract Wiener space $(\mathcal{W}, \mathcal{H}, \mu)$. Let ρ be in $[1, \frac{3}{2})$, p in $(2\rho, 3)$ and $\gamma > p$. Suppose that the covariance function of X has finite ρ-variation in the sense of Condition 1. Then X lifts to a geometric p-rough path $\mathbf{X}$, and for any collection of* Lip-γ *vector fields $V = (V^1, \dots, V^d)$ on $\mathbb{R}^e$ with $\gamma > p$ the RDE*

$$dY_t = V(Y)\,d\mathbf{X}, \qquad Y(0) = y_0$$

has a unique solution. The flow $U_{t\leftarrow 0}^{\mathbf{X}(\omega)}(\cdot)$ induced by the solution to this RDE is differentiable. Let this derivative be given by

$$J_{t\leftarrow 0}^{\mathbf{X}(\omega)}(y_0) \cdot a := \left\{\frac{d}{d\varepsilon} U_{t\leftarrow 0}^{\mathbf{X}(\omega)}(y_0 + \varepsilon a)\right\}_{\varepsilon=0}.$$

And let $M_{\mathbf{X}(\cdot)}^{(y_0,V)} : \mathcal{W} \to \mathbb{R}_+$ *denote the random variable*

$$M_{\mathbf{X}(\cdot)}^{(y_0,V)}(\omega) \equiv M_{\mathbf{X}(\omega)}^{(y_0,V)} := \big|J_{t\leftarrow 0}^{\mathbf{X}(\omega)}(y_0)\big|_{p\text{-var};[0,T]}.$$

Then for all y_0 *in* $\mathbb{R}^e$ *and all* $r < 2/\rho$ *we have*

$$\exp\big[\big(\log M_{\mathbf{X}(\cdot)}^{(y_0,V)}\big)^r\big] \qquad \text{is in } \bigcap_{q>0} L^q(\mu).$$

PROOF. Remark 5.1 guarantees the existence of a unique natural lift $\mathbf{X}$ for X. Furthermore, we know that if $V = (V^1, \dots, V^d)$ is any collection of Lip-γ vector fields (and $\gamma > p$), then the solution flow obtained by driving $\mathbf{X}$ along V is differentiable. Lemma 4.6 and Proposition 4.11 together yield (4.10) from which it follows that for any $\alpha > 0$ and y_0 in $\mathbb{R}^e$

$$M_{\mathbf{X}(\cdot)}^{(y_0,V)} \le c_1 \|\mathbf{X}(\omega)\|_{p\text{-var};[0,T]} \exp\big[c_1 N_{\alpha,I,p}\big(\mathbf{X}(\omega)\big)\big],$$

where $I = [0, T]$ and c_1 is a nonrandom constant which depends on α, p, γ and $|V|_{\text{Lip-}\gamma}$. Without loss of generality we take $c_1 > 1$. Then for three further (again nonrandom) constants c_2 and c_3 an easy calculation gives

$$\begin{aligned}\big(\log M_{\mathbf{X}(\cdot)}^{(y_0,V)}\big)^r &\le c_2 + c_3\big(\log\|\mathbf{X}(\omega)\|_{p\text{-var};[0,T]}\big)^r + c_4 N_{\alpha,I,p}\big(\mathbf{X}(\omega)\big)^r \\ &\le c_5 + c_3 \log\|\mathbf{X}(\omega)\|_{p\text{-var};[0,T]}^r + c_4 N_{\alpha,I,p}\big(\mathbf{X}(\omega)\big)^r.\end{aligned}$$

Hence, we have

$$\exp\big[\big(\log M_{\mathbf{X}(\cdot)}^{(y_0,V)}\big)^r\big] \le c_5 \|\mathbf{X}(\omega)\|_{p\text{-var};[0,T]}^{c_3 r} \exp\big[c_4 N_{\alpha,I,p}\big(\mathbf{X}(\omega)\big)^r\big]. \tag{6.8}$$

By Theorem 6.3 and the remark following it, the random variable

$$\exp\big[c_5 N_{\alpha,I,p}\big(\mathbf{X}(\omega)\big)^r\big]$$

on the right-hand side of (6.8) is $L^q(\mu)$ for all $q > 0$ provided $r < 2/\rho$. On the other hand $\|\mathbf{X}(\omega)\|_{p\text{-var};[0,T]}$ has a Gaussian tail (see [12]), and hence also has finite moments of all order. Using these two observations together with Cauchy–Schwarz inequality in (6.8) gives the desired conclusion. □

The above result applies (in particular) to fractional Brownian motion, $H > 1/3$. But in the case of fBm we can leverage the specific embedding properties to state an alternative version of the theorem which applies when $H > 1/4$.

THEOREM 6.6 (Fractional Brownian motion). *Let* $(X_t)_{t\in[0,T]} = (X_t^1, \dots, X_t^d)_{t\in[0,T]}$ *be fractional Brownian motion with Hurst parameter* $H > 1/4$. *Let* $(\mathcal{W}, \mathcal{H}, \mu)$ *denote the abstract associated with* X. *Let* $\gamma > p > 1/H$. *Then* X *lifts to a geometric* p*-rough path* $\mathbf{X}$, *and if* $V = (V^1, \dots, V^d)$ *is a collection of* Lip-γ *vector fields on* $\mathbb{R}^e$, *the RDE*

$$dY_t = V(Y)\, d\mathbf{X}, \qquad Y(0) = y_0$$

has a unique solution. The flow $U^{\mathbf{X}(\omega)}_{t\leftarrow 0}(\cdot)$ induced by the solution to this RDE is differentiable. Let this derivative be given by

$$J^{\mathbf{X}(\omega)}_{t\leftarrow 0}(y_0)\cdot a := \left\{\frac{d}{d\varepsilon}U^{\mathbf{X}(\omega)}_{t\leftarrow 0}(y_0+\varepsilon a)\right\}_{\varepsilon=0},$$

and let $M^{(y_0,V)}_{\mathbf{X}(\cdot)}:\mathcal{W}\to\mathbb{R}_+$ denote the random variable

$$M^{(y_0,V)}_{\mathbf{X}(\cdot)}(\omega)\equiv M^{(y_0,V)}_{\mathbf{X}(\omega)} := \left|J^{\mathbf{X}(\omega)}_{t\leftarrow 0}(y_0)\right|_{p\text{-}\mathrm{var};[0,T]}.$$

Then for any $r<2H+1$, we have that

$$\exp\left[\left(\log M^{(y_0,V)}_{\mathbf{X}(\cdot)}\right)^r\right] \textit{ is in } \bigcap_{q>0} L^q(\mu)$$

for all y_0 in $\mathbb{R}^e$.

PROOF. The argument is the similar to that of the last theorem; we have to verify the hypothesis of Theorem 6.3. Notice first that if $r_1<r_2$, then a simple calculation gives

$$\exp\left[\left(\log M^{(y_0,V)}_{\mathbf{X}(\cdot)}\right)^{r_1}\right]\le 1+\exp\left[\left(\log M^{(y_0,V)}_{\mathbf{X}(\cdot)}\right)^{r_2}\right].$$

It is therefore sufficient to prove the result for $1<r<2H+1$. Fix such any such r. The fact that X lifts to a geometric p-rough path for any $p>H^{-1}$ is by now a familiar one. Since for any $\tilde{p}>p$ we have

$$\left|J^{\mathbf{X}(\omega)}_{t\leftarrow 0}(y_0)\right|_{\tilde{p}\text{-}\mathrm{var};[0,T]}\le\left|J^{\mathbf{X}(\omega)}_{t\leftarrow 0}(y_0)\right|_{p\text{-}\mathrm{var};[0,T]},$$

it is sufficient to prove the result for any p satisfying

$$H^{-1}<p<\max\left[4,2(r-1)^{-1}\right].$$

Fix any such p in this interval and let q be given by

$$q=\left(\frac{1}{p}+\frac{1}{2}\right)^{-1}.$$

It follows that

$$r<\frac{2}{q}<1+2H.$$

The calculations of Corollary 5.5 then ensure that $\mathcal{H}\hookrightarrow C^{q\text{-}\mathrm{var}}([0,T],\mathbb{R}^d)$. This allows us to apply Theorem 6.3 to deduce that

$$\exp\left[N_{\alpha,I,p}\left(\mathbf{X}(\omega)\right)^r\right] \tag{6.9}$$

is μ-integrable. The result then follows by repeating the steps of the proof of the previous theorem. □

REMARK 6.7. In particular these results imply (under the stated conditions) that $|J^{\mathbf{X}(\omega)}_{t\leftarrow 0}(y_0)|_{p\text{-}\mathrm{var};[0,T]}$ has finite moments of all order.

Acknowledgements. We would like to thank the anonymous referee for a number of suggestions that helped to improve the presentation of this paper significantly.

REFERENCES

[1] BAUDOIN, F. and HAIRER, M. (2007). A version of Hörmander's theorem for the fractional Brownian motion. *Probab. Theory Related Fields* **139** 373–395. MR2322701

[2] CASS, T. and FRIZ, P. (2010). Densities for rough differential equations under Hörmander's condition. *Ann. of Math.* (2) **171** 2115–2141. MR2680405

[3] CASS, T., FRIZ, P. and VICTOIR, N. (2009). Non-degeneracy of Wiener functionals arising from rough differential equations. *Trans. Amer. Math. Soc.* **361** 3359–3371. MR2485431

[4] CASS, T. and LYONS, T. (2010). Evolving communities and individual preferences. Unpublished manuscript.

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

[6] FRIZ, P. and OBERHAUSER, H. (2009). Rough path limits of the Wong–Zakai type with a modified drift term. *J. Funct. Anal.* **256** 3236–3256. MR2504524

[7] FRIZ, P. and OBERHAUSER, H. (2010). A generalized Fernique theorem and applications. *Proc. Amer. Math. Soc.* **138** 3679–3688. MR2661566

[8] FRIZ, P. and RIEDEL, S. (2012). Integrability of linear rough differential equations. Available at arXiv:1104.0577v3.

[9] FRIZ, P. and VICTOIR, N. (2006). A variation embedding theorem and applications. *J. Funct. Anal.* **239** 631–637. MR2261341

[10] FRIZ, P. and VICTOIR, N. (2010). Differential equations driven by Gaussian signals. *Ann. Inst. Henri Poincaré Probab. Stat.* **46** 369–413. MR2667703

[11] FRIZ, P. and VICTOIR, N. (2011). A note on higher dimensional p-variation. *Electron. J. Probab.* **16** 1880–1899. MR2842090

[12] FRIZ, P. K. and VICTOIR, N. B. (2010). *Multidimensional Stochastic Processes as Rough Paths: Theory and Applications. Cambridge Studies in Advanced Mathematics* **120**. Cambridge Univ. Press, Cambridge. MR2604669

[13] GANESH, A., O'CONNELL, N. and WISCHIK, D. (2004). *Big Queues. Lecture Notes in Math.* **1838**. Springer, Berlin. MR2045489

[14] GUASONI, P. (2006). No arbitrage under transaction costs, with fractional Brownian motion and beyond. *Math. Finance* **16** 569–582. MR2239592

[15] GUBINELLI, M., LEJAY, A. and ANTIPOLIS, S. (2009). Global existence for rough differential equations under linear growth condition. Available at arXiv:0905.2399v1.

[16] HAIRER, M. and MATTINGLY, J. C. (2006). Ergodicity of the 2D Navier–Stokes equations with degenerate stochastic forcing. *Ann. of Math.* (2) **164** 993–1032. MR2259251

[17] HAIRER, M. and PILLAI, N. (2011). Regularity of laws and ergodicity of hypoelliptic SDEs driven by rough paths. Available at arXiv:1104.5218v1.

[18] HAIRER, M. and PILLAI, N. S. (2011). Ergodicity of hypoelliptic SDEs driven by fractional Brownian motion. *Ann. Inst. Henri Poincaré Probab. Stat.* **47** 601–628. MR2814425

[19] HAMBLY, B. M. and LYONS, T. J. (1998). Stochastic area for Brownian motion on the Sierpinski gasket. *Ann. Probab.* **26** 132–148. MR1617044

[20] HU, Y. and TINDEL, S. (2011). Smooth density for some nilpotent rough differential equations. Available at arXiv:1104.1972.

[21] INAHAMA, Y. (2012). A moment estimate of the derivative process in rough path theory. *Proc. Amer. Math. Soc.* **140** 2183–2191. MR2888204

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

[23] LEDOUX, M., QIAN, Z. and ZHANG, T. (2002). Large deviations and support theorem for diffusion processes via rough paths. *Stochastic Process*. *Appl*. **102** 265–283. MR1935127

[24] LYONS, T. and QIAN, Z. (2002). *System Control and Rough Paths*. Oxford Univ. Press, Oxford. MR2036784

[25] LYONS, T. J. (1998). Differential equations driven by rough signals. *Rev. Mat*. *Iberoam*. **14** 215–310. MR1654527

[26] LYONS, T. J., CARUANA, M. and LÉVY, T. (2007). *Differential Equations Driven by Rough Paths*. *Lecture Notes in Math*. **1908**. Springer, Berlin. MR2314753

[27] NORRIS, J. (1986). Simplified Malliavin calculus. In *Séminaire de Probabilités*, *XX*, 1984/85. *Lecture Notes in Math*. **1204** 101–130. Springer, Berlin. MR0942019

[28] NUALART, D. (2006). *The Malliavin Calculus and Related Topics*, 2nd ed. Springer, Berlin. MR2200233

T. CASS
C. LITTERER
DEPARTMENT OF MATHEMATICS
IMPERIAL COLLEGE LONDON
HUXLEY BUILDING
180 QUEENSGATE
LONDON SW7 2AZ
UNITED KINGDOM

T. LYONS
OXFORD-MAN INSTITUTE OF QUANTITATIVE FINANCE
UNIVERSITY OF OXFORD
WALTON WELL ROAD
OXFORD OX2 6ED
UNITED KINGDOM