Rough flows

By Ismaël Bailleul and Sebastian Riedel

Abstract. We introduce in this work a concept of rough driver that somehow provides a rough path-like analogue of an enriched object associated with time-dependent vector fields. We use the machinery of approximate flows to build the integration theory of rough drivers and prove well-posedness results for rough differential equations on flows and continuity of the solution flow as a function of the generating rough driver. We show that the theory of semimartingale stochastic flows developed in the 80's and early 90's fits nicely in this framework, and obtain as a consequence some strong approximation results for general semimartingale flows and provide a fresh look at large deviation theorems for 'Gaussian' stochastic flows.

1. Introduction.

An elementary construction recipe of flows was recently introduced in [**Bai15**] and used there to get back the core results of Lyons' theory of rough differential equations [**Lyo98**], [**FV10**] in a very short and elementary way. This work emphasizes the fact that it may be worth considering flows of maps as the primary objects from which the individual trajectories can be built, as opposed to the classical point of view that constructs a flow from an uncountable collection of individual trajectories. Probabilists know how tricky it can be to deal with uncountably many null sets. As well-recognized now, the main success of Lyons' theory was to disentangle probability from pure dynamics in the study of stochastic differential equations by showing that the dynamics is a deterministic and continuous function of an enriched signal that is constructed from the noise in the equation by purely probabilistic means. This very clean picture led to new proofs and extensions of foundational results in the theory of stochastic differential equations, such as Stroock and Varadhan support theorem or the basics of Freidlin and Ventzel theory of large deviations for diffusions.

It was realized in the late 70's that stochastic differential equations not only define individual trajectories, they also define flows of regular homeomorphisms, depending on the regularity of the vector fields involved in the dynamics [**Elw78**], [**Bis81**], [**IW81**], [**Kun81**]. This opened the door to the study of stochastic flows of maps for themselves [**Har81**], [**Bax80**], [**LJ82**], [**LJ85**] and it did not take long time before Le Jan and Watanabe [**LJW84**] clarified definitely the situation by showing that, in a semimartingale setting, there is a one-to-one correspondence between flows of diffeomorphisms and

time-varying stochastic velocity fields, under proper regularity conditions on the objects involved. We offer in the present work an embedding of the theory of semimartingale stochastic flows into the theory of rough flows similar to the embedding of the theory of stochastic differential equations into the theory of rough differential equations. While the acquainted reader will have noticed that the latter framework can be used to deal with Brownian flows by seeing them as solution flows to some Banach space-valued rough differential equations driven by a Brownian rough path in some space of vector fields, such as done in [**Der10**], [**DD12**], the situation is not so clear for more general random velocity fields and stochastic flows of maps. Our approach provides a simple setting for dealing with the general case, and we provide in this work an elementary direct approach to the construction of stochastic flows whose scope goes beyond the realm of semimartingale calculus.

Our theory of rough flows is based on the deterministic "Approximate flow-to-flow" machinery introduced in [**Bai15**], which gives body to the following fact. To a two-index family $(\mu_{ts})_{0\leq s\leq t\leq T}$ of maps which falls short from being a flow one can associate a unique flow $(\varphi_{ts})_{0\leq s\leq t\leq T}$ close to $(\mu_{ts})_{0\leq s\leq t\leq T}$; moreover the flow φ depends continuously on the approximate flow μ. We introduce a notion of rough driver, that is an enriched version of a time-dependent vector field, that is given by the additional datum of a time-dependent second order differential operator satisfying some algebraic and analytic conditions. A notion of solution to a differential equation driven by a rough driver will be given, in the line of what was done in [**Bai15**] for rough differential equations, and the approximate flow-to-flow machinery will be seen to lead to a clean and simple well-posedness result for such equations. Importantly, the Itô map, that associates to a rough driver the solution flow to its associated equation, is continuous. This continuity result is the key to deep results in the theory of stochastic flows. We shall indeed prove that reasonable semimartingale velocity fields can be lifted to rough drivers under some mild boundedness and regularity conditions, and that the solution flow associated to the semimartingale rough driver coincides almost surely with the solution flow to the Kunita-type Stratonovich differential equation driven by the velocity field. In this sense, our theory of rough flows encompasses the theory of stochastic flows. A Wong–Zakaï theorem will be proved for a general class of semimartingale velocity fields and a sharp large deviation principle for Brownian flows will be proved as a consequence of the continuity of the Itô map. No Wong–Zakaï result was available so far for semimartingale stochastic flows.

The setting of rough drivers and rough flows is presented in Section 2, together with the approximate flow-to-flow machinery. This is the core of the deterministic machinery and everything that follows elaborates on this material, in a probabilistic setting. Some additional material on random rough drivers is in particular given in Section 2.5, where we provide some new variations around the Kolmogorov regularity theorem needed along the way and some sufficient conditions for a process to be bounded; this material may be of independent interest. We show in Section 3 that reasonable semimartingale velocity fields can be lifted to rough drivers under appropriate mild boundedness and regularity assumptions and prove that the theory of semimartingale stochastic flows of maps is naturally embedded in the theory of rough flows. As an illustration of the continuity of the Itô map, we prove in Section 3.5 a Wong–Zakaï theorem for stochastic flows of maps,

and provide in Section 4 a fresh look at large deviation theorems for Brownian stochastic flows. We proved in the follow up work [**BRS17**] that a large class of Gaussian vector fields can be lifted into rough drivers; this shows explicitly that the setting of rough drivers and rough flows goes beyond the semimartingale horizon. The point of view of rough flows presented here was also used in [**BC17**] to investigate the problem of stochastic turbulence.

The size of this work is related to the fact that it is intended to be as self-contained as possible; no a priori knowledge of the theory of stochastic flows is required for its understanding in particular. We have included as a consequence some material that is well-known to experts in stochastic flows or large deviation theory for instance. A reader interested only in the machinery of rough flows will have a complete picture by reading Sections 2.1 to 2.4 and Theorem 2.16 in Section 2.5.

NOTATIONS. We gather here for reference a number of notations that will be used throughout the text.

- We shall exclusively use the letter E to denote a Banach space. If F is another Banach space, we shall denote by $\mathrm{L}(E, F)$ the set of continuous linear maps from E to F, and for $M \in \mathrm{L}(E, F)$, we shall write $|M|$ for its operator norm. In this setting, differentiability and regularity notions are understood in the sense of Fréchet. We will also use the notation $\mathrm{L}(E) := \mathrm{L}(E, E)$.

- If E and F are two Banach spaces and $f \colon E \to F$, we will denote by $Df \colon E \to \mathrm{L}(E, F)$ its Fréchet derivative (recall that the Fréchet derivative coincides with the total derivative in the finite-dimensional case). For $x \in E$, we will also use the notation $D_x f := Df(x) \in \mathrm{L}(E, F)$. The second derivative will be denoted by $D^2 f \colon E \to \mathrm{L}(E, \mathrm{L}(E, F))$ and, more generally, the n-th derivative by

$$D^n f \colon E \to \mathrm{L}(E, \cdots, \mathrm{L}(E, F) \cdots).$$

Recall that there is a canonical embedding

$$\mathrm{L}(E, \cdots, \mathrm{L}(E, F) \cdots) \hookrightarrow \mathrm{L}(E^{\otimes_a n}, F)$$

as vector spaces, where $E^{\otimes_a n}$ denotes the n-th algebraic tensor product of E. Assume that we have a norm defined on $E^{\otimes_a n}$, and let $E^{\otimes n}$ denote the closure with respect to this norm. If the embedding

$$\mathrm{L}(E, \cdots, \mathrm{L}(E, F) \cdots) \hookrightarrow \mathrm{L}(E^{\otimes n}, F)$$

is continuous, we follow [**FH14**, p. 11] and call this norm *compatible*. Note that there are always compatible norms, cf. [**Rya02**], but different choices may lead to different spaces if the underlying space E is infinite-dimensional. Choosing a compatible norm, the n-th Fréchet derivative can be seen as a continuous function $D^n f \colon E \to \mathrm{L}(E^{\otimes n}, F)$. Again, we set $D_x^n f := D^n f(x) \in \mathrm{L}(E^{\otimes n}, F)$. We recall the reader that a vector field V on E induces a first order differential operator, also called a derivation,

$$f \mapsto (Df)(V).$$

- In the finite dimensional case $E = \mathbb{R}^d$, we will use the letter D for the total derivative and the usual notation for partial derivatives:

$$\partial_i = \partial_{x_i} = \frac{\partial}{\partial x_i}, \quad \partial_i^n = \frac{\partial^n}{\partial x_i^n}, \quad \partial^\alpha = \partial_1^{\alpha_1} \cdots \partial_d^{\alpha_d}, \quad \partial_{i_1,\dots,i_n}^n = \partial_{i_1} \cdots \partial_{i_n}$$

  where $i, i_1, \dots, i_n \in \{1, \dots, d\}$, $n \in \mathbb{N}_0$ and $\alpha = (\alpha_1, \dots, \alpha_d) \in \mathbb{N}_0^d$. If a function f is of the form $f(x,y)$ with $x \in \mathbb{R}^{d_1}$ and $y \in \mathbb{R}^{d_2}$, we use the notation ∂_x^α resp. ∂_y^β to denote partial derivatives in the x resp. y-variable, $\alpha \in \mathbb{N}_0^{d_1}$, $\beta \in \mathbb{N}_0^{d_2}$. In the same situation, ∂_x resp. ∂_y means the total derivative in the x resp. y-variable.

- For functions $x \colon [0, T] \to E$, we will use the notation $x_{ts} = x_t - x_s$ for the increments of x.

- As is common, we shall use Einstein's summation convention that $a^i b_i := \sum_i a^i b_i$.

- A flow on E is a family $(\varphi_{ts})_{0 \le s \le t \le T}$ of maps from E to itself such that $\varphi_{tt} = \mathrm{Id}$, for all $0 \le t \le T$, and $\varphi_{tu} \circ \varphi_{us} = \varphi_{ts}$, for all $0 \le s \le u \le t \le T$. The letter T, here and below, will always stand for a finite time horizon.

2. Rough flows.

2.1. Flows and approximate flows.

We introduced in [**Bai15**] a simple machinery for constructing flows on E from approximate flows that can be understood as a generalization of Lyons' workhorse [**Lyo98**] for constructing a rough path from an almost rough path; this is the core tool for the construction of the rough integral. Roughly speaking, the "Approximate flow-to-flow" machinery says that if we are given a family of maps $(\mu_{ts})_{0 \le s \le t \le T}$ from E to itself, and if the maps μ are close to defining a flow, in the sense that $\mu_{tu} \circ \mu_{us} - \mu_{ts}$ is small in a quantitative way, for $s \le u \le t$ with $t - s$ small, then there exists a unique flow close to μ. In the rough paths setting, Lyons almost multiplicative functionals involve a family $a = (a_{ts})_{0 \le s \le t \le T}$ of elements of a tensor algebra such that $a_{tu} a_{us} - a_{ts}$ has some given size whenever $s \le u \le t$ with $t - s$ small, with the product on the tensor space used here. Despite their similarity, Lyons' setting differs from the present setting in that multiplication in a tensor algebra satisfies the distributivity property $ab - ac = a(b - c)$, which obviously does not hold if a, b, c are maps and the product stands for composition. This seemingly minor point makes a real difference though, so it is fortunate that one can still get an analogue of Lyons' theorem in a function space setting. This comes at a little price on the regularity of the set of maps μ that one can consider. As usual, for $0 < r \le 1$, we denote by $\mathcal{C}^r$ the space of r-Hölder functions, with the understanding that they are Lipschitz continuous for $r = 1$. Following Stein and Lyons, for $r > 1$, a function of class $C^{[r]}$, with $[r]$-th derivative that is $(r - [r])$-Hölder, is said to be r-Lipschitz.

DEFINITION 2.1.   Let $0 < r \le 1$ be given. A $\mathcal{C}^r$-*approximate flow* on E is a family $\left(\mu_{ts}\right)_{0 \le s \le t \le T}$ of $(1 + \rho)$-Lipschitz maps from E to itself, for some $0 < \rho < 1$, depending

continuously on (s, t) in the topology of uniform convergence and enjoying the following two properties.

- **Perturbation of the identity.**

  There exists a constant α with

  $$0 < 1 - \rho < \alpha < 1,$$

  such that one has for any $0 \leq s \leq t \leq T$, and any $x \in E$, the decomposition

  $$D_x \mu_{ts} = \mathrm{Id} + A_x^{ts} + B_x^{ts}, \tag{2.1}$$

  for some $\mathrm{L}(E)$-valued ρ-Lipschitz maps A^{ts} with ρ-Lipschitz norm bounded above by $c|t - s|^\alpha$, and some $\mathrm{L}(E)$-valued $\mathcal{C}^1$-bounded maps B^{ts}, with $\mathcal{C}^1$-norm bounded above by $o_{t-s}(1)$.

- **$\mathcal{C}^r$-approximate flow property.**

  There exists some positive constants c_1 and $a > 1$, such that one has

  $$\left\| \mu_{tu} \circ \mu_{us} - \mu_{ts} \right\|_{\mathcal{C}^r} \leq c_1 \, |t - s|^a \tag{2.2}$$

  for all $0 \leq s \leq u \leq t \leq T$.

Note that one requires a quantitative bound on A while we only require a qualitative information on B, at the price of some more regularity for the latter. This fine decomposition of the differential of μ_{ts}, as opposed to assuming only $D\mu_{ts} = \mathrm{Id} + A^{ts}$, makes the setting more flexible. The exponents $(1 + \rho)$ and α may depend on μ.

The introduction of the notion of approximate flow is justified by the following result proved in [**Bai15**]. Given a partition $\pi_{ts} = \{s = s_0 < s_1 < \cdots < s_{n-1} < s_n = t\}$ of $(s, t) \subset [0, T]$, set

$$\mu_{\pi_{ts}} = \mu_{s_n s_{n-1}} \circ \cdots \circ \mu_{s_1 s_0}.$$

THEOREM 2.2 (Constructing flows on E). *A $\mathcal{C}^r$-approximate flow, with $1/a < r$, defines a unique flow $(\varphi_{ts})_{0 \leq s \leq t \leq T}$ on E to which one can associate a positive constant δ such that the inequality*

$$\left\| \varphi_{ts} - \mu_{ts} \right\|_\infty \leq |t - s|^a$$

holds for all $0 \leq s \leq t \leq T$ with $t - s \leq \delta$; this flow satisfies the inequality

$$\left\| \varphi_{ts} - \mu_{\pi_{ts}} \right\|_\infty \leq 2 \, c_1 \, T \, \left| \pi_{ts} \right|^{ar-1} \tag{2.3}$$

for any partition π_{ts} of any interval $(s, t) \subset [0, T]$, with mesh $\left| \pi_{ts} \right| \leq \delta$. Moreover, the $\mathcal{C}^r$ norm of the maps φ_{ts} is uniformly bounded by a function of the constant c_1 that appears in (2.2), for all $0 \leq s \leq t \leq T$.

This statement generalises Gubinelli's sewing lemma [**Gub04**], such as reshaped by Feyel and de la Pradelle in [**FdLP06**], to the non-commutative, non-associative setting of

maps on E. (The non-commutative sewing lemma of Feyel, de la Pradelle and Mokobodski [**FdLPM08**] requires associativity and cannot be used in the present setting.)

Theorem 2.2 is stated in [**Bai15**] for $\mathcal{C}^1$-approximate flows; the proof given there works verbatim for $\mathcal{C}^r$-approximate flows provided $1/a < r$; a $\mathcal{C}^1$-map is then understood in that setting as a Lipschitz map. The crucial point in the above statement is the fact that if μ depends continuously in $\mathcal{C}^r$ on some parameter then φ also happens to depend continuously on that parameter, in $\mathcal{C}^0$, as a direct consequence of estimate (2.3). As made clear in [**Bai15**], Theorem 2.2 can be seen as the cornerstone of the theory of rough differential equations, with the continuity of the Itô-Lyons solution map given as a consequence of the aforementioned continuity of φ on a parameter. We shall see in the present work that Theorem 2.2 is all we need to get back and extend the core results of the theory of stochastic flows intensively developed in the 80's and early 90's. We shall need for that purpose to introduce a notion of enriched velocity field that will somehow play the role in our setting of weak geometric Hölder p-rough paths, with $2 \leq p < 3$, in rough paths theory.

We shall thus pick a regularity exponent $2 \leq p < 3$ here, once and for all.

2.2. Rough drivers.

Let $(V(\cdot, t))_{0 \leq t \leq T}$ be a time-dependent vector field on E, with time increments

$$V_{ts}(\cdot) := V(\cdot, t) - V(\cdot, s).$$

To get a hand on the definition of a weak geometric p-rough driver given below, think of V_{ts} as given by the formula

$$V_{ts} = V X_{ts}, \tag{2.4}$$

where $V(x) \in \mathrm{L}(\mathbb{R}^\ell, \mathbb{R}^d)$, for all $x \in \mathbb{R}^d$, and $\boldsymbol{X} = (X, \mathbb{X})$ is a p-rough path over $\mathbb{R}^\ell$. Write V_i for the image by V of the ith vector in the canonical basis of $\mathbb{R}^\ell$. A solution path $x_{\bullet}$ to the rough differential equation

$$dx_t = V(x_t)\,\boldsymbol{X}(dt)$$

can be characterized as a path satisfying some uniform Euler–Taylor expansion of the form

$$f(x_t) = f(x_s) + X_{ts}^i \big(V_i f\big)(x_s) + \mathbb{X}_{ts}^{jk}\big(V_j V_k f\big)(x_s) + O\Big(|t - s|^{3/p}\Big)$$

for all sufficiently regular real-valued functions f on $\mathbb{R}^d$. Recall we identify vector fields and first order differential operators, so

$$V_j V_k f = (D^2 f)(V_j, V_k) + (Df)((DV_k)(V_j)).$$

The present Section will make it clear that the operators $X_{ts}^i V_i = V X_{ts}$ and $\mathbb{X}_{ts}^{jk} V_j V_k = (DV)V\,\mathbb{X}_{ts}$ are all we need in this formula to run the theory, with no need to separate their space part, given by V and $(DV)V$, from their time part $\boldsymbol{X}_{ts}$.

DEFINITION 2.3. Let $2 \leq p < 3$, and $p - 2 < \rho \leq 1$ be given. A weak geometric (p, ρ)-*rough driver* is a family $(\boldsymbol{V}_{ts})_{0 \leq s \leq t \leq T}$, with

$$\boldsymbol{V}_{ts} := \big(V_{ts}, \mathbb{V}_{ts}\big),$$

and $\mathbb{V}_{ts}$ a second order differential operator, such that

(i) the vector fields V_{ts} are additive as functions of time

$$V_{ts} = V_{tu} + V_{us}$$

for all $s \leq u \leq t$, and each V_{ts} is of class $\mathcal{C}^{2+\rho}$ on E, with

$$\sup_{0 \leq s < t \leq T} \frac{\|V_{ts}\|_{\mathcal{C}^{2+\rho}}}{|t - s|^{1/p}} < \infty,$$

(ii) the second order differential operators

$$W_{ts} := \mathbb{V}_{ts} - \frac{1}{2} V_{ts} V_{ts},$$

coincide with the derivations (first order differential operator) given by some vector fields on E, denoted by the same letters W_{ts}, and we have

$$\sup_{0 \leq s < t \leq T} \frac{\|W_{ts}\|_{\mathcal{C}^{1+\rho}}}{|t - s|^{2/p}} < \infty,$$

(iii) we have

$$\mathbb{V}_{ts} = \mathbb{V}_{tu} + V_{us} V_{tu} + \mathbb{V}_{us},$$

for any $0 \leq s \leq u \leq t \leq T$.

(The formulation of item (ii) takes care of the fact that in a general Banach space, not all derivations are associated with a vector field. See e.g. example 8.3 in Section 8.4 of [**BGN04**]. In finite dimension, all derivations are associated with a vector field.) With in mind the model weak geometric p-rough driver given by formula (2.4), the requirement $p - 2 < \rho$ appears as a natural assumption to impose, given known well-posedness results on rough differential equations [**Dav07**]; the first order condition on the operators W_{ts} justifies that we call $\boldsymbol{V}$ a *weak geometric p-rough driver*, and condition (iii) stands for an analogue of *Chen's relation*. We shall freely talk about rough drivers rather than weak geometric (p, ρ)-rough drivers in the sequel.

DEFINITION 2.4. We define the (pseudo-)norm of $\boldsymbol{V}$ to be

$$\|\boldsymbol{V}\|_{p,\rho} := \sup_{0 \leq s < t \leq T} \left\{ \frac{\|V_{ts}\|_{\mathcal{C}^{2+\rho}}}{|t - s|^{1/p}} \vee \frac{\|W_{ts}\|_{\mathcal{C}^{1+\rho}}}{|t - s|^{2/p}} \right\}. \tag{2.5}$$

and define an associated (pseudo-)metric on the set $\mathcal{D}_{p,\rho}$ of weak geometric (p,ρ)-rough drivers setting

$$d_{p,\rho}(\boldsymbol{V}, \boldsymbol{V}') := \left\| \boldsymbol{V} - \boldsymbol{V}' \right\|_{p,\rho}.$$

Like the space of rough paths the space of rough drivers is not a linear space. We will also need the homogeneous metric

$$\mathfrak{d}_{p,\rho}(\boldsymbol{V}, \boldsymbol{V}') := \sup_{0 \leq s < t \leq T} \left\{ \frac{\left\| V_{ts} - V'_{ts} \right\|_{\mathcal{C}^{2+\rho}}}{|t-s|^{1/p}} \vee \sqrt{\frac{\left\| W_{ts} - W'_{ts} \right\|_{\mathcal{C}^{1+\rho}}}{|t-s|^{2/p}}} \right\}. \tag{2.6}$$

We will often drop the subindices p, ρ when it is clear from the context in which space we are working in. Note that d and $\mathfrak{d}$ induce the same topology in the space of rough drivers.

Note that one should add the $\mathcal{C}^{2+\rho}$-norm of $V(\cdot, 0)$ in formula (2.5) to define a proper norm on the space of rough drivers. This has no consequences as rough drivers are only used via their increments. Note also that given a rough driver $\boldsymbol{V}$ and $0 < a \leq T$, one defines another rough driver $\boldsymbol{V}^a := (V^a, \mathbb{V}^a)$, on the time interval $[0, a]$, setting

$$\begin{aligned} V^a_{ts} &= V_{a-s,a-t}, \\ \mathbb{V}^a_{ts} &= -\mathbb{V}_{a-s,a-t} + V_{a-s,a-t} V_{a-s,a-t}, \end{aligned} \tag{2.7}$$

for all $0 \leq s \leq t \leq a$. It is indeed elementary to check that these operators satisfy the algebraic conditions (ii) and (iii), with

$$W^a_{ts} := \mathbb{V}^a_{ts} - \frac{1}{2} V^a_{ts} V^a_{ts} = -W_{a-s,a-t}; \tag{2.8}$$

that they satisfy the above analytic requirements is obvious. This rough driver is called the *time reversal of the rough driver $\boldsymbol{V}$, from time a*. Note that $\left\| \boldsymbol{V}^a \right\| \leq \left\| \boldsymbol{V} \right\|$.

2.3. Rough flows.

We shall adopt below a definition of a solution flow to the equation

$$d\varphi = \boldsymbol{V}(\varphi \,; dt) \tag{2.9}$$

similar to the above definition of a solution path to a rough differential equation. A solution flow will be required to satisfy some uniform Euler–Taylor expansion of the form

$$f \circ \varphi_{ts} - \left\{ f + V_{ts} f + \mathbb{V}_{ts} f \right\} = O\left(|t-s|^{3/p} \right),$$

for all sufficiently regular real-valued functions f on $\mathbb{R}^d$. It is actually elementary to construct a family of maps $(\mu_{ts})_{0 \leq s \leq t \leq T}$ which enjoys the above Euler–Taylor expansion property. The key point is that this family will turn out to be a $\mathcal{C}^\rho$-approximate flow, if ρ is not too small, so we shall get the existence and uniqueness of a solution flow from its very definition and Theorem 2.2.

Given $0 \leq s \leq t \leq T$, consider the ordinary differential equation

$$\dot{y}_u = V_{ts}(y_u) + W_{ts}(y_u), \quad 0 \leq u \leq 1, \tag{2.10}$$

and denote by μ_{ts} its well-defined time 1 map, associating to any $x \in E$ the value at time 1 of the solution of equation (2.10) started from x. Elementary results on ordinary differential equations imply that if one considers y_u as a function of x, for $0 \leq u \leq 1$, then we have

$$\left\| y_u - \mathrm{Id} \right\|_{\mathcal{C}^1} \leq c \left\| V \right\| \left| t - s \right|^{1/p} \tag{2.11}$$

for some universal positive constant c. Proposition 2.5 below shows that the maps μ_{ts} have the awaited Euler–Taylor expansion expected from a solution flow to equation (2.9).

PROPOSITION 2.5.   *We have*

$$\left\| f \circ \mu_{ts} - \left\{ f + V_{ts}f + \mathbb{V}_{ts}f \right\} \right\|_\infty \leq c \left\| f \right\|_{\mathcal{C}^{2+\rho}} |t - s|^{3/p}, \tag{2.12}$$

for any $f \in \mathcal{C}^{2+\rho}$ and any $0 \leq s \leq t \leq T$.

The proof of this statement is straightforward and relies on the the following formula. For all $x \in E$ and all $f \in \mathcal{C}^2$, we have

$$f\left(\mu_{ts}(x)\right) = f(x) + \int_0^1 \left(V_{ts}f\right)(y_u)\, du + \int_0^1 \left(W_{ts}f\right)(y_u)\, du$$

$$= f(x) + \left(V_{ts}f\right)(x) + \left(\mathbb{V}_{ts}f\right)(x) + \epsilon_{ts}^f(x)$$

where

$$\epsilon_{ts}^f(x) := \int_0^1 \int_0^u \left\{ \left(V_{ts}V_{ts}f\right)(y_r) - \left(V_{ts}V_{ts}f\right)(x) \right\} dr\,du + \int_0^1 \int_0^u \left(W_{ts}V_{ts}f\right)(y_r)\, dr\,du$$

$$+ \int_0^1 \left\{ \left(W_{ts}f\right)(y_u) - \left(W_{ts}f\right)(x) \right\} du.$$

The inequality

$$\left\| \epsilon_{ts}^f \right\|_{\mathcal{C}^\rho} \leq c \left(1 + \|V\|^3\right) \|f\|_{\mathcal{C}^{2+\rho}} |t - s|^{3/p}, \tag{2.13}$$

justifies Proposition 2.5.

THEOREM 2.6.   *The family of maps $(\mu_{ts})_{0 \leq s \leq t \leq T}$ is a $\mathcal{C}^\rho$-approximate flow which depends continuously on $((s,t), V)$ in $\mathcal{C}^0$-topology.*

PROOF.   The family $(\mu_{ts})_{0 \leq s \leq t \leq T}$ satisfies the regularity assumptions (2.1) as a direct consequence of classical results on the dependence of solutions to ordinary differential equations with respect to parameters, including the initial condition for the equation. These results also imply the continuous dependence of μ_{ts} on $((s,t), V)$ in $\mathcal{C}^\rho$-topology. To show that the family μ defines a $\mathcal{C}^\rho$-approximate flow, write, for $0 \leq s \leq u \leq t \leq T$,

$$\mu_{tu}\big(\mu_{us}(x)\big) = \mu_{us}(x) + V_{tu}\big(\mu_{us}(x)\big) + \big(\mathbb{V}_{tu}\mathrm{Id}\big)\big(\mu_{us}(x)\big) + \epsilon_{tu}^{\mathrm{Id}}\big(\mu_{us}(x)\big)$$

$$= x + V_{us}(x) + \big(\mathbb{V}_{us}\mathrm{Id}\big)(x) + \epsilon_{us}^{\mathrm{Id}}(x)$$

$$+ V_{tu}(x) + \big(V_{us}V_{tu}\big)(x) + \big(\mathbb{V}_{us}V_{tu}\big)(x) + \epsilon_{us}^{V_{tu}}(x)$$

$$+ \big(\mathbb{V}_{tu}\mathrm{Id}\big)(x) + \Big(\big(\mathbb{V}_{tu}\mathrm{Id}\big)\big(\mu_{us}(x)\big) - \big(\mathbb{V}_{tu}\mathrm{Id}\big)(x)\Big) + \epsilon_{tu}^{\mathrm{Id}}\big(\mu_{us}(x)\big)$$

$$= \mu_{ts}(x) + \Big\{\big(\mathbb{V}_{us}V_{tu}\big)(x) + \Big(\big(\mathbb{V}_{tu}\mathrm{Id}\big)\big(\mu_{us}(x)\big) - \big(\mathbb{V}_{tu}\mathrm{Id}\big)(x)\Big) + \epsilon_{us}^{V_{tu}}(x)$$

$$+ \epsilon_{us}^{\mathrm{Id}}(x) + \epsilon_{tu}^{\mathrm{Id}}\big(\mu_{us}(x)\big)\Big\}.$$

The approximate flow property then follows from the regularity assumptions on V_{ts} and $\mathbb{V}_{ts}$, and estimate (2.13). $\qquad\square$

With the notations used in the definition of an approximate flow, the exponent a that appears here in the approximate flow identity (2.2) is $a = 3/p$.

DEFINITION 2.7.   A *flow* $(\varphi_{ts})_{0 \le s \le t \le T}$ is said to *solve the rough differential equation*

$$d\varphi = \boldsymbol{V}(\varphi\,;dt) \tag{2.14}$$

if there exists a possibly $\boldsymbol{V}$-dependent positive constant δ such that the inequality

$$\big\|\varphi_{ts} - \mu_{ts}\big\|_\infty \le |t - s|^{3/p} \tag{2.15}$$

holds for all $0 \le s \le t \le T$ with $t - s \le \delta$. Flows solving a differential equation of the form (2.14) are called *rough flows*. If equation (2.14) is well-posed, the map which associates to a rough driver $\boldsymbol{V}$ the solution flow to equation (2.14) is called the *Itô map*.

Following Cass and Weidner [**CW17**], [**Bai18**], one can equivalently take the Taylor expansion property

$$\varphi_{ts} = \mathrm{Id} + V_{ts}\mathrm{Id} + \mathbb{V}_{ts}\mathrm{Id} + O\big(|t - s|^{3/p}\big)$$

as a defining property of a solution flow to the rough differential equation (2.14). The following well-posedness result comes as a direct consequence of Theorem 2.2 and Theorem 2.6. A family of maps is said to be uniformly $\mathcal{C}^\rho$ is it has uniformly bounded $\mathcal{C}^\rho$-norm.

THEOREM 2.8.   *Assume $\rho > p/3$. Then the differential equation on flows*

$$d\varphi = \boldsymbol{V}(\varphi\,;dt)$$

has a unique solution flow; it takes values in the space of uniformly $\mathcal{C}^\rho$-homeomorphisms of E, with uniformly $\mathcal{C}^\rho$-inverses, and depends continuously on $\boldsymbol{V}$ in the topology of uniform convergence.

PROOF.   It follows from the proof of Theorem 2.6 that one can choose as a constant c_1 in inequality (2.2) a multiple of $1 + \|\boldsymbol{V}\|^4$, so we have from Theorem 2.2 the estimate

$$\left\| \varphi_{ts} - \mu_{\pi_{ts}} \right\|_\infty \le c \left(1 + \|\boldsymbol{V}\|^4 \right) T \left| \pi_{ts} \right|^{\rho(3/p)-1},$$

for any partition π_{ts} of $(s,t) \subset [0,T]$ with mesh $\left| \pi_{ts} \right|$ small enough, say no greater than δ. Note that the exponent $\rho(3/p) - 1$ is positive. As these bounds are uniform in (s,t), and for $\boldsymbol{V}$ in a bounded set of the space of rough drivers, and each $\mu_{\pi_{ts}}$ is a continuous function of $\boldsymbol{V}$, the flow φ depends continuously on $((s,t), \boldsymbol{V})$.

To prove that φ is a homeomorphism, note that it follows from (2.8) that, for $0 \le a \le b \le t$, each μ_{ba} is a diffeomorphism with inverse given by the time one map $\mu^t_{t-a,t-b}$ of the ordinary differential equation

$$\dot{y}_u = -V_{ba}(y_u) - W_{ba}(y_u) = V^t_{t-a,t-b}(y_u) + W^t_{t-a,t-b}(y_u), \quad 0 \le u \le 1,$$

associated with the time reversed rough driver $\boldsymbol{V}^t$. As μ^t has the same properties as μ, the maps

$$\left(\mu_{\pi_{ts}} \right)^{-1} = \mu^{-1}_{s_1 s_0} \circ \cdots \circ \mu^{-1}_{s_n s_{n-1}} = \mu^t_{s_n s_{n-1}} \circ \cdots \circ \mu^t_{s_1 s_0}$$

converge uniformly to some continuous map φ^{-1}_{ts}, as the mesh of the partition π_{ts} tends to 0; these limit maps φ^{-1}_{ts} satisfy by construction $\varphi_{ts} \circ \varphi^{-1}_{ts} = \mathrm{Id}$.

As Theorem 2.2 provides a uniform control of the $\mathcal{C}^\rho$-norm of the maps φ_{ts}, the same control holds for their inverses since $\|\boldsymbol{V}^t\| \le \|\boldsymbol{V}\|$. We propagate this control from the set $\{(s,t) \in [0,T]^2 \,;\, s \le t, \ t - s \le \delta\}$ to the whole of $\{(s,t) \in [0,T]^2 \,;\, s \le t\}$ using the flow property of φ. $\qquad\square$

Note that the solution flow to the rough differential equation

$$d\psi = \boldsymbol{V}^T(\psi \,; dt),$$

driven by the time reversal of the rough driver $\boldsymbol{V}$, from time T, provides the inverse flow of φ, in the sense that

$$\varphi^{-1}_{ts} = \psi_{T-s,\, T-t},$$

for all $0 \le s \le t \le T$. Last, note that it is elementary to adapt the above results to add a globally Lipschitz drift in the dynamics; the above results hold in that setting as well.

REMARK 2.9. 1. Note that like for the theory of rough paths, the technical shape of the theory of rough drivers depends on that regularity exponent. Only two objects are needed in the definition of a rough driver when $2 \le p < 3$; for $3 \le p < 4$, we would need to introduce an additional object in the definition of a rough driver, that would thus consist of three objects; and so on. There is no real difficulty other than notational in giving a general theory, but as our applications of rough flows to the study of stochastic flows only require to develop the theory in the case where $2 \le p < 3$, we stick to that setting and invite the reader to make up herself her mind about what the general theory looks like.

2. The rough drivers introduced here are somewhat a dual version of similar objects that were introduced very recently in [**BG17**] by one of the authors and Gubinelli in the

study of the well-posedness of a general family of linear hyperbolic symmetric systems of equations driven by time-dependent vector fields that are only distributions in the time direction. The latter work deals with evolutions in function spaces and uses functional analytic tools in the setting of controlled paths to make a first step towards a general theory of rough PDEs, in the lines of the classical PDE approach based on duality, a priori estimates and compactness results. The present work does not overlap with the latter.

How the story goes on.   The entire technical core of the theory of rough flows is contained in Section 2.1 and Section 2.3. The remainder of this work is dedicated to

- showing how one can lift semimartingale velocity fields into rough drivers—Section 2.5 and 3.3,

- showing that stochastic and rough flows coincide for semimartingale velocity fields—Section 3.4,

- proving a Wong–Zakaï-type approximation theorem for semimartingale stochastic flows—Section 3.5,

- proving Schilder- and Freidlin–Ventzel-type large deviation theorems for flows generated by Gaussian rough drivers—Section 4.1.

We emphasize here that we proved in the subsequent work [**BRS17**] that one can lift into rough drivers a whole class of Gaussian velocity fields, showing that the setting of rough flows goes beyond the setting of semimartingale calculus.

2.4.   An Itô formula for rough flows.

With a view to identifying stochastic and rough flows in Section 3.4, we prove here an elementary Itô formula analogue to Friz and Hairer's Itô formula in [**FH14**]. As a matter of fact, Theorem 2.11 below states that any $1/p$-Hölder path in a Banach space satisfies an Itô formula, outside the setting of rough or controlled paths. To state and prove it recall Feyel and de la Pradelle sewing lemma [**FdLP06**], that can be seen as a precursor of the construction theorem for flows, Theorem 2.2. Given a partition $\pi_{ts} = \{s = s_0 < s_1 < \cdots < s_{n-1} < s_n = t\}$ of an interval $[s,t]$, and an E-valued 2-index map $z = (z_{ts})_{0 \le s \le t \le T}$, set

$$z_{\pi_{ts}} := z_{s_n s_{n-1}} + \cdots + z_{s_1 s_0}.$$

Theorem 2.10 ([**FdLP06**]).   *Let $(z_{ts})_{0 \le s \le t \le T}$ be an E-valued 2-index continuous map to which one can associate some positive constants c_1 and $a > 1$ such that*

$$\left| \left(z_{tu} + z_{us} \right) - z_{ts} \right| \le c_1 \left| t - s \right|^a \tag{2.16}$$

holds for all $0 \le s \le u \le t \le T$. Then there exists a unique continuous function $Z : [0,T] \to \mathbb{R}$, with increments $Z_{ts} := Z_t - Z_s$, to which one can associate a positive constant δ such that the inequality

$$\left| Z_{ts} - z_{ts} \right| \le \left| t - s \right|^a,$$

holds for all $0 \le s \le u \le t \le T$, with $t - s \le \delta$; this map Z satisfies the inequality

$$\left| Z_{ts} - z_{\pi_{ts}} \right| \le 2\, c_1\, T \left| \pi_{ts} \right|^{a-1}$$

for any partition π_{ts} of any interval $[s, t] \subset [0, T]$, with mesh $\left| \pi_{ts} \right| \le \delta$. It follows in particular that Z depends continuously on any parameter in uniform topology if z does.

A map z satisfying condition (2.16) is said to be *almost-additive*, and we write

$$Z_{ts} =: \int_s^t z_{du}.$$

We equip the tensor product space $E \otimes E$ with a compatible tensor norm that makes the natural embedding $\mathrm{L}(E, \mathrm{L}(E, \mathbb{R})) \subset \mathrm{L}(E \otimes E, \mathbb{R})$ continuous. Given such a choice, one can identify the second differential of a $\mathcal{C}^2$-real-valued function on E to an element of $\mathrm{L}(E \otimes E, \mathbb{R})$ that is symmetric; this is what we do below.

THEOREM 2.11 (Itô formula). *Let $F : [0, T] \times E \to \mathbb{R}$ be a $\mathcal{C}^1$-function of time with time derivative $\partial_t F(t, x)$ bounded and continuous, uniformly in $x \in E$. Assume also that F is of class $\mathcal{C}^3$ in the sense of Fréchet as a function of its E-component, with derivatives $F^{(1)}, F^{(2)}, F^{(3)}$ and $\partial_t F^{(1)}$, bounded uniformly in time[1]. Let $(x_t)_{0 \le s \le t \le T}$ be $1/p$-Hölder E-valued map. Then the continuous 2-index map*

$$z_{ts} := F^{(1)}_{(s, x_s)} (x_t - x_s) + \frac{1}{2} F^{(2)}_{(s, x_s)} (x_t - x_s)^{\otimes 2}$$

is almost-additive, and we have

$$F\big(t, x_t\big) = F\big(s, x_s\big) + \int_s^t \big(\partial_r F\big)(r, x_r)\, dr + \int_s^t z_{du}, \tag{2.17}$$

for any $0 \le s \le t \le T$.

PROOF. The proof is a straightforward application of Feyel-de la Pradelle' sewing lemma, Theorem 2.10. Given $0 \le s \le u \le t \le T$, the algebraic identity

$$\begin{aligned} z_{tu} + z_{us} &= F^{(1)}_{(s, x_s)} (x_t - x_s) + \left(F^{(1)}_{(u, x_u)} - F^{(1)}_{(s, x_s)} \right)(x_t - x_u) \\ &\quad + \frac{1}{2} F^{(2)}_{(u, x_u)} (x_t - x_u)^{\otimes 2} + \frac{1}{2} F^{(2)}_{(s, x_s)} (x_u - x_s)^{\otimes 2}, \end{aligned}$$

the regularity assumptions on F and the symmetric character of $F^{(2)}_{(s, x)}$, for any $x \in E$, we have

---

[1] More precisely, $F^{(n)}_{(t, y)} = D^n F(t, \cdot)(y) \in \mathrm{L}(E^{\otimes n}, \mathbb{R})$ for $n = 1, 2, 3$, $t \in [0, T]$ and $y \in E$.

$$z_{tu} + z_{us} = F^{(1)}_{(s,x_s)}(x_t - x_s) + F^{(2)}_{(s,x_s)}(x_u - x_s) \otimes (x_t - x_u)$$

$$+ O\Big(|t - s|^{3/p}\Big) + O\big(\|x_u - x_s\|^2\big)\,\|x_t - x_u\|$$

$$+ \frac{1}{2}F^{(2)}_{(u,x_u)}(x_t - x_u)^{\otimes 2} + \frac{1}{2}F^{(2)}_{(s,x_s)}(x_u - x_s)^{\otimes 2}$$

$$= z_{ts} + O\Big(|t - s|^{3/p}\Big).$$

Itô's formula (2.17) follows by noting that we have for all $n \geq 1$

$$F(t, x_t) - F(s, x_s) = \sum_{i=0}^{n-1} \Big\{ F\big(s_{i+1}, x_{s_{i+1}}\big) - F\big(s_i, x_{s_i}\big) \Big\}$$

$$= o_n(1) + \sum_{i=0}^{n-1}(s_{i+1} - s_i)\,(\partial_s F)\big(s_i, x_{s_i}\big) + \sum_{i=0}^{n-1} \Big\{ F\big(s_i, x_{s_{i+1}}\big) - F\big(s_i, x_{s_i}\big) \Big\},$$

with

$$F\big(s_i, x_{s_{i+1}}\big) - F\big(s_i, x_{s_i}\big)$$

$$= F^{(1)}_{(s_i, x_{s_i})}\big(x_{s_{i+1}} - x_{s_i}\big) + \frac{1}{2}F^{(2)}_{(s_i x_{s_i})}\big(x_{s_{i+1}} - x_{s_i}\big)^{\otimes 2} + O\Big(\big|x_{s_{i+1}} - x_{s_i}\big|^3\Big)$$

$$= z_{s_{i+1}s_i} + O\Big(|s_{i+1} - s_i|^{3/p}\Big). \qquad \square$$

As an example, consider the solution flow φ to a rough differential equation on $\mathbb{R}^d$

$$d\varphi = \boldsymbol{V}(\varphi\,;dt).$$

Write φ_t for φ_{t0}, and consider it as an element of the space E of continuous paths from $[0, T]$ to $\mathcal{C}(\mathbb{R}^d, \mathbb{R}^d)$, equipped with the norm of uniform convergence, with $\mathcal{C}(\mathbb{R}^d, \mathbb{R}^d)$ endowed with a norm inducing uniform convergence on compact sets. It satisfies by its very definition and Proposition 2.5 the Euler–Taylor expansion

$$\varphi_t = \varphi_s + \big(V_{ts}\mathrm{Id}\big) \circ \varphi_s + \big(\mathbb{V}_{ts}\mathrm{Id}\big) \circ \varphi_s + O\big(|t - s|^{3/p}\big)$$

so it is a $1/p$-Hölder path in that space. Now, given some points $y_1, \ldots, y_k$ in $\mathbb{R}^d$ and a $\mathcal{C}^3_b$-real-valued function f on $(\mathbb{R}^d)^k$, one can think of the function

$$F(\phi) = f\big(\phi(y_1), \ldots, \phi(y_k)\big), \tag{2.18}$$

for $\phi \in E$, as a typical time-independent example of function satisfying the conditions of Theorem 2.11. One then has

$$F(\varphi_{s_{i+1}}) - F(\varphi_{s_i}) = f\Big(\varphi_{s_{i+1}s_i}\big(\varphi_{s_i}(y_1)\big), \ldots, \varphi_{s_{i+1}s_i}\big(\varphi_{s_i}(y_k)\big)\Big) - f\big(\varphi_{s_i}(y_1), \ldots, \varphi_{s_i}(y_k)\big)$$

$$= \sum_{m=1}^{k} \Big(\big(V^{\{m\}}_{s_{i+1}s_i} + \mathbb{V}^{\{m\}}_{s_{i+1}s_i}\big)f\Big)\big(\varphi_{s_i}(y_1), \ldots, \varphi_{s_i}(y_k)\big) + O_c\Big(|s_{i+1} - s_i|^{3/p}\Big),$$

where the upper index $\{m\}$ means that the operators act on the mth component of f. The above sum defines an almost-additive continuous function z_{ts}^f, taken here at time (s_{i+1}, s_i), so we have

$$f\big(\varphi_t(y_1), \ldots, \varphi_t(y_k)\big) = f\big(\varphi_s(y_1), \ldots, \varphi_s(y_k)\big) + \int_s^t z_{du}^f$$

for all times $0 \leq s \leq t \leq T$.

2.5. A Kolmogorov-type regularity theorem.

We shall use below the theory of rough drivers in a setting where the drivers are random. Like in the theory of rough paths, the primary object we are given is not the random rough driver itself, or the random rough path, but rather a genuine random vector field, or random path, which needs to be enhanced in a first step into a random rough driver, or random rough path. This first, purely probabilistic, step can typically be done using some Kolmogorov-type continuity arguments. We give in this Section some variations on this theme that will be used to enhance vector field-valued martingales into rough drivers in Section 3.3; *a reader interested only in these applications is advised to skip the technical details and only have a look at Theorem* 2.16; for the other readers, it is our hope that this somewhat long Section contains some material interesting in itself; it provides moment conditions under which one can get back uniform in time estimates on quantities of the form $(t-s)^{-\alpha}\|X_{ts}\|_{\mathcal{C}^a}$, such as required by the definition of a rough driver.

The next Lemma gives sufficient conditions for a process defined on a possibly unbounded domain to be bounded. Recall the equivalence of having Gaussian tails to square-root growth of moments, cf. [**FV10**, Lemma A.17].

LEMMA 2.12. *Let (E, d_E) be a complete, separable metric space. Let D be an open subset of $\mathbb{R}^d$, $X \colon D \to (E, d_E)$ a continuous stochastic process, $e \in E$ and $\kappa > 0$. Set*

$$D_n := \big\{x \in D \,:\, n - 1 \leq |x| < n\big\}$$

and $\mathbf{N} := \{n \in \mathbb{N} \,:\, D_n \neq \emptyset\}$. Let $(a_n)_{n \in \mathbf{N}}$ be a sequence of non-negative real numbers and $(x_n)_{n \in \mathbf{N}}$ a sequence of elements in D such that $x_n \in D_n$ for every $n \in \mathbf{N}$.

(i) *Assume that there are $q \in (q, \infty)$ and $\gamma \in (d/q, 1]$ such that*

$$\sup_{x,y \in D_n} \Big\|d_E\big(X(x), X(y)\big)\Big\|_{L^q} \leq \kappa a_n |x - y|^\gamma$$

and that

$$\Big\|d_E\big(X(x_n), e\big)\Big\|_{L^q} \leq \kappa a_n$$

for every $n \in \mathbf{N}$. Set $(b_n) := (a_n n^\gamma)$ and assume that $\|b\|_{\ell^q} \leq K < \infty$. Then there is a constant $C = C(q, \gamma)$ such that

$$\Big\| \sup_{x \in D} d_E\big(X(x), e\big)\Big\|_{L^q} \leq CK\kappa.$$

(ii) *Let $\gamma \in (0, 1]$ and assume that for every $q \geq 1$ there is a constant c_q such that for every $n \in \mathbf{N}$,*

$$\sup_{x,y \in D_n} \left\| d_E\big(X(x), X(y)\big) \right\|_{L^q} \leq \kappa c_q a_n |x - y|^\gamma$$

and that

$$\left\| d_E(X(x_n), e) \right\|_{L^q} \leq \kappa c_q a_n$$

where $c_q = \mathcal{O}(\sqrt{q})$ when $q \to \infty$. Assume that $a_n = \mathcal{O}(n^{-\gamma}(1 + \log n)^{-1/2})$. Then for every $q \geq 1$ there is some constant $C = C(q, \gamma)$ such that

$$\left\| \sup_{x \in D} d_E\big(X(x), e\big) \right\|_{L^q} \leq C\kappa$$

with $C = \mathcal{O}(\sqrt{q})$ when $q \to \infty$. In particular, the random variable $\sup_{x \in D} d_E(X(x), e)$ has a Gaussian tail.

PROOF.   Without loss of generality, one may assume $\kappa = 1$, otherwise we consider the metric $\tilde{d}_E = d_E/\kappa$ instead, and $\mathbf{N} = \mathbb{N}$ – otherwise we add small, disjoint balls to D and define X to be constant and equal to e on these balls. We first prove Claim (i).

Let $\alpha > d/q$ and set $p(u) = u^{\alpha + d/q}$. By the Garsia–Rodemich–Rumsey Lemma (cf. e.g. [**Sch09**, Lemma 2.4 (i)]), for every $x, y \in D_n$,

$$\frac{d_E\big(X(x), X(y)\big)}{|x - y|^{\alpha - d/q}} \leq C V_n^{1/q}$$

where

$$V_n = \int_{D_n \times D_n} \frac{\left| d_E\big(X(u), X(v)\big) \right|^q}{|u - v|^{\alpha q + d}} \, du \, dv.$$

Thus, by a change of variables,

$$\mathbb{E} \left| \sup_{x,y \in D_n} \frac{d_E\big(X(x), X(y)\big)}{|x - y|^{\alpha - d/q}} \right|^q \leq C^q a_n^q \int_{D_n \times D_n} |u - v|^{(\gamma - \alpha)q - d} \, du \, dv$$

$$\leq C^q a_n^q n^{d + (\gamma - \alpha)q} \int_{(0,1)^2} |u - v|^{(\gamma - \alpha)q - d} \, du \, dv.$$

Let $\beta \in (0, \gamma - d/q)$ and set $\alpha = d/q + \beta < \gamma$. Then the integral is finite, and we have shown that

$$\left\| \sup_{x,y \in D_n} \frac{d_E\big(X(x), X(y)\big)}{|x - y|^\beta} \right\|_{L^q} \leq C a_n n^{(\gamma - \beta)}.$$

By the triangle inequality,

$$\left\| \sup_{x \in D_n} d_E\big(X(x), e\big) \right\|_{L^q} \le C a_n (n^\gamma + 1) \le 2C b_n.$$

Thus we obtain

$$\mathbb{E}\left(\sup_{x \in D} d_E\big(X(x), e\big)^q \right) = \mathbb{E}\left(\sup_n \sup_{x \in D_n} d_E\big(X(x), e\big)^q \right) \le \sum_{n=1}^{\infty} \mathbb{E}\left(\sup_{x \in D_n} d_E\big(X(x), e\big)^q \right)$$

$$\le 2^q C^q \sum_{n=1}^{\infty} b_n^q < \infty$$

and claim (i) is shown.

Now we prove claim (ii). Note that the constant in the Garsia–Rodemich–Rumsey Lemma may be chosen non-increasing in q. Therefore, we can argue similarly as before to see that for every $q \ge 1$ and $n \in \mathbb{N}$,

$$\left\| \sup_{x \in D_n} d_E\big(X(x), e\big) \right\|_{L^q} \le C_q b_n$$

where $C_q = \mathcal{O}(\sqrt{q})$. This shows that the random variable has Gaussian tails, i.e. there is some constant C such that for every $n \in \mathbb{N}$

$$\mathbb{P}\left(\sup_{x \in D_n} d_E\big(X(x), e\big) \ge t \right) \le C \exp\left(-\frac{t^2}{C b_n^2} \right)$$

for every $t \ge 0$. Hence

$$\mathbb{P}\left(\sup_{x \in D} d_E\big(X(x), e\big) \ge t \right) \le C \sum_{n=1}^{\infty} \exp\left(-\frac{t^2}{C b_n^2} \right) \le C \sum_{n=1}^{\infty} \exp\left(-\frac{t^2}{C}(1 + \log n) \right)$$

$$\le C \exp\left(-\frac{t^2}{C} \right) \sum_{n=1}^{\infty} n^{-t^2/C}$$

and the sum is finite for t large enough. This proves that $\sup_{x \in D} d_E(X(x), e)$ has Gaussian tails. □

COROLLARY 2.13. *Let D be an open subset of $\mathbb{R}^d$, $(E, \| \cdot \|)$ a separable Banach space, $X \colon D \to E$ a continuous stochastic process and $q > 1$.*

(i) *Assume that there are $\kappa > 0$ and $\gamma \in (0, 1]$ such that for every $x, y \in D$ with $0 < |x - y| \le 1$,*

$$\big\| X(x) - X(y) \big\|_{L^q} \le \kappa |x - y|^\gamma \tag{2.19}$$

and that there is an $\eta \in (0, \infty)$ such that for every $x \in D$,

$$\| X(x) \|_{L^q} \le \frac{\kappa}{1 + |x|^\eta} \tag{2.20}$$

with q sufficiently large satisfying

$$q > \frac{1}{\eta} + d\left(\frac{1}{\eta} + \frac{1}{\gamma}\right).$$

Then the random variable $\sup_{x \in D} \|X(x)\|$ is almost surely finite. Moreover, there is a constant $C = C(\gamma, q, \eta)$ such that

$$\left\|\sup_{x \in D} \|X(x)\|\right\|_{L^q} \leq C\kappa. \tag{2.21}$$

(ii) Assume that (2.19) and (2.20) hold for every $q \geq 1$ with $\kappa = \kappa(q) \leq \sqrt{q}\hat{\kappa}$ and some $\eta \in (0, \infty)$.

Then $\sup_{x \in D} \|X(x)\|$ has Gaussian tails and there is a constant $C = C(\gamma, \eta)$ such that

$$\left\|\sup_{x \in D} \|X(x)\|\right\|_{L^q} \leq C\sqrt{q}\hat{\kappa}$$

for every $q \geq 1$.

PROOF.   We start with proving (i). Let $x, y \in D$ such that $|x - y| \geq 1$. Then, by (2.20),

$$\left\|X(x) - X(y)\right\|_{L^q} \leq \left\|X(x)\right\|_{L^q} + \left\|X(y)\right\|_{L^q} \leq 2\kappa \leq 2\kappa|x - y|^\gamma$$

which shows that

$$\left\|X(x) - X(y)\right\|_{L^q} \leq 2\kappa|x - y|^\gamma \tag{2.22}$$

holds for every $x, y \in D$. Now let $x, y \in D$ such that $n - 1 \leq |x|, |y| < n$. Interpolating between the inequality

$$\left\|X(x) - X(y)\right\|_{L^q} \leq \left\|X(x)\right\|_{L^q} + \left\|X(y)\right\|_{L^q} \leq \frac{2\kappa}{1 + (n-1)^\eta}$$

and inequality (2.22), we see that for every $\lambda \in [0, 1]$,

$$\left\|X(x) - X(y)\right\|_{L^q} \leq C\kappa n^{-(1-\lambda)\eta}|x - y|^{\gamma\lambda}$$

and

$$\left\|X(x)\right\|_{L^q} \leq C\kappa n^{-(1-\lambda)\eta}$$

for every $x, y \in D_n$. Set $\gamma' := \gamma\lambda$ and $a_n := n^{-(1-\lambda)\eta}$. In order to obtain $(a_n n^{\gamma'}) \in \ell^q(\mathbb{N})$, we must have $q(\lambda\gamma - (1-\lambda)\eta) < -1$ which is equivalent to

$$\frac{1}{q} < \eta - \lambda(\eta + \gamma).$$

The condition $\gamma' > d/q$ is equivalent to

$$\frac{1}{q} < \frac{\lambda\gamma}{d}.$$

Choosing $\lambda^* = \eta/(\gamma/d + \eta + \gamma) \in (0,1)$, we have

$$\eta - \lambda^*(\eta + \gamma) = \frac{\lambda^*\gamma}{d} = \frac{\gamma\eta}{\gamma + d(\gamma + \eta)}$$

which is indeed smaller than $1/q$ by assumption. Hence we may apply Lemma 2.12 to conlude (i). The claim (ii) follows by applying Lemma 2.12 (ii). □

EXAMPLE. Consider the Gaussian process $X\colon (0,\infty) \to \mathbb{R}$ where $X_t = B_t/(1 + t^\alpha)$, for a standard Brownian motion B, and $\alpha \in (1/2, 1]$. Then, if $t > 0$,

$$\|X_t\|_{L^q} \lesssim \sqrt{q}\|X_t\|_{L^2} \lesssim \frac{\sqrt{q}}{1 + t^{\alpha - 1/2}}$$

and for $s < t$,

$$
\begin{aligned}
\|X_t - X_s\|_{L^2} &= \frac{\|B_t(1 + s^\alpha) - B_s(1 + t^\alpha)\|_{L^2}}{(1 + s^\alpha)(1 + t^\alpha)} \leq \frac{\|B_t - B_s\|_{L^2}}{1 + t^\alpha} + \frac{\|B_s\|_{L^2}|t^\alpha - s^\alpha|}{(1 + s^\alpha)(1 + t^\alpha)} \\
&\leq \frac{|t - s|^{1/2}}{1 + t^\alpha} + \frac{2s^{1/2}t^{\alpha - 1/2}}{(1 + s^\alpha)(1 + t^\alpha)}|t - s|^{1/2} \\
&\leq |t - s|^{1/2}
\end{aligned}
$$

and

$$\|X_t - X_s\|_{L^q} \lesssim \sqrt{q}\|X_t - X_s\|_{L^2}.$$

Applying part (ii) in Corollary 2.13 shows that the random variable

$$\sup_{t \in (0,\infty)} |X_t|$$

is finite and has Gaussian tails. Note that this is sharp in the sense that the law of the iterated logarithm for a Brownian motion implies that it is not possible to choose $\alpha = 1/2$.

Next, we apply the same ideas to give conditions for Hölder continuity.

LEMMA 2.14. *Let (E, d_E) be a complete separable metric space, D an open subset of $\mathbb{R}^d$, $X\colon D \to (E, d)$ a continuous stochastic process and $\kappa > 0$. Set*

$$D_n := \Big\{x \in D : n - 1 \leq |x| < n\Big\}$$

and $\mathbf{N} := \{n \in \mathbb{N} : D_n \neq \emptyset\}$. Let $(a_n)_{n \in \mathbf{N}}$ be a sequence of non-negative real numbers.

(i) *Assume that there are $q > 1$ and $\gamma \in (d/q, 1]$ such that for every $n \in \mathbf{N}$ and every $x, y \in D_n$ with $0 < |x - y| \leq 4$,*

$$\|d_E(X(x), X(y))\|_{L^q} \le \kappa a_n |x - y|^\gamma.$$

Let $\beta \in (0, \gamma - d/q)$. Define the sequence $(b_n) := (a_n n^{\gamma-\beta})$, and assume that $\|b\|_{L^q} \le K < \infty$. Then there is a constant $C = C(q, \gamma)$ such that

$$\left\|\sup_{\substack{x,y \in D \\ 0 < |x-y| \le 1}} \frac{d_E(X(x), X(y))}{|x - y|^\beta}\right\|_{L^q} \le CK\kappa.$$

(ii) *Assume that there is some $\gamma \in (0, 1]$ and that for every $q \ge 1$ there is a positive constant c_q such that for every $n \in \mathbf{N}$ and every $x, y \in D_n$ with $0 < |x - y| \le 4$,*

$$\left\|d_E(X(x), X(y))\right\|_{L^q} \le \kappa c_q a_n |x - y|^\gamma$$

where $c_q = \mathcal{O}(\sqrt{q})$ when $q \to \infty$. Let $\beta \in (0, \gamma)$ and assume that $a_n = \mathcal{O}(n^{-(\gamma-\beta)}(1+\log n)^{-1/2})$. Then for every $q \ge 1$ there is some constant $C = C(q, \gamma)$ such that

$$\left\|\sup_{\substack{x,y \in D \\ 0 < |x-y| \le 1}} \frac{d_E(X(x), X(y))}{|x - y|^\beta}\right\|_{L^q} \le C\kappa$$

with $C = \mathcal{O}(\sqrt{q})$ when $q \to \infty$. In particular, the random variable

$$\sup_{\substack{x,y \in D \\ 0 < |x-y| \le 1}} \frac{d_E(X(x), X(y))}{|x - y|^\beta}$$

has Gaussian tails.

Proof.   Without loss of generality, one can choose $\kappa = 1$ and $\mathbf{N} = \mathbb{N}$. For $n \in \mathbb{N}$, set $\tilde{D}_n := \{D_n \cup D_{n+1} \cup D_{n+2}\}$. We first prove (i). Fix some $n \in \mathbb{N}$ and some $k \in \mathbb{N}$. Let $\alpha > d/q$ and define

$$p_k(s) = \begin{cases} s^{\alpha+(d/q)} & \text{if } s \in [0, 4], \\ (4^{\alpha q+d} + k(s - 4))^{1/q} & \text{if } s \ge 4. \end{cases}$$

Fix $x, y \in \tilde{D}_n$ with $0 < |x - y| \le 1$. From the Garsia–Rodemich–Rumsey Lemma,

$$d_E(X(x), X(y)) \le C V_{n,k}^{1/q} |x - y|^{\alpha - d/q}$$

where

$$V_{n,k} = \int_{\tilde{D}_n \times \tilde{D}_n} \frac{\left|d_E(X(u), X(v))\right|^q}{p_k(|u - v|)^q} \, du \, dv.$$

Thus

$$
\mathbb{E}\left|\sup_{\substack{x,y\in\tilde{D}_n \\ 0<|x-y|\leq1}}\frac{d_E\big(X(x),X(y)\big)}{|x-y|^{\alpha-d/q}}\right|^q \leq C^q\int_{\tilde{D}_n\times\tilde{D}_n}\frac{\mathbb{E}\Big[d_E\big(X(u),X(v)\big)^q\Big]}{p_k(|u-v|)^q}\,du\,dv
$$

$$
\leq C^q\sum_{l=0,1,2}\int_{D_{n+l}\times D_{n+l}}\frac{\mathbb{E}\Big[d_E\big(X(u),X(v)\big)^q\Big]}{p_k(|u-v|)^q}\,du\,dv.
$$

For every $m\in\mathbb{N}$, we have

$$
\int_{D_m\times D_m}\frac{\mathbb{E}\Big[d_E\big(X(u),X(v)\big)^q\Big]}{p_k(|u-v|)^q}\,du\,dv \leq a_m^q\int_{(D_m\times D_m)\cap\{|u-v|\leq4\}}|u-v|^{(\gamma-\alpha)q-d}\,du\,dv
$$

$$
+\int_{(D_m\times D_m)\cap\{|u-v|>4\}}\frac{\mathbb{E}\Big[d_E\big(X(u),X(v)\big)^q\Big]}{4^{\alpha q+d}+k(|u-v|-4)}\,du\,dv.
$$

Moreover, by a change of variables,

$$
\int_{(D_m\times D_m)\cap\{|u-v|\leq4\}}|u-v|^{(\gamma-\alpha)q-d}\,du\,dv \leq \int_{D_m\times D_m}|u-v|^{(\gamma-\alpha)q-d}\,du\,dv
$$

$$
= m^{d+(\gamma-\alpha)q}\int_{(0,1)^2}|u-v|^{(\gamma-\alpha)q-d}\,du\,dv.
$$

Set $\alpha=d/q+\beta<\gamma$. Then this integral is finite, and sending $k\to\infty$ shows that

$$
\left\|\sup_{\substack{x,y\in\tilde{D}_n \\ 0<|x-y|\leq1}}\frac{d_E\big(X(x),X(y)\big)}{|x-y|^\beta}\right\|_{L^q} \leq C\Big(a_n n^{(\gamma-\beta)}+a_{n+1}(n+1)^{(\gamma-\beta)}+a_{n+2}(n+2)^{(\gamma-\beta)}\Big)
$$

$$
= C\big(b_n+b_{n+1}+b_{n+2}\big).
$$

Now take $x,y\in D$ with $0<|x-y|\leq1$ and assume that

$$
\frac{d_E\big(X(x),X(y)\big)}{|x-y|^\beta}\geq t.
$$

Then there is an $n\in\mathbb{N}$ such that $x\in D_n$ and since $|x-y|\leq1$, $y\in\{D_{n-1}\cup D_n\cup D_{n+1}\}$, where we set $D_0:=D_1$. Thus we have shown that for every $t\geq0$,

$$
\left\{\sup_{\substack{x,y\in D \\ 0<|x-y|\leq1}}\frac{d_E\big(X(x),X(y)\big)}{|x-y|^\beta}\geq t\right\}\subseteq\bigcup_{n\in\mathbb{N}}\left\{\sup_{\substack{x,y\in\tilde{D}_n \\ 0<|x-y|\leq1|}}\frac{d_E\big(X(x),X(y)\big)}{|x-y|^\beta}\geq t\right\}
$$

and therefore

$$
\mathbb{E}\left|\sup_{\substack{x,y\in D \\ 0<|x-y|\leq1}}\frac{d_E\big(X(x),X(y)\big)}{|x-y|^\beta}\right|^q \leq\sum_{n=1}^\infty\mathbb{E}\left|\sup_{\substack{x,y\in\tilde{D}_n \\ 0<|x-y|\leq1}}\frac{d_E\big(X(x),X(y)\big)}{|x-y|^\beta}\right|^q \leq 3C^qK^q.
$$

Now we prove (ii). Note that the constant in the Garsia–Rodemich–Rumsey Lemma may be chosen non-increasing in q. Therefore, we can argue similarly as before to see that for every $q \geq 1$ and $n \in \mathbb{N}$,

$$
\left\| \sup_{\substack{x,y \in \tilde{D}_n \\ 0 < |x-y| \leq 1}} \frac{d_E\big(X(x), X(y)\big)}{|x-y|^\beta} \right\|_{L^q} \leq C_q\big(b_n + b_{n+1} + b_{n+2}\big)
$$

where $C_q = \mathcal{O}(\sqrt{q})$. This shows that the random variable has Gaussian tails, i.e. there is some constant C such that for every $n \in \mathbb{N}$

$$
\mathbb{P}\left(\sup_{\substack{x,y \in \tilde{D}_n \\ 0 < |x-y| \leq 1}} \frac{d_E\big(X(x), X(y)\big)}{|x-y|^\beta} \geq t \right) \leq C \exp\left(-\frac{t^2}{C(b_n + b_{n+1} + b_{n+2})^2} \right)
$$

for every $t \geq 0$. Hence

$$
\begin{aligned}
\mathbb{P}\left(\sup_{\substack{x,y \in D \\ 0 < |x-y| \leq 1}} \frac{d_E\big(X(x), X(y)\big)}{|x-y|^\beta} \geq t \right) &\leq C \sum_{n=1}^{\infty} \exp\left(-\frac{t^2}{C(b_n + b_{n+1} + b_{n+2})^2} \right) \\
&\leq C \sum_{n=1}^{\infty} \exp\left(-\frac{t^2}{C}(1 + \log n) \right) \\
&\leq C \exp\left(-\frac{t^2}{C} \right) \sum_{n=1}^{\infty} n^{-t^2/C}
\end{aligned}
$$

and the sum is finite for t large enough. This proves that

$$
\sup_{\substack{x,y \in D \\ 0 < |x-y| \leq 1}} \frac{d_E\big(X(x), X(y)\big)}{|x-y|^\beta}
$$

has Gaussian tails. $\qquad\square$

EXAMPLE.   Let $X \colon (0, \infty) \to \mathbb{R}$ be the Gaussian process defined as

$$
X_t = \frac{B_t}{\sqrt{t \log(1+t)}},
$$

B being a standard Brownian motion. Then

$$
\|X_t - X_s\|_{L^2} \leq \frac{|t-s|^{1/2}}{\sqrt{t \log(1+t)}} + \frac{\sqrt{t \log(1+t)} - \sqrt{s \log(1+s)}}{\sqrt{\log(1+s)}\sqrt{s \log(1+s)}}.
$$

By the mean value theorem,

$$
t \log(1+t) - s \log(1+s) \leq (\log(1+t) + 1)(t - s)
$$

and therefore

$$\sqrt{t\log(1+t)} - \sqrt{s\log(1+s)} \le \sqrt{t\log(1+t) - s\log(1+s)} \le \sqrt{(\log(1+t)+1)}|t-s|^{1/2}.$$

If $(n-1) \le s \le t \le n$, we have for any $q \ge 2$

$$\|X_t - X_s\|_{L^q} \lesssim \sqrt{q}\|X_t - X_s\|_{L^2} \lesssim \sqrt{q}a_n|t-s|^{1/2}$$

with $a_n = \mathcal{O}(n^{-1/2}(1+\log n)^{-1/2})$. Part (ii) of Lemma 2.14 shows that for any $\beta \in (0, 1/2)$, the random variable

$$\sup_{\substack{s,t\in(0,\infty)\\0<|t-s|\le 1}} \frac{|X_t - X_s|}{|t-s|^\beta}$$

is finite and has Gaussian tails.

COROLLARY 2.15. *Let D be an open subset of $\mathbb{R}^d$, and $(E, \|\cdot\|)$ be a separable Banach space. Let $X\colon D \to E$ be a continuous stochastic process and $q > 1$.*

(i) *Assume that there are constants $\kappa > 0$, $\gamma \in (0, 1]$ and $\beta \in (0, \gamma)$ such that for every $x, y \in D$ with $0 < |x - y| \le 1$,*

$$\big\|X(x) - X(y)\big\|_{L^q} \le \kappa|x-y|^\gamma \tag{2.23}$$

and that there is $\eta \in (0, \infty)$ such that for every $x \in D$,

$$\big\|X(x)\big\|_{L^q} \le \frac{\kappa}{1+|x|^\eta} \tag{2.24}$$

where $q > 1$ satisfies

$$q > \frac{\gamma}{\eta(\gamma-\beta)} + d\left(\frac{\gamma}{\eta(\gamma-\beta)} + \frac{1}{\gamma-\beta}\right).$$

Then the random variable

$$\sup_{\substack{x,y\in D\\0<|x-y|\le 1}} \frac{\|X(x) - X(y)\|}{|x-y|^\beta}$$

is almost surely finite. Moreover, there is a constant $C = C(\gamma, \eta, \beta, q)$ such that

$$\left\|\sup_{\substack{x,y\in D\\0<|x-y|\le 1}} \frac{\|X(x) - X(y)\|}{|x-y|^\beta}\right\|_{L^q} \le C\kappa. \tag{2.25}$$

(ii) *Assume that (2.23) and (2.24) hold for every $q \ge 1$ with $\kappa = \kappa(q) \le \sqrt{q}\hat{\kappa}$ and some $\eta \in (0, \infty)$.*

Then for every $\beta \in (0, \gamma)$,

$$\sup_{\substack{x,y\in D \\ 0<|x-y|\leq 1}} \frac{\big\|X(x)-X(y)\big\|}{|x-y|^\beta}$$

has Gaussian tails, and there is a constant $C = C(\gamma,\eta,\beta)$ such that

$$\left\| \sup_{\substack{x,y\in D \\ 0<|x-y|\leq 1}} \frac{\big\|X(x)-X(y)\big\|}{|x-y|^\beta} \right\|_{L^q} \leq C\sqrt{q}\hat{\kappa}$$

holds for every $q \geq 1$.

Proof.   The proof is similar to the proof of Corollary 2.13, using Lemma 2.14 above. We leave the details to the reader. $\square$

If D is an open subset of $\mathbb{R}^d$, and $(E,\|\cdot\|)$ is a normed space, $f\colon D\to E$ a function and $\rho \in (0,1]$, we define

$$\|f\|_{\widetilde{\mathcal{C}}^\rho} := \max\left\{ 2\sup_{x\in D}\big\|f(x)\big\|,\quad \sup_{\substack{x,y\in D \\ 0<|x-y|\leq 1}} \frac{\big\|f(x)-f(y)\big\|}{|x-y|^\rho} \right\}.$$

Let $f,g\colon D\to\mathbb{R}^m$. Then we define the function $(f\odot g)\colon D\to\mathbb{R}^{m\times m}$ by setting $(f\odot g)^{ij}(x) = f^i(x)g^j(x)$. Note that $\|\cdot\|_{\widetilde{\mathcal{C}}^\rho}$ is equivalent to $\|\cdot\|_{\mathcal{C}^\rho}$ and that we have

$$\|f\odot g\|_{\widetilde{\mathcal{C}}^\rho} \leq \|f\|_{\widetilde{\mathcal{C}}^\rho}\|g\|_{\widetilde{\mathcal{C}}^\rho} \tag{2.26}$$

provided we equip $\mathbb{R}^m$ and $\mathbb{R}^{m\times m}$ with the sup norm.

In the following, we will consider stochastic processes $V\colon D\times[0,T]\to\mathbb{R}^m$ and $W\colon D\times\{0\leq s\leq t\leq T\}\to\mathbb{R}^{m\times m}$ for which we assume that for every $s<u<t\in[0,T]$ and every $x\in D$,

$$W_{ts}(x) - W_{us}(x) - W_{tu}(x) = V_{us}(x)\odot V_{tu}(x) \tag{2.27}$$

holds almost surely. The next theorem is the main result of this Section.

Theorem 2.16 (Kolmogorov criterion for rough drivers).   *Let D be an open subset of $\mathbb{R}^d$, $\kappa > 0$ and $\gamma_1,\gamma_2 \in (0,1]$.*

(i) *Let $V\colon D\times I\to\mathbb{R}^m$ be a stochastic process and $q > 1$. Assume that for every $x,y\in D$ with $0 < |x-y| \leq 1$ and $s < t \in I$,*

$$\big\|V_{ts}(x)-V_{ts}(y)\big\|_{L^q} \leq \kappa|t-s|^{\gamma_1}|x-y|^{\gamma_2} \tag{2.28}$$

and that there is $\eta \in (0,\infty)$ such and that for every $x\in D$ and $s < t\in I$,

$$\big\|V_{ts}(x)\big\|_{L^q} \leq \frac{\kappa|t-s|^{\gamma_1}}{1+|x|^\eta}. \tag{2.29}$$

Let $\alpha \in (0, \gamma_1)$, $\beta \in (0, \gamma_2)$ and assume that

$$q > \max \left\{ \frac{\gamma_2}{\eta(\gamma_2 - \beta)} + d \left(\frac{\gamma_2}{\eta(\gamma_2 - \beta)} + \frac{1}{\gamma_2 - \beta} \right), \frac{1}{\gamma_1 - \alpha} \right\}. \tag{2.30}$$

Then there is a continuous modification of the process V. Moreover, there is a constant $C = C(\gamma_1, \gamma_2, \alpha, \beta, \eta, d, T, q)$ such that

$$\left\| \sup_{s < t \in I} \frac{\|V_{ts}\|_{\mathcal{C}^\beta}}{|t - s|^\alpha} \right\|_{L^q} \leq C\kappa. \tag{2.31}$$

(ii) *In addition, let $W \colon D \times \{0 \leq s \leq t \leq T\} \to \mathbb{R}^{m \times m}$ be a stochastic process for which the relation (2.26) holds and $q > 2$. Assume that for every $x, y \in D$ with $0 < |x - y| \leq 1$ and $s < t \in I$,*

$$\|W_{ts}(x) - W_{ts}(y)\|_{L^{q/2}} \leq \kappa^2 |t - s|^{2\gamma_1} |x - y|^{\gamma_2} \tag{2.32}$$

and that there is an $\eta \in (0, \infty)$ such that for every $x \in D$ and $s < t \in I$,

$$\left\| W_{ts}(x) \right\|_{L^{q/2}} \leq \frac{\kappa^2 |t - s|^{2\gamma_1}}{1 + |x|^{2\eta}}. \tag{2.33}$$

Let $\alpha \in (0, \gamma_1)$, $\beta \in (0, \gamma_2)$ and assume that q is sufficiently large such that (2.30) holds. Then there is a continuous modification of the process W. Moreover, there is a constant $C = C(\gamma_1, \gamma_2, \alpha, \beta, \eta, d, T, q)$ such that

$$\left\| \sup_{s < t \in I} \frac{\|W_{ts}\|_{\mathcal{C}^\beta}}{|t - s|^{2\alpha}} \right\|_{L^{q/2}} \leq C^2 \kappa^2. \tag{2.34}$$

(iii) *Assume in addition that (2.28), (2.32) and the growth conditions (2.29) and (2.33) hold for every $q \geq 2$ with constant $\kappa \leq \sqrt{q}\hat{\kappa}$. Then for every $\alpha \in (0, \gamma_1)$ and $\beta \in (0, \gamma_2)$, the random variables*

$$\sup_{s < t \in I} \frac{\|V_{ts}\|_{\mathcal{C}^\beta}}{|t - s|^\alpha} \qquad and \qquad \sqrt{\sup_{s < t \in I} \frac{\|W_{ts}\|_{\mathcal{C}^\beta}}{|t - s|^{2\alpha}}}$$

have Gaussian tails.

PROOF. Without loss of generality we may assume $\kappa = 1$, otherwise we can replace V and W by V/κ and W/κ^2 respectively. Furthermore, we will prove the result for the $\| \cdot \|_{\widetilde{\mathcal{C}}^\beta}$ norm, claimed results follow by equivalence of norms.

We start with proving (i). Fix $s < t$. Using (2.28) and the classical Kolmogorov theorem [**Kal02**, Theorem 3.23], there is a continuous modification of the process $x \mapsto V_{ts}(x)$ on D. The estimates (2.28) and (2.29) and Corollary 2.15 imply that

$$\left\|\sup_{\substack{x,y\in D\\0<|x-y|\leq 1}}\frac{|V_{ts}(x)-V_{ts}(y)|}{|x-y|^{\beta}}\right\|_{L^q}\leq C|t-s|^{\gamma_1}$$

and Corollary 2.13 gives

$$\left\|\sup_{x\in D}|V_{ts}(x)|\right\|_{L^q}\leq C|t-s|^{\gamma_1}.$$

Note in particular that the constant on the right hand side of both equations is independent of s and t. We can repeat this procedure for every $s<t$ and obtain a process $t\mapsto V_t$ which, for every $t\in[0,T]$, takes values in $\mathcal{C}_b^{\beta}$ almost surely, and for which

$$\left\|\left\||V_t-V_s\right\|_{\mathcal{C}^{\beta}}^{\sim}\right\|_{L^q}\leq C|t-s|^{\gamma_1} \tag{2.35}$$

holds for every $s<t$. Applying again the Kolmogorov theorem for Banach space valued processes gives the claim.

We proceed with (ii). As in (i), for every $s<t$ there are modifications of the process $x\mapsto W_{ts}(x)$ such that

$$\left\|\left\|W_{ts}\right\|_{\mathcal{C}^{\beta}}^{\sim}\right\|_{L^{q/2}}\leq C|t-s|^{2\gamma_1}.$$

Using the algebraic relation (2.27), the estimate (2.35) for V and the compatibility of the $\|\cdot\|_{\mathcal{C}^{\beta}}^{\sim}$ norms given in (2.26), we can mimic the proof of the Kolmogorov criterion for rough paths ([**FH14**, Theorem 3.1]) to conlude.

Assertion (iii) follows similarly by using part (ii) in the Corollaries 2.15 and 2.13. $\quad\square$

Finally, we give a Kolmogorov criterion for the distance between rough drivers, whose proof is very similar to the proof of Theorem 2.16, using the Kolmogorov criterion for rough path distance [**FH14**, Theorem 3.3]; we leave it to the reader.

Theorem 2.17 (Kolmogorov criterion for rough driver distance). *Let $\kappa>0$, $\gamma_1,\gamma_2\in(0,1]$ and (V,W), $(\hat{V},\hat{W})$ processes as in Theorem 2.16. Set $\Delta V:=V-\hat{V}$ and $\Delta W:=W-\hat{W}$.*

(i) *Assume that (V,W) and $(\hat{V},\hat{W})$ satisfy the same moment conditions as in Theorem 2.16 with q sufficiently large as in (2.30). Moreover, assume that there is an $\varepsilon>0$ such that*

$$\left\|\Delta V_{ts}(x)-\Delta V_{ts}(y)\right\|_{L^q}\leq\varepsilon\kappa|t-s|^{\gamma_1}|x-y|^{\gamma_2} \tag{2.36}$$

and

$$\left\|\Delta W_{ts}(x)-\Delta W_{ts}(y)\right\|_{L^{q/2}}\leq\varepsilon\kappa^2|t-s|^{2\gamma_1}|x-y|^{\gamma_2} \tag{2.37}$$

for every $x,y\in D$ such that $0<|x-y|\leq 1$ and every $s<t\in[0,T]$ and that

$$\left\|\Delta V_{ts}(x)\right\|_{L^q} \leq \varepsilon \frac{\kappa |t-s|^{\gamma_1}}{1+|x|^{\eta}} \tag{2.38}$$

and

$$\left\|\Delta W_{ts}(x)\right\|_{L^{q/2}} \leq \varepsilon \frac{\kappa^2 |t-s|^{2\gamma_1}}{1+|x|^{2\eta}} \tag{2.39}$$

hold for every $x \in D$ and every $s < t \in [0,T]$. Then there are continuous modifications of the processes (V,W) and $(\hat{V},\hat{W})$. Moreover, there is a constant $C = C(\gamma_1,\gamma_2,\alpha,\beta,\eta,d,T,q)$ such that

$$\left\| \sup_{s<t\in I} \frac{\left\|V_{ts} - \hat{V}_{ts}\right\|_{\mathcal{C}^\beta}}{|t-s|^\alpha} \right\|_{L^q} \leq \varepsilon C \kappa \tag{2.40}$$

and

$$\left\| \sup_{s<t\in I} \frac{\left\|W_{ts} - \hat{W}_{ts}\right\|_{\mathcal{C}^\beta}}{|t-s|^{2\alpha}} \right\|_{L^{q/2}} \leq \varepsilon C^2 \kappa^2. \tag{2.41}$$

(ii) *Assume in addition that (V,W) and $(\hat{V},\hat{W})$ satisfy the same moment conditions as in Theorem 2.16 for every $q \geq 2$ and that (2.36), (2.37), (2.38) and (2.39) also hold for every $q \geq 2$ with common constant $\kappa \leq \sqrt{q}\hat{\kappa}$. Then there is a constant $C = C(\gamma_1,\gamma_2,\alpha,\beta,\eta,d,T)$ such that*

$$\varepsilon^{-1} \left\| \sup_{s<t\in I} \frac{\left\|V_{ts} - \hat{V}_{ts}\right\|_{\mathcal{C}^\beta}}{|t-s|^\alpha} \right\|_{L^q} + \varepsilon^{-1/2} \left\| \sqrt{\sup_{s<t\in I} \frac{\left\|W_{ts} - \hat{W}_{ts}\right\|_{\mathcal{C}^\beta}}{|t-s|^{2\alpha}}} \right\|_{L^q} \leq \sqrt{q}\hat{\kappa}C$$

for every $q \geq 2$.

3. Stochastic and rough flows.

The theory of stochastic flows grew out of the pioneering works of the Russian school [**BF61**], [**IA72**] on the dependence of solutions to stochastic differential equations with respect to parameters and the proof by Bismut [**Bis81**] and Kunita [**Kun81**] that stochastic differential equations generate continuous flows of diffeomorphisms under proper regularity conditions on the driving vector fields. The Brownian character of these random flows, that is the fact that they are continuous with stationary and independent increments, was inherited from the Brownian character of their driving noise. The next natural step consisted in the study of Brownian flows for themselves. After the works of Harris [**Har81**], Baxendale [**Bax80**] and Le Jan [**LJ82**], they appeared to be generated by stochastic differential equations driven by infinitely many Brownian motions, or better, to be in one-to-one correspondence with vector field-valued Brownian motions. A probabilistic integration theory of such random time-varying velocity fields was developed to establish that correspondence, and it was extended by Le Jan and Watanabe [**LJW84**] to a large class of continuous semimartingale flows and continuous semimartin-

gale velocity fields. Kunita [**Kun86a**], [**Kun86b**] studied the problem of convergence of stochastic flows, with applications to averaging and homogenization results, and promoted the use of stochastic flows to implement a version of the characteristic method in the setting of first and second order stochastic partial differential equations, notably those coming from the nonlinear filtering theory.

We shall show in this Section that the theory of semimartingale stochastic flows can be embedded into the theory of rough flows developed in Section 2. We review in Section 3.2 the basics of the theory of stochastic flows and show in Section 3.3 that sufficiently regular (semi)martingale velocity fields can be lifted to rough drivers; this is done using our Kolmogorov-type criterion for rough drivers, Theorem 2.16. The identification of (semi)martingale flows generated by (semi)martingale velocity fields to rough flows associated with the corresponding rough driver is done through the Itô formula, on which one can read the local characteristics of a semimartingale flow.

3.1. Notations for function spaces.

The study of stochastic flows classically requires the introduction of a number of function spaces, that we recall here.

Let E and F be Banach spaces. The derivative of a function f from E to F is understood in the Fréchet sense. We shall equip tensor products of Banach spaces with a compatible tensor norm which makes the canonical embedding

$$\mathrm{L}\big(E, L(E, F)\big) \hookrightarrow \mathrm{L}\big(E \otimes E, F\big)$$

continuous. The n-th derivative of f can be seen as a function $D^n f \colon E \to \mathrm{L}(E^{\otimes n}, F)$. For $n \in \mathbb{N}_0$ and $\rho \in (0, 1]$, we define

$$\|f\|_{n+\rho} := \|f\|_{\mathcal{C}^{n+\rho}} := \sum_{i=0}^{n} \sup_{x \in E} \big\|D^i f(x)\big\| + \sup_{0 < \|x-y\| \leq 1} \frac{\big\|D^n f(x) - D^n f(y)\big\|}{\|x - y\|^\rho}.$$

We define $\mathcal{C}_b^{n,\rho}(E, F)$ to be the space of n-times continuously differentiable functions $f \colon E \to F$ such that $\|f\|_{\mathcal{C}^{n+\rho}} < \infty$.

Next, we consider the finite dimensional case. Let D be a domain of $\mathbb{R}^d$, i.e. an open connected subset, and $A \subseteq D$ any subset, $n \in \mathbb{N}_0$ and $\rho \in (0, 1]$. For a function $f \colon D \to \mathbb{R}^k$, set

$$\|f\|_{n+\rho;A} := \sum_{0 \leq |\alpha| \leq n} \sup_{x \in A} \big|\partial^\alpha f(x)\big| + \sum_{|\alpha|=n} \sup_{\substack{x,y \in A \\ 0 < |x-y| \leq 1}} \frac{\big|\partial^\alpha f(x) - \partial^\alpha f(y)\big|}{|x - y|^\rho}.$$

We also set $\|f\|_{n+\rho} := \|f\|_{n+\rho;D}$. Note that this is consistent with the notation above when $D = \mathbb{R}^d$. Let $\mathcal{C}^{n,\rho}(D, \mathbb{R}^k)$ be the space of n-times continuously differentiable functions $f \colon D \to \mathbb{R}^k$ such that $\|f\|_{n+\rho,K} < \infty$ for every compact subset $K \subset D$. Note that although the (semi-)norms we defined here differ slightly from those used by Kunita in his book [**Kun90**], they are actually equivalent on compact sets, hence the spaces coincide. We also define

$$\mathcal{C}_b^{n,\rho}(D, \mathbb{R}^k) := \big\{ f \in \mathcal{C}^{n,\rho}(D, \mathbb{R}^k) \, : \, \|f\|_{n+\rho} < \infty \big\}.$$

For a function $g\colon D \times D \to \mathbb{R}^k$, we similarly define

$$\|g\|^{\wedge}_{n+\rho;A} := \sum_{0 \le |\alpha| \le n} \sup_{x,y \in A} \left|\partial_x^\alpha \partial_y^\alpha g(x,y)\right|$$

$$+ \sum_{|\alpha|=n} \sup_{\substack{x,y,x',y' \in A \\ 0 < |x-x'|,|y-y'| \le 1}} \frac{\left|\partial_x^\alpha \partial_y^\alpha g(x,y) - \partial_x^\alpha \partial_y^\alpha g(x',y) - \partial_x^\alpha \partial_y^\alpha g(x,y') + \partial_x^\alpha \partial_y^\alpha g(x',y')\right|}{|x-x'|^\rho |y-y'|^\rho}.$$

As above, set $\|g\|^{\wedge}_{n+\rho} := \|g\|^{\wedge}_{n+\rho;D}$. We denote by $\widehat{\mathcal{C}}^{n,\rho}(D \times D, \mathbb{R}^k)$ the space of functions $g\colon D \times D \to \mathbb{R}^k$ which are n-times continuously differentiable with respect to each x and y and for which $\|g\|^{\wedge}_{n+\rho;K} < \infty$ for every compact subset $K \subset D$. Set

$$\widehat{\mathcal{C}}^{n,\rho}_b(D \times D, \mathbb{R}^k) := \left\{ g \in \widehat{\mathcal{C}}^{n,\rho}(D \times D, \mathbb{R}^k) \,:\, \|g\|^{\wedge}_{n+\rho} < \infty \right\}.$$

We will sometimes use the shorter notation $\mathcal{C}^{n,\rho}$ and $\mathcal{C}^{n,\rho}_b$, $\widehat{\mathcal{C}}^{n,\rho}$ and $\widehat{\mathcal{C}}^{n,\rho}_b$ respectively when domain and codomain of the function spaces are clear from the context.

3.2. Semimartingale stochastic flows.

We describe in this Section the basics of the theory of semimartingale stochastic flows, and refer the reader to [**LJW84**] or [**Kun90**] for a complete account; we refer to Kunita's book for precise regularity and growth assumptions on the different objects involved. Readers familiar with this material can go directly to Section 3.3.

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0 \le t \le T}, \mathbb{P})$ be a filtered probability space; denote by $\mathfrak{Diff}$, resp. $\mathfrak{F}$, the complete separable metric spaces of $\mathcal{C}^{k_0}$ diffeomorphisms of $\mathbb{R}^d$, resp. $\mathcal{C}^{k_0}$ vector fields on $\mathbb{R}^d$, for some integer $k_0 \ge 2$.

DEFINITION 3.1. A $\mathfrak{Diff}$-valued continuous $(\mathcal{F}_t)_{0 \le t \le T}$-adapted random process $(\phi_t)_{0 \le t \le T}$ is called a $\mathfrak{Diff}$-*valued semimartingale stochastic flow of maps* if the real-valued processes $f(\phi_\bullet(x))$ are real-valued $(\mathcal{F}_t)_{0 \le t \le T}$-semimartingales for all $x \in \mathbb{R}^d$, and all $f \in \mathcal{C}_c^\infty(\mathbb{R}^d)$. Such a $\mathfrak{Diff}$-valued semimartingale is said to be *regular* if for every $x,y \in \mathbb{R}^d$, and $f,g \in \mathcal{C}_c^\infty(\mathbb{R}^d)$, the bounded variation part of $f(\phi_\bullet(x))$ and the bracket $\langle f(\phi_\bullet(x)), g(\phi_\bullet(y)) \rangle$ are absolutely continuous with respect to Lebesgue measure dt.

Their densities $w_t^f(x)$ and $\{f,g\}_t(x,y)$ can be chosen to be jointly measurable and continuous in f,g in the $\mathcal{C}^2$-norm [**LJW84**]. Set

$$\left(\mathcal{L}_t f\right)(x) := w_t^f\left(\phi_t^{-1}(x)\right), \qquad \langle f,g \rangle_t(x,y) := \{f,g\}_t\left(\phi_t^{-1}(x), \phi_t^{-1}(y)\right),$$

so that the processes

$$M_t^f(x) := f\left(\phi_t(x)\right) - f\left(\phi_0(x)\right) - \int_0^t \left(\mathcal{L}_s f\right)\left(\phi_s(x)\right) ds, \qquad x \in \mathbb{R}^d, \ f \in \mathcal{C}_c^\infty(\mathbb{R}^d)$$

are continuous $(\mathcal{F}_t)_{0 \le t \le T}$-local martingales with bracket

$$\left\langle M^f(x), M^g(y) \right\rangle_t = \int_0^t \langle f,g \rangle_t\left(\phi_s(x), \phi_s(y)\right) ds.$$

We have

$$\big(\mathcal{L}_t f\big)(x) = \lim_{h \downarrow 0} \mathbb{E}\left[\frac{f\big(\phi_{t+h,t}(x)\big) - f(x)}{h} \bigg| \mathcal{F}_t \right]$$

and

$$\langle f, g \rangle (x,y) = \lim_{h \downarrow 0} \frac{1}{h} \mathbb{E}\Big[\big\{ f\big(\phi_{t+h,t}(x)\big) - f(x) \big\} \big\{ g\big(\phi_{t+h,t}(y)\big) - g(y) \big\} \Big| \mathcal{F}_t \Big],$$

with limits in L^1 whenever they exist. Under proper regularity conditions [**LJW84**], the operators $\langle f, g \rangle_s(x,y)$ can be seen to be random differential operators of the form

$$\langle f, g \rangle_s(x,y) = A_s^{ij}(x,y)\, \partial_{i,j}^2,$$

for some process $A_s(x,y)$ with values in the space $\mathrm{Symm}(d)$ of symmetric $d \times d$ matrices. The operators $\mathcal{L}_s$ can moreover be expressed in terms of A_s and its differential with respect to the space variables, so that *the data of the processes A_s and $\mathcal{L}_s$ is equivalent to the data of the process A_s and an $\mathfrak{F}$-valued process $b_\bullet$*. The family of random operators $\langle \cdot, \cdot \rangle_t$ and *the drift b_t* are called the *local characteristics* of the $\mathfrak{Diff}$-valued semimartingale $\phi_\bullet$. As an example, for the semimartingale flow generated by a stochastic differential equation of the form

$$dx_t = V_i(x_t) \circ dB_t^i,$$

driven by an ℓ-dimensional Brownian motion, we have

$$\mathcal{L}_t f = \frac{1}{2} \sum_{i=1}^{\ell} V_i^2 f$$

and

$$\langle f, g \rangle_s(x,y) = \big(V_i f\big)(x)\big(V_j g\big)(y),$$

and the drift b_s in the local characteristic is given here by the time-independent vector field

$$b_s(x) = \frac{1}{2}\big(V_i V_i\big)(x) = \frac{1}{2}\big(D_x V_i\big)V_i(x).$$

The infinitesimal counterpart of a $\mathfrak{Diff}$-valued semimartingale is given by the following notion.

DEFINITION 3.2.   A *semimartingale velocity field* is an $\mathfrak{F}$-valued process $(V_t)_{0 \le t \le T}$ such that the processes $(V_\bullet f)(x)$ are real-valued semimartingale for all $x \in \mathbb{R}^d$ and all $f \in \mathcal{C}_c^\infty(\mathbb{R}^d)$. It is called *regular* if one can write

$$V_t = M_t + \int_0^t v_s\, ds$$

for a vector field-valued adapted process $v_\bullet$, and an $\mathfrak{F}$-valued local martingale $M_\bullet$ for which there exists a Symm(d)-valued process $a_s(x, y)$ with

$$\left\langle \partial^\alpha M_\bullet(x), \partial^\beta M_\bullet(y) \right\rangle_t = \int_0^t \partial_x^\alpha \, \partial_y^\beta \, a_s(x, y) \, ds$$

for a range of multiindices α, β depending on the regularity assumptions on a_s. The pair $(a_\bullet, v_\bullet)$ is called the *local characteristics* of the semimartingale velocity field $V_\bullet$.

A theory of Stratonovich integration can be constructed for making sense of integrals of the form

$$\int_0^t V_{\circ ds}(x_s),$$

for some progressively measurable process $x_\bullet$ and some regular semimartingale velocity field $V_\bullet$, as a limit in probability of symmetric Riemann sums. This requires some almost sure regularity properties on the local characteristics $(a_\bullet, v_\bullet)$ of $V_\bullet$, and some almost sure bound on $\int_0^t |a_s(x_s, x_s)| \, ds$ and $\int_0^t |v_s(x_s)| \, ds$ – see e.g. Section 2.3 of [**Kun86b**]. Under these conditions, the integral Stratonovich equation

$$x_t = x_0 + \int_0^t V_{\circ ds}(x_s) \tag{3.1}$$

can be seen to have a unique solution started from any point $x_0 \in \mathbb{R}^d$.

THEOREM 3.3 ([**LJW84**]). *These solutions can be gathered into a semimartingale stochastic flow whose local characteristics are $(a_\bullet, v_\bullet + c_\bullet)$, where the time-dependent vector field c_s has coordinates*

$$c_s^i(x) := \frac{1}{2} \sum_{j=1}^d \partial_{y_j} a_s^{ij}(x, y)\big|_{y=x}$$

in the canonical basis of $\mathbb{R}^d$. Conversely, one can associate to any regular stochastic flow of diffeomorphisms $\phi_\bullet$ a semimartingale velocity field $V_\bullet$, with the same local characteristics as $\phi_\bullet$, and such that $\phi_\bullet$ coincides with the stochastic flow generated by the Stratonovich equation

$$x_t = x_0 + \int_0^t V_{\circ ds}(x_s) - \int_0^t c_s(x_s) \, ds.$$

The optimal regularity assumptions on the velocity fields and stochastic flows of maps are given in Theorems 4.4.1 and 4.5.1 of Kunita's book. We shall use the full strength of these two statements in Section 3.4 to identify semimartingale stochastic flows of maps and the rough flows associated with the lift of the semimartingale velocity fields into a rough driver.

The correspondence between semimartingale stochastic flows and semimartingale velocity fields via an *Itô equation* of the form

$$x_t = x_0 + \int_0^t V_{ds}(x_s)$$

is exact, with no need to add the drift $\int_0^t c_s\,ds$. We state it here under the above form as we shall see below that rough flows are naturally associated with Stratonovich differential equations.

The main difficulty in this business is to deal with the local martingale part of the dynamics, which is where probability theory is really needed. As a consequence, we shall concentrate our efforts on local martingale velocity fields in the sequel, the remaining changes to deal with regular semimartingale velocity fields being essentially cosmetic. As above, we shall freely identify in the sequel vector fields with first order differential operators. In order to keep consistent notations, we shall also denote by

$$\int_0^t \alpha_s\,m_{ds} \qquad \text{and} \qquad \int_0^t \alpha_s\,m_{\circ ds}$$

the Itô and Stratonovich integrals of an adapted process α_s with respect to a local martingale m_s.

3.3. Local martingale rough drivers.

The aim of this Section is to give conditions under which a local martingale velocity field can be lifted into a rough driver. Let D be an open connected subset of $\mathbb{R}^d$ and let M stand for a local martingale velocity field. We prove in this Section that such a field can be lifted to a rough driver $\boldsymbol{M} = (M_{ts}, \mathbb{M}_{ts})_{0 \le s \le t \le T}$, with $M_{ts} := M_t - M_s$, under regularity and boundedness assumptions on the local characteristic of M. At a heuristic level, if M is differentiable in space, the second level operator $\mathbb{M}_{ts}$ associated with $M_{ts} = M_t - M_s$ is given by the formula

$$\mathbb{M}_{ts} = \int_s^t M_{us} M_{\circ du}$$
$$= \left(\int_s^t M_{us}^i \partial_i M_{\circ du}^k \right) \partial_k + \left(\int_s^t M_{us}^j\, M_{\circ du}^k \right) \partial_{jk}^2,$$

with obvious notations for the operators ∂_k and ∂_{jk}^2. In the following, we will use the notation

$$\left(M_{ts}.M_{ts} \right) := \left(DM_{ts} \right)(M_{ts}).$$

As the classical rules of Stratonovich integration give

$$\left(\int_s^t M_{us}^j\, M_{\circ du}^k \right) \partial_{jk}^2 = \frac{1}{2}\, M_{ts}^j M_{ts}^k \partial_{jk}^2 = \frac{1}{2}\, M_{ts} M_{ts} - \frac{1}{2} \left(M_{ts}.M_{ts} \right),$$

we see that the operators $\mathbb{M}_{ts}$ can be decomposed as

$$\mathbb{M}_{ts} = W_{ts} + \frac{1}{2}\, M_{ts} M_{ts}, \tag{3.2}$$

where

$$W_{ts} = \left(\int_s^t M_{us}^i \partial_i M_{\circ du}^k \right) \partial_k - \frac{1}{2} \left(M_{ts}.M_{ts} \right)$$

$$= \frac{1}{2} \left(\int_s^t M_{us}^i \partial_i M_{du}^k - M_{du}^i \partial_i M_{us}^k \right) \partial_k$$

$$= \frac{1}{2} \int_s^t \left(M_{us}.M_{du} - M_{du}.M_{us} \right)$$

is a martingale velocity field defined pointwisely by an Itô integral. The proof that the process $\boldsymbol{M}_{ts} := (M_{ts}, \mathbb{M}_{ts})$ has a modification which is a p-rough driver for every $2 < p < 3$ will require two elementary intermediate results that heavily rest on a now classical modified version of Kolmogorov's regularity criterion that we recall here for the reader's convenience—this is different from the content of the above Section 2.5; they can be found in Section 3 of Kunita's book [**Kun90**].

> THEOREM 3.4.   1. *Let $M_\bullet(x)$, $x \in D$, be a family of continuous local martingales started from 0, such that the joint quadratic variation $\langle M_\bullet(x), M_\bullet(y) \rangle$ has a continuous modification in $\widehat{\mathcal{C}}^{0,\delta}$. Then the process M has a modification that is a continuous process with values in $\mathcal{C}^{0,\epsilon}$, for every $\epsilon < \delta$.*
>
> 2. *Let M be a local martingale velocity field started from 0, with local characteristic the random field $a_t(x,y)$. Assume that has a continuous modification that belongs almost surely to $\widehat{\mathcal{C}}^{m,\delta}$, for some integer $m \geq 1$, and $0 < \delta \leq 1$. Then, for every $0 < \epsilon < \delta$, the velocity field M has a modification that is a continuous process with values in $\mathcal{C}^{m,\epsilon}$; we still denote it by M. Furthermore, for each multi-index α, with $|\alpha| \leq m$, the time varying random field $\partial^\alpha M$ is a local martingale velocity field with quadratic variation*
>
> $$d\langle \partial^\alpha M_\bullet(x), \partial^\alpha M_\bullet(y) \rangle_t = \partial_x^\alpha \partial_y^\alpha a_t(x,y)\, dt.$$
>
> 3. *Let here M and N be two local martingale velocity fields with values in $\mathcal{C}^{m,\delta}$. Then their joint quadratic variation*
>
> $$\langle M_\bullet(x), N_\bullet(y) \rangle_t$$
>
> *has a continuous modification taking values in $\widehat{\mathcal{C}}^{m,\epsilon}$ for every $\epsilon < \delta$. Furthermore, if $m \geq 1$, this modification satisfies the identity*
>
> $$\partial_x^\alpha \partial_y^\beta \langle M_\bullet(x), N_\bullet(y) \rangle_t = \langle \partial^\alpha M_\bullet(x), \partial^\beta N_\bullet(y) \rangle_t,$$
>
> *for all $|\alpha|, |\beta| \leq m$.*

> PROOF.   Cf. Theorem 3.1.1, Theorem 3.1.2 and Theorem 3.1.3 in [**Kun90**].   □

These regularity results will be instrumental in the proof of the following intermediate result. Note that N in equation (3.3) below is seen as a vector field, not a differential

operator, so $(MN)(x) = (D_x N)(M(x))$.

PROPOSITION 3.5. *Let $M, N \colon D \times [0,T] \to \mathbb{R}$ be continuous $\mathcal{C}^{m,\delta}$-valued local mar-tingale fields, for $m \in \mathbb{N}_0$ and $\delta \in (0,1]$. Assume M is adapted to the filtration generated by N. Then the pointwisely defined Itô integral*

$$t \mapsto \int_0^t \left(M_s\, N_{ds} \right)(x) \tag{3.3}$$

has a continuous modification taking values in $\mathcal{C}^{m,\alpha}$-process for every $\alpha < \delta$. Moreover, if $m \geq 1$, the derivative is almost surely given by the formula

$$\partial_i \left(\int_0^t M_s\, N_{ds} \right) = \int_0^t \partial_i M_s\, N_{ds} + \int_0^t M_s\, \partial_i N_{ds}. \tag{3.4}$$

Both assertions also hold for the Stratonovich integral.

The proof of this result is somewhat lengthy but rests on classical considerations based on the regularization theorem, Theorem 3.4.

PROOF.   Set

$$U_t(x) := \left(\int_0^t M_s\, N_{ds} \right)(x).$$

This is a continuous local martingale field with joint quadratic variation given by

$$\left\langle U_\bullet(x), U_\bullet(y) \right\rangle_t = \int_0^t M_s(x)\, M_s(y)\, d\left\langle N_\bullet(x), N_\bullet(y) \right\rangle_s. \tag{3.5}$$

Fix $t \in [0,T]$, some compact set $K \subset D$ and set

$$g(x,y) := \left\langle U_\bullet(x), U_\bullet(y) \right\rangle_t.$$

We first consider the case $m = 0$. Choose $\alpha < \delta' < \delta$. Then, for $x, x', y, y' \in K$, we have

$$\left| g(x,y) - g(x',y) - g(x,y') + g(x',y') \right|$$

$$= \left| \left\langle U_\bullet(x) - U_\bullet(x'), U_\bullet(y) - U_\bullet(y') \right\rangle_t \right|$$

$$\leq \left| \int_0^t \left(M_s(x) - M_s(x') \right)\left(M_s(y) - M_s(y') \right) d\left\langle N_\bullet(x,), N_\bullet(y) \right\rangle_s \right|$$

$$+ \left| \int_0^t \left(M_s(x) - M_s(x') \right) M_s(x')\, d\left\langle N_\bullet(x), N_\bullet(y) - N_\bullet(y') \right\rangle_s \right|$$

$$+ \left| \int_0^t M_s(x') \left(M_s(y) - M_s(y') \right) d\left\langle N_\bullet(x) - N_\bullet(x'), N_\bullet(y) \right\rangle_s \right|$$

$$+ \left| \int_0^t M_s(x') M_s(y')\, d\left\langle N_\bullet(x) - N_\bullet(x'), N_\bullet(y) - N_\bullet(y') \right\rangle_s \right|.$$

For the first integral, we use Kunita's extended Cauchy–Schwarz inequality, as stated in [**Kun90**, Theorem 2.2.13], to see that

$$\left| \int_0^t \left(M_s(x) - M_s(x') \right) \left(M_s(y) - M_s(y') \right) d\langle N_\bullet(x), N_\bullet(y) \rangle_s \right|$$

$$\leq \left(\int_0^t \left(M_s(x) - M_s(x') \right)^2 d\langle N_\bullet(x) \rangle_s \right)^{1/2} \left(\int_0^t \left(M_s(y) - M_s(y') \right)^2 d\langle N_\bullet(y) \rangle_s \right)^{1/2}$$

$$\leq |x - x'|^\delta |y - y'|^\delta \sup_{s \in [0,T]} \|M_s\|_{\delta-\text{Höl};K}^2 \sup_{z \in K} \langle N_\bullet(z) \rangle_T.$$

Similarly, for the second integral,

$$\left| \int_0^t \left(M_s(x) - M_s(x') \right) M_s(x') \, d\langle N_\bullet(x), N_\bullet(y) - N_\bullet(y') \rangle_s \right|$$

$$\leq \left(\int_0^t \left(M_s(x) - M_s(x') \right)^2 d\langle N_\bullet(x) \rangle_s \right)^{1/2} \left(\int_0^t M_s(x')^2 \, d\langle N_\bullet(y) - N_\bullet(y') \rangle_s \right)^{1/2}$$

$$\leq |x-x'|^\delta |y-y'|^{\delta'} \sup_{s \in [0,T]} \left\| M_s \right\|_{\infty;K} \sup_{s \in [0,T]} \left\| M_s \right\|_{\delta-\text{Höl};K} \sup_{z \in K} \langle N_\bullet(z) \rangle_T^{1/2} \left\| \langle N_\bullet, N_\bullet \rangle_T \right\|_{\delta';K}^{1/2}.$$

From point 3 in Theorem 3.4 we know that there is a version of the joint quadratic variation of N such that $\|\langle N_\bullet, N_\bullet \rangle_T\|_{\delta';K}^\wedge < \infty$. The other integrals are estimated similarly. This shows that

$$\sup_{\substack{x,x',y,y' \in K \\ x \neq x', y \neq y'}} \frac{\left| g(x,y) - g(x',y) - g(x,y') + g(x',y') \right|}{|x - x'|^{\delta'} |y - y'|^{\delta'}} < \infty.$$

Clearly $\|g\|_{\infty;K} < \infty$, thus we have shown that the joint quadratic variation of U has a modification which is a continuous $\widehat{\mathcal{C}}^{0,\delta'}$-process. Point 1 in Theorem 3.4 shows that U has a modification which is a continuous $\mathcal{C}^{0,\alpha}$-process. Now let $m \geq 1$. From point 3 in Theorem 3.4 we may deduce that the joint quadratic variation of N has a modification which is a continuous $\tilde{\mathcal{C}}^{m,\delta'}$-process with

$$\partial_x^\beta \partial_y^\gamma \langle N_\bullet(x), N_\bullet(y) \rangle = \langle \partial^\beta N_\bullet(x), \partial^\gamma N_\bullet(y) \rangle$$

for every $|\beta|, |\gamma| \leq m$. We may apply Proposition 5.1 in Appendix iteratively in equation (3.5) to show that $\langle U_\bullet(x), U_\bullet(y) \rangle_t$ has a modification which is m-times differentiable with respect to x and y, and we can calculate the derivatives using the product rule stated in Proposition 5.1. As above, one can show that the m-th derivative has the claimed Hölder regularity, and we can conclude with point 2 of Theorem 3.4 that U has a modification which is a continuous $\mathcal{C}^{m,\alpha}$-process. The Itô-Stratonovich conversion formula

$$\int_0^t M_s \, N_{\circ ds} = \int_0^t M_s \, N_{ds} + \frac{1}{2} \langle M_\bullet, N_\bullet \rangle_t$$

and point 3 in Theorem 3.4 show that the same is true for the Stratonovich integral.

We now come to equation (3.4). In the following, we use $\| \cdot \|_{L^1}$ for the L^1-norm

with respect to $\mathbb{P}$. For $n \in \mathbb{N}$, set

$$\tau_n = \inf\left\{t \in [0,T] \ : \ \left\|M_t\right\|_{\mathcal{C}^{1,\delta}} + \left\|N_t\right\|_{\mathcal{C}^{1,\delta}} \geq n\right\}.$$

The random times τ_n define an increasing sequence of stopping times such that $\mathbb{P}(\tau_n < T) \to 0$ a $n \to 0$. Let $x \in D$, $t > 0$ and choose h such that $x + he_i \in D$. Then

$$\left\|\frac{1}{h}\left\{\int_0^{t \wedge \tau_n} \left(M_s\, N_{ds}\right)(x + he_i) - \int_0^{t \wedge \tau_n} \left(M_s N_{ds}\right)(x)\right\}\right.$$
$$\left. - \int_0^{t \wedge \tau_n} \left(\partial_i M_s\, N_{ds}\right)(x) - \int_0^{t \wedge \tau_n} \left(M_s\, \partial_i N_{ds}\right)(x)\right\|_{L^1}$$
$$= \left\|\int_0^{t \wedge \tau_n} \frac{M_s(x + he_i) - M_s(x)}{h}\, N_{ds}(x + he_i) + \int_0^{t \wedge \tau_n} M_s(x)\, \frac{N_{ds}(x + he_i) - N_{ds}(x)}{h}\right.$$
$$\left. - \int_0^{t \wedge \tau_n} \partial_i M_s(x)\, N_{ds}(x) + \int_0^{t \wedge \tau_n} M_s(x)\, \partial_i N_{ds}(x)\right\|_{L^1}$$
$$\leq \left\|\int_0^{t \wedge \tau_n} \left(\frac{M_s(x + he_i) - M_s(x)}{h}\right) \left(N_{ds}(x + he_i) - N_{ds}(x)\right)\right\|_{L^1}$$
$$+ \left\|\int_0^{t \wedge \tau_n} \left(\frac{M_s(x + he_i) - M_s(x)}{h} - \partial_i M_s(x)\right) N_{ds}(x)\right\|_{L^1}$$
$$+ \left\|\int_0^{t \wedge \tau_n} M_s(x) \left(\frac{N_{ds}(x + he_i) - N_{ds}(x)}{h} - \partial_i N_{ds}(x)\right)\right\|_{L^1}.$$

We aim to show that the integrals on the right hand side vanish as $h \to 0$, using the Burkholder–Davis–Gundy inequality. We start with the first integral. Note that since N is a $\mathcal{C}^{1,\delta}$-process, also the stopped process is a $\mathcal{C}^{1,\delta}$-process and its joint quadratic variation has a modification which is a $\tilde{\mathcal{C}}^{1,\delta'}$-process for any $\delta' < \delta$. In particular,

$$\left\langle N_\bullet(x + he_i) - N_\bullet(x)\right\rangle_{t \wedge \tau_n} \to 0$$

almost surely as $h \to 0$. From the Burkholder–Davis–Gundy inequality,

$$\mathbb{E}\left|\left\langle N_\bullet(x + he_i) - N_\bullet(x)\right\rangle_{t \wedge \tau_n}\right|^{p/2} \leq C_p\, \mathbb{E}\left|\sup_{s \in [0, t \wedge \tau_n]} \left|N_s(x + he_i) - N_s(x)\right|\right|^p \leq C\, 2^p n^p$$

which shows that $\left\langle N_\bullet(x + he_i) - N_\bullet(x)\right\rangle_{t \wedge \tau_n} \to 0$ in L^p for any $p \geq 1$. Using the Burkholder–Davis–Gundy inequality for the first integral gives

$$\mathbb{E}\left|\int_0^{t \wedge \tau_n} \left(\frac{M_s(x + he_i) - M_s(x)}{h}\right) \left(N_{ds}(x + he_i) - N_{ds}(x)\right)\right|$$
$$\leq C\, \mathbb{E}\left|\int_0^{t \wedge \tau_n} \left(\frac{M_s(x + he_i) - M_s(x)}{h}\right)^2 d\langle N_\bullet(x + he_i) - N_\bullet(x)\rangle_s\right|^{1/2}$$
$$\leq C\, \mathbb{E}\left|\sup_{s \in [0, t \wedge \tau_n]} \left\|\partial_i M_s\right\|_\infty^2 \langle N_\bullet(x + he_i) - N_\bullet(x)\rangle_{t \wedge \tau_n}\right|^{1/2}$$

$$\le Cn\,\mathbb{E}\left|\left\langle N_\bullet(x+he_i)-N_\bullet(x)\right\rangle_{t\wedge\tau_n}\right|^{1/2}\to 0$$

as $h\to 0$. For the second integral, the Burkholder–Davis–Gundy inequality gives

$$\mathbb{E}\left|\int_0^{t\wedge\tau_n}\left(\frac{M_s(x+he_i)-M_s(x)}{h}-\partial_i M_s(x)\right)\,N_{ds}(x)\right|$$

$$\le C\,\mathbb{E}\left|\int_0^{t\wedge\tau_n}\left(\frac{M_s(x+he_i)-M_s(x)}{h}-\partial_i M_s(x)\right)^2\,d\langle N_\bullet(x)\rangle_s\right|^{1/2}$$

For $h\to 0$, we can use dominated convergence twice for the expectation and the Lebesgue–Stieltjes integral to see that this term indeed converges to 0. Now we come to the third integral. Set

$$N_t^h(x):=\frac{N_t(x+he_i)-N_t(x)}{h}.$$

As before, we can use the Burkholder–Davis–Gundy inequality to see that

$$\mathbb{E}\left|\int_0^{t\wedge\tau_n}M_s(x)\left(N_{ds}^h(x)-\partial_i N_{ds}(x)\right)\right|\le Cn\mathbb{E}\left|\left\langle N_\bullet^h(x)-\partial_i N_\bullet(x)\right\rangle_{t\wedge\tau_n}\right|^{1/2}$$

Taking a modification which is a $\tilde{\mathcal{C}}^{1,\delta'}$-process of the joint quadratic variation of N gives, using point 3 of Theorem 3.4, that

$$\left\langle N_\bullet^h(x)-\partial_i N_\bullet(x)\right\rangle_{t\wedge\tau_n}\to 0$$

almost surely as $h\to 0$. Using the Burkholder–Davis–Gundy inequality as above shows that the convergence also holds in L^p for any $p\ge 1$. To conclude, we have shown that

$$\frac{1}{h}\left(\int_0^{t\wedge\tau_n}M_s(x+he_i)\,N_{ds}(x+he_i)-\int_0^{t\wedge\tau_n}M_s(x)\,N_{ds}(x)\right)$$

$$\to\int_0^{t\wedge\tau_n}\partial_i M_s(x)\,N_{ds}(x)+\int_0^{t\wedge\tau_n}M_s(x)\,\partial_i N_{ds}(x)$$

in L^1 as $h\to 0$. Since we already know that the convergence holds almost surely, the limits coincide and we have shown that

$$\partial_i\left(\int_0^{t\wedge\tau_n}M_s(x)\,N_{ds}(x)\right)=\int_0^{t\wedge\tau_n}\partial_i M_s(x)\,N_{ds}(x)+\int_0^{t\wedge\tau_n}M_s(x)\,\partial_i N_{ds}(x)\qquad(3.6)$$

holds almost surely for every $n\in\mathbb{N}$. Now,

$$\mathbb{P}\left(\partial_i\left(\int_0^t M_s(x)\,N_{ds}(x)\right)\ne\partial_i\left(\int_0^{t\wedge\tau_n}M_s(x)\,N_{ds}(x)\right)\right)\le\mathbb{P}(\tau_n<T)\to 0$$

as $n\to\infty$, therefore

$$\partial_i \left(\int_0^{t \wedge \tau_n} M_s(x) \, N_{ds}(x) \right) \to \partial_i \left(\int_0^t M_s(x) \, N_{ds}(x) \right)$$

in probability as $n \to \infty$. The same is true for the other two integrals on the right hand side of (3.6) which shows (3.4) for the Itô integral. For the Stratonovich integral, the assertion follows from the Itô–Stratonovich conversion formula and the equality

$$\partial_i \big\langle M_\bullet(x), N_\bullet(x) \big\rangle_t = \big\langle \partial_i M_\bullet(x), N_\bullet(x) \big\rangle_t + \big\langle M_\bullet(x), \partial_i N_\bullet(x) \big\rangle_t. \qquad \square$$

Recall the definition of the space $\hat{\mathcal{C}}_b^{n,\delta}$.

PROPOSITION 3.6. *Let $M, N \colon D \times [0,T] \to \mathbb{R}$ be continuous local martingale fields adapted to the same filtration. Assume that the quadratic variation of the processes is given by*

$$d\langle M_\bullet(x), M_\bullet(y) \rangle_t = a_t(x,y) \, dt \quad resp.$$
$$d\langle N_\bullet(x), N_\bullet(y) \rangle_t = b_t(x,y) \, dt$$

for every $x, y \in D$ and every $t \in [0,T]$. Moreover, assume that there is a $\delta \in (0,1]$ such that a and b have continuous modifications in the space $\hat{\mathcal{C}}_b^{0,\delta}$. Let $p > 2$ and $\rho \in (0,\delta)$.

(i) *Assume that there are $\kappa > 0$ and some $\eta \in (0,\infty)$ such that*

$$\sup_{u \in [0,T]} \left\| \sqrt{a_u(x,x)} \right\|_{L^q(\Omega)} \leq \frac{\kappa}{1 + |x|^\eta} \tag{3.7}$$

holds for every $x \in D$ where

$$q > \max \left\{ \frac{\delta}{\eta(\delta - \rho)} + d \left(\frac{\delta}{\eta(\delta - \rho)} + \frac{1}{\delta - \rho} \right), \frac{1}{1/2 - 1/p} \right\}. \tag{3.8}$$

Then the process M has a modification which satisfies

$$\sup_{0 \leq s < t \leq T} \frac{\|M_{ts}\|_{\mathcal{C}^\rho}}{|t-s|^{1/p}} < \infty$$

almost surely.

(ii) *Set*

$$U_{ts}(x) := \int_s^t (M_{us} N_{du})(x).$$

Assume that there are a constant $\kappa > 0$ and some $\eta \in (0,\infty)$ such that

$$\sup_{u \in [0,T]} \left\| \sqrt{a_u(x,x)} \right\|_{L^q(\Omega)} + \sup_{u \in [0,T]} \left\| \sqrt{b_u(x,x)} \right\|_{L^q(\Omega)} \leq \frac{\kappa}{1 + |x|^\eta} \tag{3.9}$$

for every $x \in D$ with q sufficiently large as in (3.8).

Then the process U has a modification which satisfies

$$\sup_{0 \le s < t \le T} \frac{\|U_{ts}\|_{\mathcal{C}^\rho}}{|t-s|^{2/p}} < \infty$$

almost surely.

(iii) *Assume that* $\sup_{t \in [0,T]} \|a_t\|_\delta^\wedge$ *and* $\sup_{t \in [0,T]} \|b_t\|_\delta^\wedge$ *are almost surely bounded random variables and that (3.9) holds for some* $\kappa > 0$ *and* $\eta \in (0, \infty)$ *uniformly for every* $q \ge 1$.

Then the random variables

$$\sup_{0 \le s < t \le T} \frac{\|M_{ts}\|_{\mathcal{C}^\rho}}{|t-s|^{1/p}} \qquad and \qquad \sqrt{\sup_{0 \le s < t \le T} \frac{\|U_{ts}\|_{\mathcal{C}^\rho}}{|t-s|^{2/p}}}$$

have Gaussian tails.

PROOF. We start with (i). In a first step, we assume that $\sup_{t \in [0,T]} \|a_t\|_\delta^\wedge$ is an almost surely bounded random variable and that (3.7) holds uniformly for every $q \ge 1$. Under these assumptions, we aim to show that M has a modification such that

$$\sup_{0 \le s < t \le T} \frac{\|M_{ts}\|_{\mathcal{C}^\rho}}{|t-s|^{1/p}} \tag{3.10}$$

is finite almost surely and has Gaussian tails. Set

$$\alpha_1 := \left\| \sup_{u \in [0,T]} \|a_u\|_\delta^\wedge \right\|_{L^\infty}. \tag{3.11}$$

Let $s < t$ and $x, y \in D$ with $0 < |x-y| \le 1$. By the Burkholder–Davis–Gundy inequality with optimal constants (cf. [**CK91**, Theorem A]), for every $q \ge 2$,

$$\mathbb{E}|M_{ts}(x) - M_{ts}(y)|^q \le C^q q^{q/2} \mathbb{E} \left| \int_s^t a_u(x,x) - a_u(x,y) - a_u(y,x) + a_u(y,y) \, du \right|^{q/2}$$

$$\le C^q q^{q/2} |t-s|^{q/2} \alpha_1^{q/2} |x-y|^{q\delta}$$

where we used the estimate

$$\sup_{u \in [0,T]} |a_u(x,x) - a_u(x,y) - a_u(y,x) + a_u(y,y)| \le \alpha_1 |x-y|^{2\delta}$$

in the last step. Similarly, for $s < t$ and $x \in D$,

$$\mathbb{E}|M_{ts}(x)|^q \le C q^{q/2} \mathbb{E} \left| \int_s^t a_u(x,x) \, du \right|^{q/2} \le C q^{q/2} |t-s|^{q/2} \left(\frac{\kappa}{1+|x|^\eta} \right)^q$$

for every $q \ge 2$. Hence we can conclude with Theorem 2.16 (iii) that M has a modification such that (3.10) is finite and has Gaussian tails.

Now we drop the assumption that $\alpha_1 < \infty$. Consider the stopped process $M_t^n := M_{t \wedge \tau_n}$ where

$$\tau_n = \inf \left\{ t \in [0,T] \ : \ \sup_{u \in [0,t]} \|a_u\|_\delta^\wedge \geq n \right\},$$

hence

$$\langle M_\bullet^n(x), M_\bullet^n(y) \rangle_t = \int_0^{t \wedge \tau_n} a_u(x,y) \, du.$$

Fix $n \in \mathbb{N}$. As before, we have the estimates

$$\|M_{ts}^n(x) - M_{ts}^n(y)\|_{L^q} \leq C\sqrt{n}\sqrt{q}|t-s|^{1/2}|x-y|^\delta$$

for every $x, y \in D$, $s < t$ and $q \geq 1$, and

$$\|M_{ts}^n(x)\|_{L^q} \leq C\sqrt{q}|t-s|^{1/2} \left(\frac{\kappa}{1+|x|^\eta} \right)$$

for every $x \in D$, $s < t$ and q as in (3.8). Thus Theorem 2.16 implies that there is a modification of M^n such that

$$\sup_{0 \leq s < t \leq T} \frac{\|M_{ts}^n\|_{\mathcal{C}^\rho}}{|t-s|^{1/p}}$$

is finite almost surely for every $n \in \mathbb{N}$. Now,

$$\mathbb{P}\left(\sup_{0 \leq s < t \leq T} \frac{\|M_{ts}\|_{\mathcal{C}^\rho}}{|t-s|^{1/p}} = \infty \right) \leq \mathbb{P}\left(\sup_{0 \leq s < t \leq T} \frac{\|M_{ts}\|_{\mathcal{C}^\rho}}{|t-s|^{1/p}} \neq \sup_{0 \leq s < t \leq T} \frac{\|M_{ts}^n\|_{\mathcal{C}^\rho}}{|t-s|^{1/p}} \right)$$

$$\leq \mathbb{P}(\tau_n < T) \to 0$$

as $n \to \infty$, which shows that indeed

$$\sup_{0 \leq s < t \leq T} \frac{\|M_{ts}\|_{\mathcal{C}^\rho}}{|t-s|^{1/p}} < \infty$$

almost surely. This shows (i) and the first part of (iii).

We proceed with (ii). As above, we first assume that

$$\alpha_2 := \left\| \sup_{u \in [0,T]} \|a_u\|_\delta^\wedge \right\|_{L^\infty} + \left\| \sup_{u \in [0,T]} \|b_u\|_\delta^\wedge \right\|_{L^\infty} < \infty$$

and that (3.9) holds uniformly for every $q \geq 1$. Let $x, y \in D$ such that $0 < |x-y| \leq 1$ and $s < t$. By the triangle inequality, for every $q \geq 2$,

$$\left\| U_{ts}(x) - U_{ts}(y) \right\|_{L^{q/2}} \leq \left\| \int_s^t (M_{us}(x) - M_{us}(y)) N_{du}(x) \right\|_{L^{q/2}}$$

$$+ \left\| \int_s^t M_{us}(y)(N_{du}(x) - N_{du}(y)) \right\|_{L^{q/2}}.$$

Note that

$$\left\langle \int_0^\bullet (M_u(x) - M_u(y)) N_{du}(x) \right\rangle_t = \int_0^t |M_u(x) - M_u(y)|^2 \langle N_\bullet(x) \rangle_{du}$$

$$\leq \sup_{u \in [0,t]} |M_u(x) - M_u(y)|^2 \langle N_\bullet(x) \rangle_t.$$

Therefore, applying twice the Burkholder–Davis–Gundy inequality,

$$\left\| \int_s^t (M_{us}(x) - M_{us}(y)) N_{du}(x) \right\|_{L^{q/2}} \leq C\sqrt{q} \left\| \sqrt{\langle N_\bullet(x) \rangle_{ts}} \sup_{u \in [s,t]} |M_{us}(x) - M_{us}(y)| \right\|_{L^{q/2}}$$

$$\leq C\sqrt{q} \| \sqrt{\langle N_\bullet(x) \rangle_{ts}} \|_{L^\infty} \left\| \sup_{u \in [s,t]} |M_{us}(x) - M_{us}(y)| \right\|_{L^{q/2}}$$

$$\leq Cq \| \sqrt{\langle N_\bullet(x) \rangle_{ts}} \|_{L^\infty} \| \sqrt{\langle M_\bullet(x) - M_\bullet(y) \rangle_{ts}} \|_{L^{q/2}}.$$

Now we have the estimate

$$\| \sqrt{\langle N_\bullet(x) \rangle_{ts}} \|_{L^\infty} = \left\| \sqrt{\int_s^t b_u(x,x) \, du} \right\|_{L^\infty} \leq \sqrt{\alpha_2} |t - s|^{1/2}.$$

Furthermore, as seen above,

$$\| \sqrt{\langle M_\bullet(x) - M_\bullet(y) \rangle_{ts}} \|_{L^{q/2}} \leq \sqrt{\alpha_2} |t - s|^{1/2} |x - y|^\delta$$

which implies

$$\left\| \int_s^t (M_{us}(x) - M_{us}(y)) N_{du}(x) \right\|_{L^{q/2}} \leq C\alpha_2 q |t - s| |x - y|^\delta.$$

Similarly,

$$\left\| \int_s^t M_{us}(y)(N_{du}(x) - N_{du}(y)) \right\|_{L^{q/2}} \leq C\alpha_2 q |t - s| |x - y|^\delta,$$

hence we have shown that

$$\| U_{ts}(x) - U_{ts}(y) \|_{L^{q/2}} \leq C\alpha_2 q |t - s| |x - y|^\delta$$

holds for every $q \geq 2$, every $s < t$ and every $x, y \in D$ such that $0 < |x - y| \leq 1$.

Similarly, for every $q \geq 2$, every $x \in D$ and every $s < t$,

$$\| U_{ts}(x) \|_{L^{q/2}} \leq Cq \| \sqrt{\langle M_\bullet(x) \rangle_{ts}} \|_{L^q} \| \sqrt{\langle N_\bullet(x) \rangle_{ts}} \|_{L^q}$$

$$\leq \frac{Cq|t - s|}{1 + |x|^{2\eta}}.$$

Hence we can apply Theorem 2.16 to show that the process U has a modification such that

$$\sqrt{\sup_{0\leq s<t\leq T}\frac{\|U_{ts}\|_{\mathcal{C}^\rho}}{|t-s|^{2/p}}}$$

is finite almost surely and has Gaussian tails. It remains to prove (ii) when $\alpha_2 = \infty$. In this case, we can consider the stopped processes $M_t^n := M_{t\wedge\sigma_n}$ and $N_t^n := N_{t\wedge\sigma_n}$ where

$$\sigma_n = \inf\left\{ t \in [0,T] \, : \, \sup_{u\in[0,t]}\|a_u\|_\delta^\wedge + \sup_{u\in[0,t]}\|b_u\|_\delta^\wedge \geq n \right\}$$

and proceed as above. We leave the details to the reader. $\square$

THEOREM 3.7. *Let M be a continuous local martingale velocity field in $\mathcal{C}^{2,\delta}(D,\mathbb{R}^d)$ with continuous local characteristic a in $\hat{\mathcal{C}}_b^{2,\delta}$ for some $\delta \in (0,1]$.*

(i) *Let $\rho \in (0,\delta)$ and $p \in (2,3)$. Assume that there aren $\eta \in (0,\infty)$ and a constant $\kappa > 0$ such that*

$$\sum_{0\leq|\alpha|\leq 2}\sup_{u\in[0,T]}\left\|\sqrt{\partial_x^\alpha\partial_y^\alpha a_u(x,y)|_{x=y=z}}\right\|_{L^q} \leq \frac{\kappa}{1+|z|^\eta} \qquad (3.12)$$

for every $z \in D$ where q satisfies

$$q > \max\left\{ \frac{\delta}{\eta(\delta-\rho)} + d\left(\frac{\delta}{\eta(\delta-\rho)} + \frac{1}{\delta-\rho} \right), \frac{1}{1/2-1/p} \right\}.$$

Then $\boldsymbol{M} = (M,\mathbb{M})$, $\mathbb{M}$ being defined as in (3.2), has a modification which is a weak geometric (p,ρ)-rough driver. We call $\boldsymbol{M}$ the natural lift *of M.*

(ii) *Assume that $\sup_{t\in[0,T]}\|a_t\|_{2+\delta}^\wedge$ is an almost surely bounded random variable and that there are an $\eta \in (0,\infty)$ and a constant $\kappa > 0$ such that (3.12) holds uniformly for every $z \in D$ and every $q \geq 1$.*

Then for every $p \in (2,3)$ and $\rho \in (0,\delta)$, $\boldsymbol{M} = (M,\mathbb{M})$ has a modification which is a weak geometric (p,ρ)-rough driver, and the random variables

$$\sup_{0\leq s<t\leq T}\frac{\|M_{ts}\|_{\mathcal{C}^{2+\rho}}}{|t-s|^{1/p}} \qquad and \qquad \sqrt{\sup_{0\leq s<t\leq T}\frac{\|W_{ts}\|_{\mathcal{C}^{1+\rho}}}{|t-s|^{2/p}}},$$

W being defined as in (3.2), have Gaussian tails.

PROOF. The claim for M follows by applying Proposition 3.6 to M and its derivatives. For W, we use the product rule in Proposition 3.5 for calculating the derivatives and apply Proposition 3.6 afterwards. The estimates for W together with M yield the claimed estimates for $\mathbb{M}$. We leave the details to the reader. $\square$

REMARK 3.8. In the special case where M or a have compact support, i.e. when there exists a deterministic compact set $K \subset D$ such that M resp. a are supported on K almost surely, assertion (i) of Theorem 3.7 holds without any moment conditions on a. Indeed, this follows from the fact that the stopped processes in Proposition 3.6 trivially

satisfy the growth condition stated in the Kolmogorov theorem 2.16, and this was the only point where these assumptions were needed.

3.4. Stochastic and rough flows.

We keep in this Section the notations of the previous Sections, and denote in particular by $(\mathcal{F}_t)_{0 \le t \le T}$ a filtration to which the semimartingale velocity field M is adapted. Assume that the local characteristic a of M satisfies the boundedness assumptions of point (i) in Theorem 3.7. Then we can use Theorem 3.7 to define the natural lift $\boldsymbol{M}$ of M into a rough driver, and one can make sense of the rough flow φ as pathwise solution to the equation

$$d\varphi = \boldsymbol{M}(\varphi \, ; dt)$$

using Theorem 2.8. It follows from equation (2.3), giving φ_{ts} as a limit of compositions of μ_{ba}'s, that φ is a semimartingale stochastic flow of homeomorphisms. One can read its local characteristics on the Itô formula that it satisfies. Given $x, y \in \mathbb{R}^d$ and $f, g \in \mathcal{C}_b^3$, we have

$$f\big(\varphi_{ts}(x)\big) = f(x) + \big(M_{ts}f\big)(x)$$
$$+ \frac{1}{2}\left\{\left(\int_s^t M_{us}.M_{du} - M_{du}.M_{us}\right)f\right\}(x) + \frac{1}{2}\big(M_{ts}^2 f\big)(x) + O\big(|t-s|^{3/p}\big),$$

with an $O(\cdot)$ term depending only on $\|\boldsymbol{M}\|$ and $\|f\|_{\mathcal{C}^3}$, with a similar formula for $g(\varphi_{ts}(y))$. We read on this identity that

$$\lim_{h \downarrow 0} \mathbb{E}\left[\frac{f\big(\varphi_{t+h,t}(x)\big) - \varphi(x)}{h}\,\Big|\,\mathcal{F}_t\right] \overset{L^1}{=} \lim_{h \downarrow 0} \mathbb{E}\left[\frac{1}{2}\big(M_{t+h,t}^2 f\big)(x)\,\Big|\,\mathcal{F}_t\right],$$

and

$$\lim_{h \downarrow 0} \frac{1}{h}\mathbb{E}\Big[\big\{f\big(\varphi_{t+h,t}(x)\big) - f(x)\big\}\big\{g\big(\varphi_{t+h,t}(y)\big) - g(y)\big\}\,\Big|\,\mathcal{F}_t\Big]$$
$$= \lim_{h \downarrow 0} \frac{1}{h}\mathbb{E}\Big[\big(M_{t+h,t}f\big)(x)\big(M_{t+h,t}g\big)(y)\,\Big|\,\mathcal{F}_t\Big] = \langle f, g\rangle_t(x, y).$$

So the semimartingale stochastic flow φ has the same local characteristics as the semimartingale stochastic flow generated by the Stratonovich differential equation

$$dx_t = M_{\circ dt}(x_t); \tag{3.13}$$

they coincide by Theorem 3.3, such as stated in Theorems 4.4.1 and 4.5.1 in Kunita's book, as assumption (3.12) on the local characteristic a of M is clearly stronger than the optimal assumptions of Kunita.

THEOREM 3.9. *Let M be a continuous local martingale velocity field in $\mathcal{C}^{2,\delta}(\mathbb{R}^d, \mathbb{R}^d)$, for some $\delta \in (2/3, 1]$, with continuous local characteristic a in $\widehat{\mathcal{C}}_b^{2,\delta}$. Let $\boldsymbol{M}$ be the rough driver associated with M by Theorem 3.7. Under the condition that M or a have compact support or that the growth condition (3.12) holds, the rough flow solution*

to the differential equation

$$d\varphi = \boldsymbol{M}(\varphi\,;dt)$$

coincides with the stochastic flow generated by the Stratonovich differential equation (3.13).

3.5. Strong approximations.

We give in this Section an example of use of the continuity of the Itô map, in the setting of rough drivers and rough flows, by proving a Wong–Zakaï type theorem for semimartingale stochastic flows of maps. That is, we prove that such flows are limits in probability of flows generated by ordinary differential equations. Granted the continuity of the Itô map, the core of the proof consists in showing that a rough lift of a continuous piecewise linear time interpolation of a semimartingale velocity field M converges in probability to $\boldsymbol{M}$ in the topology of rough drivers.

As in the last Section, let

$$M \colon [0,T] \to \mathcal{C}^{2,\delta}\big(D,\mathbb{R}^d\big)$$

be a continuous local martingale velocity field with quadratic variation

$$\big\langle M_\bullet^i(x), M_\bullet^j(y) \big\rangle_t = \int_0^t a_s^{ij}(x,y)\,ds$$

and $\delta \in (0,1]$. Let $\mathcal{D} = \{0 = t_0 < t_1 < \ldots < t_n = T\}$ be a partition of the interval $[0,T]$ and define the piecewise linear approximation of M with respect to $\mathcal{D}$ as

$$M_t^{\mathcal{D}} := M_{t_i} + (t - t_i)\,\frac{M_{t_{i+1}} - M_{t_i}}{t_{i+1} - t_i} \quad \text{if } t \in [t_i, t_{i+1}].$$

Note that $\mathcal{D} \mapsto M^{\mathcal{D}}$ commutes with the spatial derivate, i.e.

$$\partial_i\big(M^{\mathcal{D}}\big) = \big(\partial_i M\big)^{\mathcal{D}} =: \partial_i M^{\mathcal{D}}.$$

Define the mesh size of the partition by the formula $|\mathcal{D}| := \max_i |t_{i+1} - t_i|$. We define the first order differential operator

$$W_{ts}^{\mathcal{D}} := \frac{1}{2}\left(\int_s^t M_{us}^{\mathcal{D};i}\partial_i M_{du}^{\mathcal{D};k} - \int_s^t M_{du}^{\mathcal{D};i}\partial_i M_{us}^{\mathcal{D};k}\right)\partial_k$$

$$= \frac{1}{2}\int_s^t (M_{us}^{\mathcal{D}}.M_{du}^{\mathcal{D}} - M_{du}^{\mathcal{D}}.M_{us}^{\mathcal{D}})$$

via usual Riemann–Stieltjes integration. Then we set

$$\mathbb{M}_{ts}^{\mathcal{D}} := W_{ts}^{\mathcal{D}} + \frac{1}{2}\,M_{ts}^{\mathcal{D}}M_{ts}^{\mathcal{D}}$$

and

$$\boldsymbol{M}^{\mathcal{D}} := \big(M^{\mathcal{D}}, \mathbb{M}^{\mathcal{D}}\big). \tag{3.14}$$

Our aim is to prove that $\boldsymbol{M}^{\mathcal{D}}$ converges towards the natural lift $\boldsymbol{M}$ of M when $|\mathcal{D}| \to 0$ in probability (or even in $L^p(\Omega)$) in the topology of rough drivers. Note that

$$W_{ts} = \frac{1}{2} \int_s^t (M_{us}.M_{du} - M_{du}.M_{us}) = \frac{1}{2} \int_s^t (M_{us}.M_{\circ du} - M_{\circ du}.M_{us}),$$

hence it is enough to prove that the Riemann–Stieltjes integrals of the approximated processes converge towards the Stratonovich integrals (in the right topology), and this is what we are going to do.

LEMMA 3.10. *Let* $M = (M^1, \ldots, M^d) \colon [0, T] \to \mathbb{R}^d$ *be a continuous local martingale and assume that*

$$\big\| \langle M \rangle_T \big\|_{L^{q/2}} \le K < \infty$$

for some $q \ge 2$ *and* $K > 0$. *Let* $\boldsymbol{M}$ *and* $\boldsymbol{M}^{\mathcal{D}}$ *be the associated rough paths lifts to* M *and* $M^{\mathcal{D}}$, *i.e.* $\boldsymbol{M} = (M, \mathbb{M})$ *and* $\boldsymbol{M}^{\mathcal{D}} = (M^{\mathcal{D}}, \mathbb{M}^{\mathcal{D}})$ *where*

$$\mathbb{M} = \int_s^t M_{us} M_{\circ du} \quad and \quad \mathbb{M}^{\mathcal{D}} = \int_s^t M_{us}^{\mathcal{D}} M_{du}^{\mathcal{D}}$$

are iterated Stratonovich, resp. Riemann–Stieltjes, integrals. Set

$$\varepsilon := \left\| \sup_{0 < v - u \le |\mathcal{D}|} \| \boldsymbol{M}_{vu} \| \right\|_{L^q}$$

and assume that $\varepsilon \le 1$.

Then for every $\delta \in (0, 1/3)$ *there is a constant* $C = C(\delta, q, K)$ *such that*

$$\big\| M_{ts} - M_{ts}^{\mathcal{D}} \big\|_{L^q} \le C \varepsilon^{\delta/8} \big\| \langle M \rangle_{ts} \big\|_{L^{q/2}}^{(1-\delta)/2} \quad and$$
$$\big\| \mathbb{M}_{ts} - \mathbb{M}_{ts}^{\mathcal{D}} \big\|_{L^{q/2}} \le C \varepsilon^{\delta/4} \big\| \langle M \rangle_{ts} \big\|_{L^{q/2}}^{1-\delta}$$

for every $s, t \in [0, T]$.

PROOF. Let d denote the Carnot-Caratheodoy metric on the step-two free nilpotent Lie group G_d^2 over $\mathbb{R}^d$ (cf. [**FV10**, Chapter 7]). By interpolation and [**FV10**, Proposition 8.15], for fixed $s < t$ and $2 < p' < p < 3$,

$d(\boldsymbol{M}_{ts}, \boldsymbol{M}_{ts}^{\mathcal{D}})$

$\le \Big(d_{0-\text{Höl}}(\boldsymbol{M}, \boldsymbol{M}^{\mathcal{D}})^{1-(p'/p)} d_{p'-\text{var};[s,t]}(\boldsymbol{M}, \boldsymbol{M}^{\mathcal{D}})^{p'/p} \Big)$

$\lesssim \Big(d_\infty(\boldsymbol{M}, \boldsymbol{M}^{\mathcal{D}})^{1-(p'/p)} + d_\infty(\boldsymbol{M}, \boldsymbol{M}^{\mathcal{D}})^{1/2 - p'/(2p)} (\|\boldsymbol{M}\|_\infty^{1/2 - p'/(2p)} + \|\boldsymbol{M}^{\mathcal{D}}\|_\infty^{1/2 - p'/(2p)}) \Big)$

$\quad \times \Big(\|\boldsymbol{M}\|_{p'-\text{var};[s,t]}^{p'/p} + \|\boldsymbol{M}^{\mathcal{D}}\|_{p'-\text{var};[s,t]}^{p'/p} \Big).$

Taking the q-th moment, using Hölder's inequality and Cauchy–Schwarz, we obtain

$$\mathbb{E}\Big[d\big(\boldsymbol{M}_{ts}, \boldsymbol{M}_{ts}^{\mathcal{D}}\big)^q\Big]$$

$$\lesssim \left(\mathbb{E}\Big[d_\infty\big(\boldsymbol{M}, \boldsymbol{M}^{\mathcal{D}}\big)^q\Big] + \sqrt{\mathbb{E}\Big[d_\infty\big(\boldsymbol{M}, \boldsymbol{M}^{\mathcal{D}}\big)^q\Big]\Big(\mathbb{E}\big[\|\boldsymbol{M}\|_\infty^q\big] + \mathbb{E}\big[\|\boldsymbol{M}^{\mathcal{D}}\|_\infty^q\big]\Big)}\right)^{1-(p'/p)}$$

$$\times \left(\mathbb{E}\Big[\|\boldsymbol{M}\|_{p'-\mathrm{var};[s,t]}^q\Big] + \mathbb{E}\Big[\|\boldsymbol{M}^{\mathcal{D}}\|_{p'-\mathrm{var};[s,t]}^q\Big]\right)^{p'/p}.$$

By [**FV10**, Theorem 14.8 and Theorem 14.15],

$$\mathbb{E}\big[\|\boldsymbol{M}\|_\infty^q\big] \lesssim \mathbb{E}\Big[\langle M\rangle_T^{q/2}\Big] \qquad \text{and}$$

$$\mathbb{E}\big[\|\boldsymbol{M}^{\mathcal{D}}\|_\infty^q\big] \leq \mathbb{E}\Big[\|\boldsymbol{M}^{\mathcal{D}}\|_{p-\mathrm{var}}^q\Big] \lesssim \mathbb{E}\Big[\langle M\rangle_T^{q/2}\Big].$$

Using the same theorems, we also have

$$\mathbb{E}\Big[\|\boldsymbol{M}\|_{p'-\mathrm{var};[s,t]}^q\Big] \lesssim \mathbb{E}\Big[\langle M\rangle_{ts}^{q/2}\Big] \qquad \text{and} \qquad \mathbb{E}\Big[\|\boldsymbol{M}^{\mathcal{D}}\|_{p'-\mathrm{var};[s,t]}^q\Big] \lesssim \mathbb{E}\Big[\langle M\rangle_{ts}^{q/2}\Big].$$

The estimate [**FV10**, Equation (14.6) p. 400] gives

$$\mathbb{E}\Big[d_\infty\big(\boldsymbol{M}, \boldsymbol{M}^{\mathcal{D}}\big)^q\Big] \lesssim \left(\mathbb{E}\Big[\sup_{0<v-u\leq|\mathcal{D}|}\|\boldsymbol{M}_{vu}\|^q\Big]\right)^{1/4}\left(\mathbb{E}\Big[\langle M\rangle_T^{q/2}\Big]\right)^{3/4} + \mathbb{E}\Big[\sup_{0<v-u\leq|\mathcal{D}|}\|\boldsymbol{M}_{vu}\|^q\Big]$$

$$\lesssim \varepsilon^{q/4}(K^{3q/4} + 1).$$

Setting $\delta = 1 - p'/p$, we have shown that

$$\Big\|d\big(\boldsymbol{M}_{ts}, \boldsymbol{M}_{ts}^{\mathcal{D}}\big)\Big\|_{L^q} \lesssim \varepsilon^{\delta/8}|\langle M\rangle_{ts}|_{L^{q/2}}^{(1-\delta)/2}.$$

The result follows by equivalence of homogeneous norms on G^2 ([**FV10**, Theorem 7.44]).
□

PROPOSITION 3.11. *Let $M, N\colon D\times[0,T] \to \mathbb{R}$ be continuous local martingale fields adapted to the same filtration. Assume that the quadratic variation of the processes is given by*

$$d\langle M_\bullet(x), M_\bullet(y)\rangle_t = a_t(x,y) \quad resp.$$
$$d\langle N_\bullet(x), N_\bullet(y)\rangle_t = b_t(x,y)$$

for every $x, y \in D$ and every $t \in [0,T]$. Moreover, assume that there is a $\delta \in (0,1]$ such that a and b have continuous modifications in the space $\hat{\mathcal{C}}_b^{0,\delta}$. Let $p > 2$ and $\rho \in (0,\delta)$.

(i) *Assume that there are an $\eta \in (0,\infty)$ and a constant $\kappa > 0$ such that*

$$\sup_{u\in[0,T]}\Big\|\sqrt{a_u(x,x)}\Big\|_{L^q} \leq \frac{\kappa}{1+|x|^\eta}$$

for every $x \in D$ where

$$q > \max\left\{\frac{\delta}{\eta(\delta - \rho)} + d\left(\frac{\delta}{\eta(\delta - \rho)} + \frac{1}{\delta - \rho}\right), \frac{1}{1/2 - 1/p}\right\}. \qquad (3.15)$$

Let M be the modification of the process given in Proposition 3.6.
Then

$$\sup_{0 \leq s < t \leq T} \frac{\left\|M_{ts} - M_{ts}^{\mathcal{D}}\right\|_{\mathcal{C}^\rho}}{|t - s|^{1/p}} \to 0 \qquad (3.16)$$

in probability when $|\mathcal{D}| \to 0$.

(ii) *Define*

$$U_{ts}(x) := \int_s^t (M_{us} N_{\circ du})(x) \quad \text{and} \quad U_{ts}^{\mathcal{D}}(x) := \int_s^t (M_{us}^{\mathcal{D}} N_{du}^{\mathcal{D}})(x)$$

as Stratonovich and Riemann–Stieltjes integral respectively. Assume that there are an $\eta > 0$ and a constant $\kappa > 0$ such that

$$\sup_{u \in [0,T]} \left\|\sqrt{a_u(x,x)}\right\|_{L^q} + \sup_{u \in [0,T]} \left\|\sqrt{b_u(x,x)}\right\|_{L^q} \leq \frac{\kappa}{1 + |x|^\eta} \qquad (3.17)$$

for every $x \in D$ with q sufficiently large as in (3.15). Let U be the modification of the process given in Proposition 3.6.
Then

$$\sup_{0 \leq s < t \leq T} \frac{\left\|U_{ts} - U_{ts}^{\mathcal{D}}\right\|_{\mathcal{C}^\rho}}{|t - s|^{2/p}} \to 0 \qquad (3.18)$$

in probability when $|\mathcal{D}| \to 0$.

(iii) *Assume that $\sup_{u \in [0,T]} \|a_u\|_\delta^\wedge$ and $\sup_{u \in [0,T]} \|b_u\|_\delta^\wedge$ are almost surely bounded random variables and that (3.17) holds for some $\eta \in (0, \infty)$ and $\kappa > 0$ uniformly for all $q \geq 1$.*

Then the convergences in (3.16) and (3.18) hold in L^q for every $q \geq 1$.

PROOF. We will only give a proof in the case of a and b having compact support. The more general case works analogous using the stated growth conditions, as seen in the proof of Proposition 3.6. We start with (i). Assume first that $\sup_{u \in [0,T]} \|a_u\|_\delta^\wedge$ is an almost surely bounded random variable. Fix $x, y \in D$ such that $0 < |x - y| \leq 1$ and let $s < t \in [0,T]$. Define the martingale $\hat{M} := M(x) - M(y)$ and let $\hat{\boldsymbol{M}}$ denote its canonical rough path lift (given by $\hat{M}$ and its iterated Stratonovich integrals). From the Burkholder–Davis–Gundy inequality for enhanced martingales (cf. [**FV10**, Theorem 14.8]), for every $q \geq 2$,

$$\left\| \sup_{0<v-u\leq|\mathcal{D}|} \left\| \hat{\boldsymbol{M}}_{vu} \right\| \right\|_{L^q} \leq C \left\| \sqrt{\langle \hat{M} \rangle_{|\mathcal{D}|}} \right\|_{L^q} \leq C|\mathcal{D}|^{1/2}|x-y|^\delta$$

where the constant C depends on the essential supremum of $\sup_{u\in[0,T]} \|a_u\|_\delta^\wedge$. We may assume that $|\mathcal{D}|$ is sufficiently small such the right hand side of the equation is smaller than 1. By Lemma 3.10, for every $q \geq 2$ and every $\beta \in (0,1/3)$,

$$\left\| M_{ts}(x) - M_{ts}^{\mathcal{D}}(x) - M_{ts}(y) + M_{ts}^{\mathcal{D}}(y) \right\|_{L^q}$$
$$\leq C|\mathcal{D}|^{\beta/16}|x-y|^{\beta\delta/8} \left\| \langle M(x) - M(y) \rangle_{ts} \right\|_{L^{q/2}}^{(1-\beta)/2}$$
$$\leq C|\mathcal{D}|^{\beta/16}|x-y|^{\delta(1-7\beta/8)}|t-s|^{(1-\beta)/2}.$$

We have already seen that

$$\left\| M_{ts}(x) - M_{ts}(y) \right\|_{L^q} \leq C|x-y|^\delta|t-s|^{1/2},$$

and by the triangle inequality,

$$\left\| M_{ts}^{\mathcal{D}}(x) - M_{ts}^{\mathcal{D}}(y) \right\|_{L^q} \leq C|x-y|^\delta|t-s|^{(1-\beta)/2}.$$

Choosing β small enough and q large enough, we can apply Theorem 2.17 to see that

$$\left\| \sup_{s<t} \frac{\left\| M_{ts} - M_{ts}^{\mathcal{D}} \right\|_{\mathcal{C}^\rho}}{|t-s|^{1/p}} \right\|_{L^q} \leq C|\mathcal{D}|^{\beta/16}$$

which shows the claim. If the essential supremum of $\sup_{u\in[0,T]} \|a_u\|_\delta^\wedge$ is not bounded, define the stopping times

$$\tau_n = \inf \left\{ t \in [0,T] : \sup_{u\in[0,t]} \|a_u\|_\delta^\wedge \geq n \right\}$$

and set $M^n := M_{t\wedge\tau_n}$. We can repeat the argument above and conclude that

$$\sup_{s<t} \frac{\left\| M_{ts}^n - M_{ts}^{n;\mathcal{D}} \right\|_{\mathcal{C}^\rho}}{|t-s|^{1/p}} \to 0$$

in probability as $|\mathcal{D}|$ tends to 0, and every $n \in \mathbb{N}$. Fix some $\varepsilon > 0$ and some $n \in \mathbb{N}$. Then

$$\mathbb{P}\left(\left| \sup_{s<t} \frac{\left\| M_{ts} - M_{ts}^{\mathcal{D}} \right\|_{\mathcal{C}^\rho}}{|t-s|^{1/p}} \right| \geq \varepsilon \right)$$

$$\leq \mathbb{P}\left(\left| \sup_{s<t} \frac{\left\| M_{ts} - M_{ts}^{\mathcal{D}} \right\|_{\mathcal{C}^\rho}}{|t-s|^{1/p}} - \sup_{s<t} \frac{\left\| M_{ts}^n - M_{ts}^{n;\mathcal{D}} \right\|_{\mathcal{C}^\rho}}{|t-s|^{1/p}} \right| \geq \frac{\varepsilon}{2} \right)$$

$$+ \mathbb{P} \left(\left| \sup_{s<t} \frac{\left\| M_{ts}^n - M_{ts}^{n;\mathcal{D}} \right\|_{\mathcal{C}^\rho}}{|t-s|^{1/p}} \right| \geq \frac{\varepsilon}{2} \right)$$

$$\leq \mathbb{P} \left(\tau_n < T \right) + \mathbb{P} \left(\left| \sup_{s<t} \frac{\left\| M_{ts}^n - M_{ts}^{n;\mathcal{D}} \right\|_{\mathcal{C}^\rho}}{|t-s|^{1/p}} \right| \geq \frac{\varepsilon}{2} \right),$$

and the first term converges to 0 a $n \to \infty$. This shows that indeed

$$\sup_{s<t} \frac{\left\| M_{ts} - M_{ts}^{\mathcal{D}} \right\|_{\mathcal{C}^\rho}}{|t-s|^{1/p}} \to 0$$

in probability when $|\mathcal{D}| \to 0$.

Now we prove (ii). As above, we first assume that $\sup_{u \in [0,T]} \|a_u\|_\delta^\wedge$ and $\sup_{u \in [0,T]} \|b_u\|_\delta^\wedge$ are almost surely bounded random variables. Let $x, y \in D$ with $0 < |x-y| \leq 1$ and $s < t$. We have already seen that in case of the Itô integral, for every $q \geq 2$,

$$\left\| \int_s^t (M_{us} N_{du})(x) - \int_s^t (M_{us} N_{du})(y) \right\|_{L^{q/2}} \leq C|t-s||x-y|^\delta.$$

Moreover, we have

$$\left\| \langle M(x), N(x) \rangle_{ts} - \langle M(y), N(y) \rangle_{ts} \right\|_{L^{q/2}}$$

$$\leq \left\| \langle M(x) - M(y), N(x) \rangle_{ts} \right\|_{L^{q/2}} + \left\| \langle M(y), N(x) - N(y) \rangle_{ts} \right\|_{L^{q/2}}$$

$$\leq \left\| \sqrt{\langle M(x) - M(y) \rangle_{ts}} \right\|_{L^q} \left\| \sqrt{\langle N(x) \rangle_{ts}} \right\|_{L^q} + \left\| \sqrt{\langle N(x) - N(y) \rangle_{ts}} \right\|_{L^q} \left\| \sqrt{\langle M(x) \rangle_{ts}} \right\|_{L^q}$$

$$\leq C|t-s||x-y|^\delta.$$

This implies that the same estimate holds for the Stratonovich integral, thus

$$\left\| U_{ts}(x) - U_{ts}(y) \right\|_{L^{q/2}} \leq C|t-s| \, |x-y|^\delta.$$

By the triangle inequality,

$$\left\| U_{ts}^{\mathcal{D}}(x) - U_{ts}^{\mathcal{D}}(y) \right\|_{L^{q/2}}$$

$$\leq \left\| \int_s^t (M_{us}^{\mathcal{D}}(x) - M_{us}^{\mathcal{D}}(y)) N_{du}^{\mathcal{D}}(x) \right\|_{L^{q/2}} + \left\| \int_s^t M_{us}^{\mathcal{D}}(y)(N_{du}^{\mathcal{D}}(x) - N_{du}^{\mathcal{D}}(y)) \right\|_{L^{q/2}}$$

Now we define the martingale $\hat{M} := ((M(x) - M(y))/|x-y|^\delta, N(x))$. We can check that for its quadratic variation, we have

$$\left\| \langle \hat{M} \rangle_{ts} \right\|_{L^{q/2}} \leq C|t-s|$$

where the constant C does not depend on x or y. From [**FV10**, Theorem 14.15], it

follows that

$$\left\|\int_s^t \left(M_{us}^{\mathcal{D}}(x) - M_{us}^{\mathcal{D}}(y)\right) N_{du}^{\mathcal{D}}(x)\right\|_{L^{q/2}} \leq C|t-s|\,|x-y|^\delta.$$

Similarly,

$$\left\|\int_s^t M_{us}^{\mathcal{D}}(y) \left(N_{du}^{\mathcal{D}}(x) - N_{du}^{\mathcal{D}}(y)\right)\right\|_{L^{q/2}} \leq C|t-s|\,|x-y|^\delta$$

and hence

$$\left\|U_{ts}^{\mathcal{D}}(x) - U_{ts}^{\mathcal{D}}(y)\right\|_{L^{q/2}} \leq C|t-s||x-y|^\delta.$$

Therefore, by the triangle inequality,

$$\left\|U_{ts}(x) - U_{ts}^{\mathcal{D}}(x) - U_{ts}(y) + U_{ts}^{\mathcal{D}}(y)\right\|_{L^{q/2}} \leq C|t-s|\,|x-y|^\delta. \tag{3.19}$$

On the other hand, we can apply Lemma 3.10 to the martingale $M(x)$, x fixed, to see that for every $q \geq 2$ and every $\beta \in (0, 1/3)$,

$$\left\|U_{ts}(x) - U_{ts}^{\mathcal{D}}(x)\right\|_{L^{q/2}} \leq C|\mathcal{D}|^{\beta/8} \left\|\langle M(x)\rangle_{ts}\right\|_{L^{q/2}}^{1-\beta} \leq C|\mathcal{D}|^{\beta/8}\,|t-s|^{1-\beta}.$$

The same estimate holds if we replace x by y, and by the triangle inequality, we also get

$$\left\|U_{ts}(x) - U_{ts}^{\mathcal{D}}(x) - U_{ts}(y) + U_{ts}^{\mathcal{D}}(y)\right\|_{L^{q/2}} \leq C|\mathcal{D}|^{\beta/8}|t-s|^{1-\beta}. \tag{3.20}$$

Interpolating the two inequalities (3.19) and (3.20), we see that for every $q \geq 2$, every $\beta \in (0, 1/3)$ and every $\lambda \in [0, 1]$, we have the estimate

$$\left\|U_{ts}(x) - U_{ts}^{\mathcal{D}}(x) - U_{ts}(y) + U_{ts}^{\mathcal{D}}(y)\right\|_{L^{q/2}} \leq C|\mathcal{D}|^{\lambda\beta/8}|x-y|^{(1-\lambda)\delta}|t-s|^{1-\lambda\beta}.$$

Choosing $\lambda > 0$ and $\beta > 0$ small enough and q large enough, we can again use Theorem 2.17 to see that

$$\left\|\sup_{s<t} \frac{\|U_{ts} - U_{ts}^{\mathcal{D}}\|_{\mathcal{C}^\rho}}{|t-s|^{2/p}}\right\|_{L^{q/2}} \leq C|\mathcal{D}|^{\lambda\beta/8}$$

which proves the claim if $\sup_{u \in [0,T]} \|a_u\|_\delta^\wedge$ and $\sup_{u \in [0,T]} \|b_u\|_\delta^\wedge$ are bounded random variables. The general case follows by the same stopping argument as above. □

THEOREM 3.12. *Let M be a continuous local martingale velocity field in $\mathcal{C}^{2,\delta}(D, \mathbb{R}^d)$, for some $\delta \in (0, 1]$, with continuous local characteristic a in $\hat{\mathcal{C}}_b^{2,\delta}$.*

(i) *Let $\rho \in (0, \delta)$ and $p \in (2, 3)$. Assume that there are an $\eta \in (0, \infty)$ and a constant $\kappa > 0$ such that*

$$\sum_{0 \leq |\alpha| \leq 2} \sup_{u \in [0,T]} \left\| \sqrt{\partial_x^\alpha \partial_y^\alpha a_u(x,y)|_{x=y=z}} \right\|_{L^q} \leq \frac{\kappa}{1 + |z|^\eta} \tag{3.21}$$

for every $z \in D$ where q satisfies

$$q > \max \left\{ \frac{\delta}{\eta(\delta - \rho)} + d \left(\frac{\delta}{\eta(\delta - \rho)} + \frac{1}{\delta - \rho} \right), \frac{1}{1/2 - 1/p} \right\}.$$

Let $\boldsymbol{M} = (M, \mathbb{M})$ be the weak geometric (p, ρ)-rough driver given in Theorem 3.7. Then

$$\boldsymbol{M}^{\mathcal{D}} \to \boldsymbol{M} \tag{3.22}$$

in probability a $|\mathcal{D}| \to 0$, with $\boldsymbol{M}^{\mathcal{D}}$ defined as in (3.14).

(ii) *Assume that* $\sup_{u \in [0,T]} \|a_u\|_\delta^\wedge$ *is almost surely bounded, and that the growth condition* (3.21) *holds for some* $\eta \in (0, \infty)$ *and a constant* $\kappa > 0$ *uniformly for all* $q \geq 1$.

Then for every $p \in (2, 3)$ *and* $\rho \in (0, \delta)$, $\boldsymbol{M} = (M, \mathbb{M})$ *has a modification which is a weak geometric* (p, ρ)-*rough driver, and the convergence in* (3.22) *holds in* L^q *for every* $q \geq 1$.

PROOF.   This follows by applying Proposition 3.11 to M, W, $M^{\mathcal{D}}$ and $W^{\mathcal{D}}$ and its derivatives. We use the product rule in Proposition 3.5 for calculating the derivatives of W and Proposition 5.1 for the derivatives of $W^{\mathcal{D}}$. The details are left to the reader.   □

REMARK 3.13.   As seen in the proof of Proposition 3.11, assertion (i) in Theorem 3.12 holds without any moment conditions on a in the case where M or a have compact support.

It follows then directly from this statement and the continuity of the Itô solution map, Theorem 2.8, that the solution flow to the equation

$$d\varphi = \boldsymbol{M}(\varphi \, ; dt)$$

satisfies a Wong–Zakaï theorem. Using that the flow coincides with the one generated by the corresponding Stratonovich SDE, we obtain the following corollary.

COROLLARY   3.14.   *Let* M *be a continuous local martingale velocity field in* $\mathcal{C}^{2,\delta}(\mathbb{R}^d, \mathbb{R}^d)$, *for some* $\delta \in (2/3, 1]$, *with local characteristic* a *in* $\widehat{\mathcal{C}}_b^{2,\delta}$. *Assume that* M *or* a *have compact support or that the growth assumption* (3.12) *holds. Let* φ *be the flow generated by the Stratonovich solution to*

$$d\varphi = M(\varphi \, ; \circ dt)$$

and $\varphi^{\mathcal{D}}$ be the pathwise solution to

$$d\varphi^{\mathcal{D}} = M^{\mathcal{D}}\big(\varphi^{\mathcal{D}} \, ; dt\big).$$

Then $\varphi^D \to \varphi$ in the space of $\mathcal{C}^\rho$ homeomorphisms uniformly in probability when $|\mathcal{D}| \to 0$ for all $\rho \in (0, \delta)$.

4. Application: Large deviations.

We provide in this Section another illustration of use of the continuity of the Itô map by proving a large deviation theorem for Brownian flows. Relatively few works were dedicated to these topics, and we mention [**BDM10**] and [**DD12**]. In [**BDM10**], Dupuis and his co-authors use Dupuis' weak convergence approach to large deviation principles to prove such a result for Brownian flows of maps, building on a general large deviation criterion proved earlier in [**BDM08**]. Dereich and Dimitroff's approach in [**DD12**] is more in the line of the present work. In the spirit of Elworthy's approach to stochastic differential equations [**Elw82**], they consider Brownian flows of maps as solutions to rough differential equations driven by a Banach space valued Brownian rough path, whose construction in a vector field setting was made possible by the previous work [**Der10**] of Dereich. The support and large deviation theorems for Brownian flows are then inherited from the corresponding results proved in [**Der10**] for the above mentioned vector field-valued Brownian rough path, via the classical Itô map. Our approach here is similar to the latter, and perhaps more direct.

Let $(E, \mathcal{H}_1, \gamma)$ be a Gaussian Banach space with norm $\| \cdot \|$, i.e. $(E, \| \cdot \|)$ is a separable Banach space, γ is a Gaussian measure defined on the Borel σ algebra and $\mathcal{H}_1$ the Cameron–Martin Hilbert space (cf. [**Bog98**] or [**Led96**, Chapter 4] for the precise definitions and further properties). Recall that $\mathcal{H}_1$ is continuously embedded in E, and for every $h \in \mathcal{H}_1$,

$$\|h\| \leq \sigma_\gamma \|h\|_{\mathcal{H}_1} = \sigma_\gamma \sqrt{\langle h, h \rangle_{\mathcal{H}_1}}$$

where

$$\sigma_\gamma^2 := \int_E \|x\|^2 \, \gamma(dx).$$

A process $X \colon [0, T] \to E$, defined on some probability space, is called an *E-valued Wiener process* if it has almost surely continuous sample paths starting from 0, has independent increments, and for every $\xi \in E^*$, the distribution of $\langle X_t - X_s, \xi \rangle$ is a centered Gaussian random variable with variance $|t - s| \|\xi\|_{\mathcal{H}_1}^2$ (cf. [**LLQ02**] and [**Der10**] for more properties about E-valued Wiener processes). The law of X on the space $\mathcal{C}([0, T], E)$ is again Gaussian, and one can see that the corresponding Cameron–Martin space $\mathcal{H}$ is given by

$$\mathcal{H} = \left\{ \int_0^\bullet f_s \, ds \, : \, f \in L^2([0, T], \mathcal{H}_1) \right\}$$

where the integral is a Bochner-integral. Moreover, if $h_t^i = \int_0^t \dot{h}_s^i \, ds$, $i = 1, 2$, the scalar product is given by

$$\langle h^1, h^2 \rangle_{\mathcal{H}} = \int_0^T \langle \dot{h}_s^1, \dot{h}_s^2 \rangle_{\mathcal{H}_1} \, ds.$$

Let $p \in [1, \infty)$. In the following, we will use the notion of *p-variation* of a path $h \colon [0, T] \to E$ which is defined as

$$\|h\|_{p-\mathrm{var};[s,t]} := \sup_{(t_i) \subset [s,t]} \left(\sum_{t_i} \left\| h_{t_{i+1}} - h_{t_i} \right\|^p \right)^{1/p}$$

where the supremum is taken over all finite partitions (t_i) of the interval $[s, t]$. If $\|h\|_{1-\mathrm{var};[0,T]} < \infty$, we say that h has finite variation.

LEMMA 4.1. *For every $h \in \mathcal{H}$ we have*

$$\sup_{0 \leq s < t \leq T} \frac{\|h_t - h_s\|}{|t - s|^{1/2}} \leq \sigma_\gamma \sqrt{\langle h, h \rangle_{\mathcal{H}}} \tag{4.1}$$

and for $[s, t] \subseteq [0, T]$,

$$\|h\|_{1-var;[s,t]} \leq \sigma_\gamma |t - s|^{1/2} \sqrt{\langle h, h \rangle_{\mathcal{H}}}. \tag{4.2}$$

In particular, every $h \in \mathcal{H}$ is $1/2$-Hölder continuous and has finite variation on $[0, T]$.

PROOF. Clearly, (4.1) follows from (4.2), hence we only prove the second estimate. Let $h \in \mathcal{H}$ with $h(t) = \int_0^t \dot{h}_s \, ds$ and let (t_i) be a partition of some interval $[s, t] \subseteq [0, T]$. Then

$$\sum_i \|h_{t_{i+1}} - h_{t_i}\| \leq \sum_i \int_{t_i}^{t_{i+1}} \|\dot{h}_u\| \, du = \int_s^t \|\dot{h}_u\| \, du$$

$$\leq |t - s|^{1/2} \left(\int_0^T \|\dot{h}_u\|^2 \, du \right)^{1/2} \leq \sigma_\gamma |t - s|^{1/2} \sqrt{\langle h, h \rangle_{\mathcal{H}}}.$$

Taking the supremum over all partitions shows the claim. $\qquad\square$

Let D be an open, relatively compact, connected subset of $\mathbb{R}^d$, $m \in \mathbb{N}_0$ and $\delta \in (0, 1]$. In the following, we would like to take the space $\mathcal{C}_b^{m,\delta}(D, \mathbb{R}^d)$ as E and consider a Gaussian measure on this space. However, $\mathcal{C}_b^{m,\delta}(D, \mathbb{R}^d)$ is not separable (which is usually the case for Hölder-type spaces). Instead, we define the space $\mathcal{C}_b^{m,0,\delta}(D, \mathbb{R}^d)$ as the closure of smooth paths from D to $\mathbb{R}^d$ with respect to the norm $\| \cdot \|_{m+\delta}$. As for Hölder spaces, using boundedness of D, one can show that these spaces are separable. From now on, let $E = \mathcal{C}_b^{m,0,\delta}(D, \mathbb{R}^d)$ and assume that there is a Gaussian Banach space $(E, \mathcal{H}_1, \gamma)$.

If v is a $\mathcal{C}_b^{m,\delta}(D, \mathbb{R}^d)$ valued path with finite variation and if $m \geq 1$, we define the pair $S(v)_{ts} = (v_{ts}, \mathbb{v}_{ts})$ by setting

$$v_{ts}(x) = v_t(x) - v_s(x) \qquad \text{and} \qquad \mathbb{v}_{ts} = w_{ts} + \frac{1}{2} v_{ts} v_{ts}$$

where w_{ts} is the first order differential operator

$$w_{ts} = \frac{1}{2} \left(\int_s^t v_{us}.v_{du} - v_{du}.v_{us} \right)$$

and the integral is a Riemann–Stieltjes integral. Note that if X is a Wiener process in $\mathcal{C}_b^{m,0,\delta}$, $S(h)$ is always defined for every Cameron–Martin path h since these paths are continuous and have bounded variation by Lemma 4.1. Moreover, the following holds:

LEMMA 4.2. *Let $(E, \mathcal{H}_1, \gamma)$ be a Gaussian Banach space with $E = \mathcal{C}_b^{2,0,\delta}(D, \mathbb{R}^d)$ and $\delta \in (0,1]$. Then, for every $h \in \mathcal{H}$, $S(h)$ is a geometric $(2,\delta)$-rough driver, and a constant C such that*

$$\sup_{s<t} \frac{\|h_t - h_s\|_{\mathcal{C}^{2+\delta}}}{|t-s|^{1/2}} \leq \sigma_\gamma \sqrt{\langle h,h \rangle} \quad and \quad \sup_{s<t} \frac{\|w_{ts}\|_{\mathcal{C}^{1+\delta}}}{|t-s|} \leq C\sigma_\gamma^2 \langle h,h \rangle$$

where

$$w_{ts} = \frac{1}{2} \left(\int_s^t h_{us}.h_{du} - h_{du}.h_{us} \right).$$

PROOF. Let $h \in \mathcal{H}$ and $S(h) = (h, \mathbb{h})$ be defined as above. The claim for h follows directly from Lemma 4.1, and the algebraic condition for $(h, \mathbb{h})$ follows from well-known identities for Riemann–Stieltjes integrals. Let $i, k \in \{1, \ldots, d\}$, $x \in D$ and $s < t$. Then, by Riemann–Stieltjes estimates and Lemma 4.1,

$$\left| \int_s^t h_{us}^i(x)\, \partial_i h_{du}^k(x) \right| \leq \sup_{u \in [s,t]} |h_{us}^i(x)| \sup_{(t_j) \subset [s,t]} \sum_j \left| \partial_i h_{t_{j+1}}^k(x) - \partial_i h_{t_j}^k(x) \right|$$

$$\leq \sup_{u \in [s,t]} \|h_{us}\|_{2+\delta} \sup_{(t_j) \subset [s,t]} \sum_j \left\| h_{t_{j+1}} - h_{t_j} \right\|_{2+\delta}$$

$$\leq \sigma_\gamma^2 |t-s| \langle h,h \rangle.$$

One can perform the same estimate for the second term in w_{ts}. By the triangle inequality, this shows that

$$\sup_{s<t} \frac{\|w_{ts}\|_{\mathcal{C}^0}}{|t-s|} \leq C \langle h,h \rangle.$$

We can calculate the derivative of w using Proposition 5.1 and perform similar estimates, also for the Hölder norm, to conclude. □

4.1. Schilder's theorem and Freidlin–Ventzel large deviations for stochastic flows.

In this Section, we will prove a large deviation result for a (p,ρ)-rough driver $\boldsymbol{X}$ in the case where the underlying vector field X is a Wiener process. We will do this by using the *extended contraction principle* (see e.g. [**DZ98**, Theorem 4.2.23]), a strategy which has proven to be useful in rough paths theory ([**LQZ02**], [**MSS06**], [**FV07**], [**FV10**]). As a corollary, we obtain a Freidlin–Ventzel-type large deviation result for the flow generated by this driver. The key step is to prove that $\boldsymbol{X}^n := \boldsymbol{X}^{\mathcal{D}_n}$ with

$\mathcal{D}_n = \{kT/n : k = 0, \ldots, n\}$ is an *exponentially good approximation* to $\boldsymbol{X}$. This is done in Lemmata 4.3 and 4.4.

For $\varepsilon > 0$, set $\delta_\varepsilon \boldsymbol{X} := (\varepsilon X, \delta_\varepsilon \mathbb{X})$ where

$$
\begin{aligned}
\delta_\varepsilon \mathbb{X}_{ts} &:= \delta_\varepsilon W_{ts} + \frac{\varepsilon^2}{2} X_{ts} X_{ts} \\
&:= \frac{1}{2} \left(\int_s^t (\varepsilon X)_{us}.(\varepsilon X)_{\circ du} - (\varepsilon X)_{\circ du}.(\varepsilon X)_{us} \right) + \frac{\varepsilon^2}{2} X_{ts} X_{ts}.
\end{aligned}
$$

We similarly define $\delta_\varepsilon \boldsymbol{X}^n$ with Riemann–Stieltjes integrals. Note that the homogeneous metric $\mathfrak{d}$ defined in (2.6) enjoys the property that $\mathfrak{d}(\delta_\varepsilon \boldsymbol{X}, \delta_\varepsilon \boldsymbol{X}^n) = \varepsilon \mathfrak{d}(\boldsymbol{X}, \boldsymbol{X}^n)$ for every $\varepsilon > 0$ (which is the reason why we call it *homogeneous* metric). Note that since X is Gaussian, the local characteristic a is almost surely deterministic. Moreover, since D is bounded, the growth condition (3.12) in Theorem 3.7 is trivially satisfied for a and possible derivatives.

LEMMA 4.3. *Let D be a relatively compact domain in $\mathbb{R}^d$ and X be a Wiener process in $\mathcal{C}_b^{2,0,\delta}$ for some $\delta \in (0,1]$. Let $\boldsymbol{X} = (X, \mathbb{X})$ denote its natural lift to a (p,ρ)-rough driver for some $\rho \in (0,\delta)$ and $p \in (2,3)$. Let $\eta > 0$ be fixed. Then the following holds*:

$$
\lim_{n \to \infty} \limsup_{\varepsilon \to 0} \varepsilon^2 \log \mathbb{P}\Big(\mathfrak{d}_{p,\rho}(\delta_\varepsilon \boldsymbol{X}, \delta_\varepsilon \boldsymbol{X}^n) > \eta \Big) = -\infty.
$$

PROOF. Let $\varepsilon > 0$ and $n \in \mathbb{N}$ be fixed. Since X is Gaussian, the quadratic variation process a is deterministic, and all estimates for X, X^n, $X - X^n$ and its iterated integrals in the proof of Proposition 3.11 hold for $q = 2$. Moreover, $|X|_{L^q} \lesssim \sqrt{q}|X|_{L^2}$. The iterated Stratonovich and Riemann–Stieltjes integrals are both elements in the second inhomogeneous Wiener chaos, therefore

$$
\left\| \int \circ dX \otimes \circ dX \right\|_{L^q} \lesssim q \left\| \int \circ dX \otimes \circ dX \right\|_{L^2} \text{ and } \left\| \int dX^n \otimes dX^n \right\|_{L^q} \lesssim q \left\| \int dX^n \otimes dX^n \right\|_{L^2},
$$

for all $q \geq 2$, see e.g. [**FV10**, Theorem D.8], and similar estimates hold for the other quantities. Therefore, we may apply Theorem 2.17 with κ equal to a constant times $\sqrt{q}$ which shows that

$$
\left\| \mathfrak{d}_{p,\rho}(\boldsymbol{X}, \boldsymbol{X}^n) \right\|_{L^q} = \alpha_n \sqrt{q}
$$

holds for all $q \geq 2$ and some constant α_n. Repeating this argument for every $n \in \mathbb{N}$, we obtain a sequence (α_n) converging to 0 as $n \to \infty$. Thus

$$
\begin{aligned}
\mathbb{P}\big(\mathfrak{d}_{p,\rho}(\delta_\varepsilon \boldsymbol{X}, \delta_\varepsilon \boldsymbol{X}^n) > \eta \big) &= \mathbb{P}\left(\mathfrak{d}_{p,\rho}(\boldsymbol{X}, \boldsymbol{X}^n) > \frac{\eta}{\varepsilon} \right) \\
&\leq \left(\frac{\varepsilon}{\eta} \right)^q q^{q/2} \alpha_n^q \\
&\leq \exp\left[q \log\left(\frac{\varepsilon \alpha_n \sqrt{q}}{\eta} \right) \right].
\end{aligned}
$$

Choosing $q = \varepsilon^{-2}$ we obtain the inequality

$$\varepsilon^2 \log \mathbb{P}\big(\mathfrak{d}_{p,\rho}(\delta_\varepsilon \boldsymbol{X}, \delta_\varepsilon \boldsymbol{X}^n) > \eta\big) \leq \log(\alpha_n/\eta)$$

from which the claim follows.   □

If $\mathcal{H}$ is the Cameron–Martin space of a $\mathcal{C}_b^{2,0,\delta}$-valued Wiener process and v is a path with values in $\mathcal{C}_b^{2,0,\delta}$, set

$$I(v) := \begin{cases} (1/2)\langle v, v\rangle_{\mathcal{H}}, & \text{if } v \in \mathcal{H} \\ +\infty & \text{otherwise.} \end{cases}$$

LEMMA 4.4. *Let $\mathcal{H}$ be the Cameron–Martin space for some $\mathcal{C}_b^{2,0,\delta}$-valued Wiener process. Choose $\Lambda > 0$. Then*

$$\lim_{|\mathcal{D}| \to 0} \sup_{\{h \in \mathcal{H} : I(h) \leq \Lambda\}} \mathfrak{d}_{p,\delta}\Big(S(h^{\mathcal{D}}), S(h)\Big) = 0$$

for every $p > 2$.

PROOF. It is easy to check (cf. [**FV10**, Proposition 5.20] and Lemma 4.2) that

$$\sup_{0 \leq s < t \leq T} \frac{\big\|h_t^{\mathcal{D}} - h_s^{\mathcal{D}}\big\|_{2+\delta}}{|t-s|^{1/2}} \leq \sqrt{3} \sup_{0 \leq s < t \leq T} \frac{\|h_t - h_s\|_{2+\delta}}{|t-s|^{1/2}} \leq \sqrt{3}\sigma_\gamma \sqrt{\langle h, h\rangle}.$$

From Lemma 4.2, we know that

$$\sup_{s<t} \frac{\|w_{ts}\|_{\mathcal{C}^{1+\delta}}}{|t-s|} \leq C\sigma_\gamma^2 \langle h, h\rangle$$

Now fix $j, k \in \{1, \ldots, d\}$, $s < t$ and $x \in D$. Then

$$\left|\int_s^t h_{su}^{\mathcal{D};j}(x)\, \partial_j h_{du}^{\mathcal{D};k}(x)\right| \leq \big\|h^{\mathcal{D};j}(x)\big\|_{1-\text{var};[s,t]} \big\|\partial_j h^{\mathcal{D};k}(x)\big\|_{1-\text{var};[s,t]}.$$

For $u \in [0, T]$, define $u_{\mathcal{D}} := \sup\{t_i \in \mathcal{D} : t_i \leq u\}$ and $u^{\mathcal{D}} := \inf\{t_i \in \mathcal{D} : t_i \geq u\}$. With this notation, using the estimates from Lemma 4.1,

$$
\begin{aligned}
\big\|h^{\mathcal{D};j}(x)\big\|_{1-\text{var};[s,t]} &\leq \big\|h^{\mathcal{D};j}(x)\big\|_{1-\text{var};[s,s^{\mathcal{D}}]} + \big\|h^{\mathcal{D};j}(x)\big\|_{1-\text{var};[s^{\mathcal{D}},t_{\mathcal{D}}]} + \big\|h^{\mathcal{D};j}(x)\big\|_{1-\text{var};[t_{\mathcal{D}},t]} \\
&\leq \frac{|s^{\mathcal{D}} - s|}{|s^{\mathcal{D}} - s_{\mathcal{D}}|} \big|h_{s^{\mathcal{D}}}^j(x) - h_{s_{\mathcal{D}}}^j(x)\big| + \big\|h^j(x)\big\|_{1-\text{var};[s^{\mathcal{D}},t_{\mathcal{D}}]} \\
&\quad + \frac{|t - t_{\mathcal{D}}|}{|t^{\mathcal{D}} - t_{\mathcal{D}}|} \big|h_{t^{\mathcal{D}}}^j(x) - h_{t_{\mathcal{D}}}^j(x)\big| \\
&\leq \frac{|s^{\mathcal{D}} - s|}{|s^{\mathcal{D}} - s_{\mathcal{D}}|^{1/2}} \sup_{0 \leq u < v \leq T} \frac{\|h_v - h_u\|_0}{|v-u|^{1/2}} + \sigma_\gamma \big|t_{\mathcal{D}} - s^{\mathcal{D}}\big|^{1/2} \sqrt{\langle h, h\rangle}
\end{aligned}
$$

$$+ \frac{|t - t_{\mathcal{D}}|}{|t^{\mathcal{D}} - t_{\mathcal{D}}|^{1/2}} \sup_{0 \le u < v \le T} \frac{\|h_v - h_u\|_0}{|v - u|^{1/2}}$$
$$\le 3\sigma_\gamma \sqrt{\langle h, h \rangle} |t - s|^{1/2}.$$

A similar estimate holds for $\partial_j h^{\mathcal{D};k}$. Therefore,

$$\sup_{0 \le s < t \le T} \frac{\left\| \int_s^t h_{su}^{\mathcal{D};j}(\cdot) \, \partial_j h_{du}^{\mathcal{D};k}(\cdot) \right\|_0}{|t - s|} \le 9\sigma_\gamma^2 \langle h, h \rangle.$$

Let $s < t$ and $x, y \in D$. We have

$$\left| \int_s^t h_{su}^{\mathcal{D};j}(x) \, \partial_j h_{du}^{\mathcal{D};k}(x) - \int_s^t h_{su}^{\mathcal{D};j}(y) \, \partial_j h_{du}^{\mathcal{D};k}(y) \right|$$
$$\le \left\| h^{\mathcal{D};j}(x) - h^{\mathcal{D};j}(y) \right\|_{1-\mathrm{var};[s,t]} \left\| \partial_j h^{\mathcal{D};k} \right\|_{1-\mathrm{var};[s,t]}$$
$$+ \left\| h^{\mathcal{D};j} \right\|_{1-\mathrm{var};[s,t]} \left\| \partial_j h^{\mathcal{D};k}(x) - \partial_j h^{\mathcal{D};k}(y) \right\|_{1-\mathrm{var};[s,t]}$$

Similar to $h^{\mathcal{D};j}(x)$, one can estimate

$$\left\| h^{\mathcal{D};j}(x) - h^{\mathcal{D};j}(y) \right\|_{1-\mathrm{var};[s,t]} \le 3\sigma_\gamma \sqrt{\langle h, h \rangle} \, |t - s|^{1/2} \, |x - y|^\delta$$

and similarly for $\partial_j h^{\mathcal{D};k}(x) - \partial_j h^{\mathcal{D};k}(y)$. Thus

$$\sup_{0 \le s < t \le T} \frac{\left\| \int_s^t h_{su}^{\mathcal{D};j}(\cdot) \, \partial_j h_{du}^{\mathcal{D};k}(\cdot) \right\|_\delta}{|t - s|} \le 9\sigma_\gamma^2 \langle h, h \rangle.$$

Using the product rule (Proposition 5.1),

$$\partial_i \left(\int_s^t h_{su}^{\mathcal{D};j}(x) \, \partial_j h_{du}^{\mathcal{D};k}(x) \right) = \int_s^t \partial_i h_{su}^{\mathcal{D};j}(x) \, \partial_j h_{du}^{\mathcal{D};k}(x) + \int_s^t h_{su}^{\mathcal{D};j}(x) \, \partial_{i,j}^2 h_{du}^{\mathcal{D};k}(x)$$

and similar estimates as above, we can show that

$$\sup_{0 \le s < t \le T} \frac{\left\| w_{ts}^{\mathcal{D}} \right\|_{1+\delta}}{|t - s|} \le C\sigma_\gamma^2 \langle h, h \rangle$$

where

$$w_{ts}^{\mathcal{D}} = \frac{1}{2} \left(\int_s^t h_{us}^{\mathcal{D}} . h_{du}^{\mathcal{D}} - h_{du}^{\mathcal{D}} . h_{us}^{\mathcal{D}} \right).$$

This implies

$$\sup_{\mathcal{D}} \sup_{h \, : \, I(h) \le \Lambda} \left\| S(h^{\mathcal{D}}) \right\|_{2,\delta} < \infty.$$

Let $p > 2$. Then

$$\sup_{0 \leq s < t \leq T} \frac{\left\| h_{ts}^{\mathcal{D}} - h_{ts} \right\|_{2+\delta}}{|t-s|^{1/p}} \leq \left(\sup_{0 \leq s < t \leq T} \frac{\left\| h_{ts}^{\mathcal{D}} - h_{ts} \right\|_{2+\delta}}{|t-s|^{1/2}} \right)^{2/p} \left(\sup_{0 \leq s < t \leq T} \left\| h_{ts}^{\mathcal{D}} - h_{ts} \right\|_{2+\delta} \right)^{1-2/p}$$

$$\leq \left(\sup_{0 \leq s < t \leq T} \frac{\left\| h_{ts}^{\mathcal{D}} \right\|_{2+\delta}}{|t-s|^{1/2}} + \sup_{0 \leq s < t \leq T} \frac{\| h_{ts} \|_{2+\delta}}{|t-s|^{1/2}} \right)^{2/p}$$

$$\times \left(\sup_{0 \leq s < t \leq T} \left\| h_{ts}^{\mathcal{D}} - h_{ts} \right\|_{2+\delta} \right)^{1-2/p}$$

and

$$\sup_{0 \leq s < t \leq T} \frac{\left\| \int_s^t h_{su}^{\mathcal{D};j}(\cdot)\, \partial_j h_{du}^{\mathcal{D};k}(\cdot) - \int_s^t h_{su}^{j}(\cdot)\, \partial_j h_{du}^{k}(\cdot) \right\|_{1+\delta}}{|t-s|^{2/p}}$$

$$\leq \left(\sup_{0 \leq s < t \leq T} \frac{\left\| \int_s^t h_{su}^{\mathcal{D};j}(\cdot)\, \partial_j h_{du}^{\mathcal{D};k}(\cdot) \right\|_{1+\delta}}{|t-s|} + \sup_{0 \leq s < t \leq T} \frac{\left\| \int_s^t h_{su}^{j}(\cdot)\, \partial_j h_{du}^{k}(\cdot) \right\|_{1+\delta}}{|t-s|} \right)^{2/p}$$

$$\times \left(\sup_{0 \leq s < t \leq T} \left\| \int_s^t h_{su}^{\mathcal{D};j}(\cdot)\, \partial_j h_{du}^{\mathcal{D};k}(\cdot) - \int_s^t h_{su}^{j}(\cdot)\, \partial_j h_{du}^{k}(\cdot) \right\|_{1+\delta} \right)^{1-2/p}.$$

Therefore, the claim follows if we can prove

$$\lim_{|\mathcal{D}| \to 0} \sup_{h\,:\,I(h) \leq \Lambda} \sup_{0 \leq s < t \leq T} \left\| h_{ts}^{\mathcal{D}} - h_{ts} \right\|_{2+\delta} = 0 \tag{4.3}$$

and

$$\lim_{|\mathcal{D}| \to 0} \sup_{h\,:\,I(h) \leq \Lambda} \sup_{0 \leq s < t \leq T} \left\| \int_s^t h_{su}^{\mathcal{D};j}(\cdot)\, \partial_j h_{du}^{\mathcal{D};k}(\cdot) - \int_s^t h_{su}^{j}(\cdot)\, \partial_j h_{du}^{k}(\cdot) \right\|_{1+\delta} = 0. \tag{4.4}$$

Concerning (4.3), note that

$$\sup_{0 \leq s < t \leq T} \left\| h_{ts}^{\mathcal{D}} - h_{ts} \right\|_{2+\delta} \leq 2 \sup_{t \in [0,T]} \left\| h_t^{\mathcal{D}} - h_t \right\|_{2+\delta}.$$

If $t \in [0,T]$ is fixed, using Lemma 4.1, we have

$$\left\| h_t^{\mathcal{D}} - h_t \right\|_{2+\delta} = \left\| (t - t_{\mathcal{D}}) \frac{h_{t^{\mathcal{D}}} - h_{t_{\mathcal{D}}}}{t^{\mathcal{D}} - t_{\mathcal{D}}} + h_{t_{\mathcal{D}}} - h_t \right\|_{2+\delta}$$

$$\leq 2 \sup_{|v-u| \leq |\mathcal{D}|} \| h_v - h_u \|_{2+\delta}$$

$$\leq 2\sqrt{|\mathcal{D}|} \sup_{|v-u| \leq |\mathcal{D}|} \frac{\| h_v - h_u \|_{2+\delta}}{|v-u|^{1/2}}$$

$$\leq 2\sigma_\gamma \sqrt{\langle h, h\rangle_{\mathcal{H}}} \, \sqrt{|\mathcal{D}|},$$

and (4.3) follows. Fix $x \in D$. By Young integration estimates (cf. e.g. [**FV10**, Theorem 6.8]),

$$\left| \int_s^t h_{su}^{\mathcal{D};j}(x) \, \partial_j h_{du}^{\mathcal{D};k}(x) - \int_s^t h_{su}^j(x) \, \partial_j h_{du}^k(x) \right|$$
$$\lesssim \left\| h^{\mathcal{D};j} - h^j \right\|_{2-\mathrm{var};[0,T]} \left\| \partial_j h^{\mathcal{D};k} \right\|_{1-\mathrm{var};[0,T]} + \left\| h^j \right\|_{1-\mathrm{var};[0,T]} \left\| \partial_j h^{\mathcal{D};k} - \partial_j h^k \right\|_{2-\mathrm{var};[0,T]}.$$

From interpolation for the p-variation [**FV10**, Proposition 5.5] and our former estimates,

$$\left\| h^{\mathcal{D};j} - h^j \right\|_{2-\mathrm{var};[0,T]}$$
$$\leq \left(\sup_{0 \leq u < v \leq T} \left\| h_{uv}^{\mathcal{D};j} - h_{uv}^j \right\|_{1+\delta} \right)^{1/2} \left(\left\| h^{\mathcal{D};j} \right\|_{1-\mathrm{var};[0,T]} + \left\| h^j \right\|_{1-\mathrm{var};[0,T]} \right)^{1/2}$$
$$\leq C\sigma_\gamma \sqrt{\langle h, h\rangle_{\mathcal{H}}} \, |\mathcal{D}|^{1/4}.$$

A similar estimate holds for $\|\partial_j h^{\mathcal{D};k} - \partial_j h^k\|_{2-\mathrm{var};[0,T]}$ and we obtain

$$\left\| \int_s^t h_{su}^{\mathcal{D};j}(\cdot) \, \partial_j h_{du}^{\mathcal{D};k}(\cdot) - \int_s^t h_{su}^j(\cdot) \, \partial_j h_{du}^k(\cdot) \right\|_0 \leq C\sigma_\gamma^2 \langle h, h\rangle_{\mathcal{H}} \, |\mathcal{D}|^{1/4}.$$

As above, one can use the product rule and obtain similar estimate for the Hölder norm of the derivative. This shows that

$$\sup_{0 \leq s < t \leq T} \left\| \int_s^t h_{su}^{\mathcal{D};j}(\cdot) \, \partial_j h_{du}^{\mathcal{D};k}(\cdot) - \int_s^t h_{su}^j(\cdot) \, \partial_j h_{du}^k(\cdot) \right\|_{1+\delta} \leq C\sigma_\gamma^2 \langle h, h\rangle_{\mathcal{H}} \, |\mathcal{D}|^{1/4}$$

and (4.4) follows. $\qquad\square$

THEOREM 4.5 (Schilder's theorem for Wiener rough drivers). *Let D be a relatively compact domain and let X be a Wiener process in $\mathcal{C}_b^{2,0,\delta}(D, \mathbb{R}^d)$ with Cameron–Martin space $\mathcal{H}$. Denote by $\boldsymbol{X}$ its natural lift to a (p, ρ)-rough driver. For $\varepsilon > 0$, set $\boldsymbol{P}_\varepsilon :=$ $\mathbb{P} \circ (\delta_\varepsilon \boldsymbol{X})^{-1}$. Then the family $\{\boldsymbol{P}_\varepsilon \,:\, \varepsilon > 0\}$ of probability measures satisfies a large deviation principle on the space of rough drivers with speed ε^{-2} and good rate function*

$$J(\boldsymbol{v}) = \begin{cases} \dfrac{1}{2}\langle v, v\rangle & \text{if } \boldsymbol{v} = (v, \mathbb{v}) \text{ and } v \in \mathcal{H} \\ +\infty & \text{otherwise.} \end{cases}$$

PROOF. The proof is standard, using the large deviation principle for Gaussian measure [**DS89**, Section 3.4], the extended contraction principle [**DZ98**, Theorem 4.2.23] and the results in Lemmata 4.3 and 4.4 (cf. e.g. [**FV10**, Theorem 13.42]). $\qquad\square$

As an immediate corollary, we obtain Freidlin–Ventzel large devations for a class of stochastic flows.

THEOREM 4.6.   *Let X be a Wiener process in $\mathcal{C}_b^{2,0,\delta}(D, \mathbb{R}^d)$, for some $\delta \in (2/3, 1]$, and let φ^ε be the flow generated by the Stratonovich solution to*

$$d\varphi^\varepsilon = \varepsilon X(\varphi^\varepsilon \, ; \circ dt).$$

Let ν^ε denote the law of φ^ε in the space of $\mathcal{C}^\rho$-homeomorphisms, $\rho \in (2/3, \delta)$. Then the family $\{\nu^\varepsilon \, : \, \varepsilon > 0\}$ of probability measures satisfies a large deviation principle with speed ε^{-2} and good rate function

$$L(\psi) = \inf \left\{ J(\boldsymbol{v}) \, : \, d\psi = \boldsymbol{v}(\psi \, ; dt) \right\}.$$

PROOF.   The Stratonovich solution equals the solution generated by the (p, ρ)-rough driver $\boldsymbol{X}$. Using Theorem 4.5 and the pathwise continuity $\boldsymbol{X} \mapsto \varphi$, we can use the usual contraction principle in large deviation theory [**DZ98**, Theorem 4.2.1] to conclude. $\square$

5. Appendix.

We provide in this appendix an elementary regularity result for integrals depending on a parameter.

Let δ_ε be a standard Dirac sequence. If I is a closed interval and $f \colon I \to \mathbb{R}$ is a continuous function, let $\bar{f} \colon \mathbb{R} \to \mathbb{R}$ denote the unique continuous extension which coincides with f on I and which is constant outside this interval. Set $f^\varepsilon := \delta_\varepsilon * \bar{f}$. If D is some subset of $\mathbb{R}^d$ and if $f \colon D \times I \to \mathbb{R}$ is a continuous function in time for every $x \in D$, set

$$f^\varepsilon(x, t) := \big(\delta_\varepsilon * \bar{f}(x, \cdot)\big)(t).$$

PROPOSITION 5.1.   *Let $D \subset \mathbb{R}^d$ be an open set and let $f \colon D \times [0, T] \to \mathbb{R}$ and $g \colon D \times [0, T] \to \mathbb{R}$ be continuous. Assume that f and g are continuously differentiable on D and that $f(x, 0) = \partial_i f(x, 0) = 0$ for every $x \in D$ and every $i = 1, \ldots, d$. Moreover, assume that there are $p, q \in [1, \infty)$ with $1/p + 1/q > 1$ such that*

$$\sup_{(t_i)} \sum_{t_i} \big\| f(\cdot, t_{i+1}) - f(\cdot, t_i) \big\|_{\mathcal{C}^1}^p \quad and \quad \sup_{(t_i)} \sum_{t_i} \big\| g(\cdot, t_{i+1}) - g(\cdot, t_i) \big\|_{\mathcal{C}^1}^q$$

*are finite, where the suprema are taken over all finite partitions of the interval $[0, T]$. Then the Young integral (cf. e.g. [**FV10**, Chapter 6] for the precise definition) $\int_0^T f(x, t) \, g(x, dt)$ exists, is continuously differentiable for all $x \in D$ and the derivative is given by*

$$\partial_i \left(\int_0^T f(x, t) \, g(x, dt) \right) = \int_0^T \partial_i f(x, t) \, g(x, dt) + \int_0^T f(x, t) \, \partial_i g(x, dt)$$

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

PROOF.   Fix $i \in \{1, \ldots, d\}$. For any $x \in D$ and $\varepsilon > 0$,

$$\partial_i \left(\int_0^T f(x,t) \, g^\varepsilon(x, dt) \right) = \partial_i \left(\int_0^T f(x,t) \partial_t(g^\varepsilon(x,t)) \, dt \right)$$

$$= \int_0^T \partial_i f(x,t) \partial_t(g^\varepsilon(x,t)) \, dt + \int_0^T f(x,t) \partial_t((\partial_i g(x,\cdot))^\varepsilon(t)) \, dt$$

$$= \int_0^T \partial_i f(x,t) \, g^\varepsilon(x, dt) + \int_0^T f(x,t)(\partial_i g(x,\cdot))^\varepsilon(dt).$$

Let $q' > q$ such that $1/p + 1/q' > 1$. Fix $x \in D$, let U be some neighbourhood of x and take $y \in U$. From Young estimates and interpolation, we obtain

$$\left| \int_0^T \partial_i f(y,t) \, g^\varepsilon(y, dt) - \int_0^T \partial_i f(y,t) \, g(y, dt) \right|$$

$$\leq C \sup_{(t_i) \subset [0,T]} \left(\sum_{t_i} \left| \partial_i f(y, t_{i+1}) - \partial_i f(y, t_i) \right|^p \right)^{1/p}$$

$$\times \sup_{(t_i) \subset [0,T]} \left(\sum_{t_i} \left| g^\varepsilon(y, t_{i+1}) - g(y, t_{i+1}) - g^\varepsilon(y, t_i) + g(y, t_i) \right|^{q'} \right)^{1/q'}$$

$$\leq C \sup_{(t_i) \subset [0,T]} \left(\sum_{t_i} \left\| f(\cdot, t_{i+1}) - f(\cdot, t_i) \right\|_{\mathcal{C}^1}^p \right)^{1/p}$$

$$\times \left\{ \sup_{(t_i) \subset [0,T]} \left(\sum_{t_i} \left\| g^\varepsilon(\cdot, t_{i+1}) - g^\varepsilon(\cdot, t_i) \right\|_{\mathcal{C}^1}^q \right)^{1/q} \right.$$

$$\left. + \sup_{(t_i) \subset [0,T]} \left(\sum_{t_i} \left\| g(\cdot, t_{i+1}) - g(\cdot, t_i) \right\|_{\mathcal{C}^1}^q \right)^{1/q} \right\}^{q/q'}$$

$$\times 2^{1 - q/q'} \sup_{0 \leq t \leq T} \left\| g^\varepsilon(\cdot, t) - g(\cdot, t) \right\|_{\mathcal{C}^0}^{1 - q/q'}.$$

It is easy to check that

$$\sup_{(t_i) \subset [0,T]} \left(\sum_{t_i} \left\| g^\varepsilon(\cdot, t_{i+1}) - g^\varepsilon(\cdot, t_i) \right\|_{\mathcal{C}^1}^q \right)^{1/q} \leq \sup_{(t_i) \subset [0,T]} \left(\sum_{t_i} \left\| g(\cdot, t_{i+1}) - g(\cdot, t_i) \right\|_{\mathcal{C}^1}^q \right)^{1/q}.$$

Therefore, we obtain a bound of the form

$$\left| \int_0^T \partial_i f(y,t) \, g^\varepsilon(y, dt) - \int_0^T \partial_i f(y,t) \, g(y, dt) \right| \leq C \sup_{0 \leq t \leq T} \left\| g^\varepsilon(\cdot, t) - g(\cdot, t) \right\|_{\mathcal{C}^0}^{1 - \beta'/\beta}$$

where C is independent of y and ε. Thus,

$$\int_0^T \partial_i f(y,t) \, g^\varepsilon(y, dt) \to \int_0^T \partial_i f(y,t) \, g(y, dt)$$

uniformly in a neighbourhood around x when $\varepsilon \to 0$. Similarly,

$$\int_0^T f(y,t)(\partial_i g(y,\cdot))^\varepsilon(dt) \to \int_0^T f(y,t)\partial_i g(y,dt)$$

uniformly in a neighbourhood around x when $\varepsilon \to 0$. This shows differentiability in x of the integral

$$\int_0^T f(x,t)\,g(x,dt)$$

and the claimed identity. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

References

[Bai15]   I. Bailleul, Flows driven by rough paths, Rev. Mat. Iberoamericana, 2015.

[Bai18]   I. Bailleul, On the definition of a solution to a rough differential equation, arXiv: 1803.06479, 2018.

[Bax80]   P. Baxendale, Wiener processes on manifolds of maps, Proc. Royal Soc. Edinburgh, **87** (1980), 127–152.

[BC17]   I. Bailleul and R. Catellier, Rough flows and homogenization in stochastic turbulence, J. Differential Equations, **263** (2017), 4894–4928.

[BDM08]   A. Budhiraja, P. Dupuis and V. Maroulas, Large deviations for infinite dimensional stochastic dynamical systems, Ann. Probab., **36** (2008), 1390–1420.

[BDM10]   A. Budhiraja, P. Dupuis and V. Maroulas, Large deviations for stochastic flows of diffeomorphisms, Bernoulli, **16** (2010), 234–257.

[BF61]   Y. N. Blagovescenskii and M. Freidlin, Some properties of diffusion processes depending on a parameter, Soviet Math., **2** (1961), 633–636.

[BG17]   I. Bailleul and M. Gubinelli, Unbounded rough drivers, Ann. Fac. Sci. Toulouse Math. (6), **26** (2017), 795–830.

[BGN04]   W. Bertram, H. Glöckner and K.-H. Neeb, Differential calculus over general base fields and rings, Expo. Math., **22** (2004), 213–282.

[Bis81]   J.-M. Bismut, Mécanique aléatoire, Lecture Notes in Math., **866**, Springer-Verlag, Berlin-New York, 1981.

[Bog98]   V. I. Bogachev, Gaussian measures, Mathematical Surveys and Monographs, **62**, Amer. Math. Soc., Providence, RI, 1998.

[BRS17]   I. Bailleul, S. Riedel and M. Scheutzow, Random dynamical systems, rough paths and rough flows, J. Differential Equations, **262** (2017), 5792–5823.

[CK91]   E. Carlen and P. Krée, L^p estimates on iterated stochastic integrals, Ann. Probab., **19** (1991), 354–368.

[CW17]   T. Cass and M. Weidner, Tree algebras over topological vector spaces in rough path theory, arXiv:1604.07352v2, 2017.

[Dav07]   A. M. Davie, Differential equations driven by rough paths: an approach via discrete approximation, Appl. Math. Res. Express, AMRX, Art. ID abm009, 2007.

[DD12]   S. Dereich and G. Dimitroff, A support theorem and a large deviation principle for Kunita flows, Stoch. Dyn., **12** (2012), no. 3, 1150022.

[Der10]   S. Dereich, Rough paths analysis of general Banach space-valued Wiener processes, J. Funct. Anal., **258** (2010), 2910–2936.

[DS89]   J.-D. Deuschel and D. W. Stroock, Large deviations, Pure and Applied Mathematics, **137**, Academic Press, Inc., Boston, MA, 1989.

[DZ98]   A. Dembo and O. Zeitouni, Large deviations techniques and applications, Applications of Mathematics, **38**, Springer-Verlag, New York, 1998.

[Elw78]   K. D. Elworthy, Stochastic dynamical systems and their flows, In: Stochastic analysis,

Proc. Internat. Conf., Northwestern Univ., Evanston, Ill., 1978, Academic Press, New York-London, 1978, 79–95.

[Elw82]   K. D. Elworthy, Stochastic differential equations on manifolds, London Mathematical Society Lecture Notes Series, Cambridge Univ. Press, Cambridge-New York, 1982.

[FdLP06]   D. Feyel and A. de La Pradelle, Curvilinear integrals along enriched paths, Electron. J. Probab., **11** (2006), 860–892.

[FdLPM08]   D. Feyel, A. de La Pradelle and G. Mokoboski, A non-commutative sewing lemma, Elec. Comm. Probab., **13** (2008), 24–34.

[FH14]   P. K. Friz and M. Hairer, A Course on Rough Paths with an introduction to regularity structures, Universitext, **XIV**, Springer, Cham, 2014.

[FV07]   P. Friz and N. Victoir, Large deviation principle for enhanced Gaussian processes, Ann. Inst. H. Poincaré Probab. Statist., **43** (2007), 775–785.

[FV10]   P. K. Friz and N. B. Victoir, Multidimensional stochastic processes as rough paths, Theory and applications, Cambridge Studies in Advanced Mathematics, **120**, Cambridge University Press, Cambridge, 2010.

[Gub04]   M. Gubinelli, Controlling rough paths, J. Funct. Anal., **216** (2004), 86–140.

[Har81]   T. Harris, Brownian motions on the homeomorphisms of the plane, Ann. Prob., **9** (1981), 232–254.

[IA72]   I. Gihman and A. Skorohod, Stochastic differential equations, Springer, Berlin, 1972.

[IW81]   N. Ikeda and S. Watanabe, Stochastic differential equations and diffusion processes, North-Holland Mathematical Library, **24**, North-Holland Publishing Co., Amsterdam-New York, Kodansha, Ltd., Tokyo, 1981.

[Kal02]   O. Kallenberg, Foundations of modern probability, Probability and its Applications, Springer-Verlag, New York, 2002.

[Kun81]   H. Kunita, On the decomposition of solutions of stochastic differential equations, In: Stochastic integrals, Proc. Sympos., Univ. Durham, Durham, 1980, Lecture Notes in Math., **851**, Springer, Berlin-New York, 1981, 213–255.

[Kun86a]   H. Kunita, Convergence of stochastic flows connectd with stochastic ordinary differential equations, Stochastics, **17** (1986), 215–251.

[Kun86b]   H. Kunita, Lectures on stochastic flows and applications, Tata Institue of Fundamental Research, Bombay, 1986.

[Kun90]   H. Kunita, Stochastic flows and stochastic differential equations, Cambridge Studies in Advanced Mathematics, **24**, Cambridge University Press, Cambridge, 1990.

[Led96]   M. Ledoux, Isoperimetry and Gaussian analysis, In: Lectures on probability theory and statistics, Saint-Flour, 1994, Lecture Notes in Math., **1648** (1996), 165–294.

[LJ82]   Y. Le Jan, Flots de diffusion dans $\mathbf{R}^d$, C. R. Acad. Sci. Paris Sér. I Math., **294** (1982), 697–699.

[LJ85]   Y. Le Jan, On isotropic Brownian motions, Z. Wahrsch. Verw. Gebiete, **70** (1985), 609–620.

[LJW84]   Y. Le Jan and S. Watanabe, Stochastic flows of diffeomorphisms, In: Stochastic analysis, Katata/Kyoto, 1982, North-Holland Math. Library, **32**, North-Holland, Amsterdam, 1984, 307–332.

[LLQ02]   M. Ledoux, T. Lyons and Z. Qian, Lévy area of Wiener processes in Banach spaces, Ann. Probab., **30** (2002), 546–578.

[LQZ02]   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.

[Lyo98]   T. J. Lyons, Differential equations driven by rough signals, Rev. Mat. Iberoamericana, **14** (1998), 215–310.

[MSS06]   A. Millet and M. Sanz-Solé, Large deviations for rough paths of the fractional Brownian motion, Ann. Inst. H. Poincaré Probab. Statist., **42** (2006), 245–271.

[Rya02]   R. A. Ryan, Introduction to tensor products of Banach spaces, Springer Monographs in Math., Springer-Verlag London, Ltd., London, 2002.

[Sch09]   M. Scheutzow, Chaining techniques and their application to stochastic flows, In: Trends in stochastic analysis, London Math. Soc. Lecture Note Ser., **353**, Cambridge Univ. Press, Cambridge, 2009, 35–63.

Ismaël Bailleul
IRMAR
263 Avenue du General Leclerc
35042 Rennes, France
E-mail: ismael.bailleul@univ-rennes1.fr

Sebastian Riedel
Institut für Mathematik
Technische Universität Berlin
Germany
E-mail: riedel@math.tu-berlin.de