Rev. Mat. Iberoam. 38 (2022), no. 2, 433–502
DOI 10.4171/RMI/1280

Noiseless regularisation by noise

Lucio Galeati and Massimiliano Gubinelli

Abstract. We analyse the effect of a *generic* continuous additive perturbation to the well-posedness of ordinary differential equations. Genericity here is understood in the sense of prevalence. This allows us to discuss these problems in a setting where we do not have to commit ourselves to any restrictive assumption on the statistical properties of the perturbation. The main result is that a generic continuous perturbation renders the Cauchy problem well-posed for arbitrarily irregular vector fields. Therefore we establish regularisation by noise "without probability".

1. Introduction

From the modelling point of view, the presence of external perturbations to otherwise autonomous evolutions is a very natural assumption. Let $d \in \mathbb{N}$ and consider the following ODE in $\mathbb{R}^d$:

$$\begin{cases} \dot{x}(t) = b(t, x(t)) + \dot{w}(t), \\ x(0) = x_0 \in \mathbb{R}^d, \end{cases} \qquad t \geqslant 0, \tag{1.1}$$

where $w \in C(\mathbb{R}_+, \mathbb{R}^d)$ is a fixed perturbation, the dot denotes differentiation with respect to time and $b\colon \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d$ is a time-dependent vector field. Provided eq. (1.1) is understood as an integral equation, and thanks to the additive nature of the perturbation, there are no particular regularity requirements, apart from continuity, which have to be imposed on the function w. A natural question is then for which classes of vector fields b eq. (1.1) is well-posed and if, for certain sets of perturbations w, one can obtain well-posedness results in classes which are known to lead to an ill-posed problem when $w = 0$.

One possible approach to this problem is to consider w a sample path of a stochastic process W defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Indeed, in recent years there has been a lot of activity in understanding the possible role of random perturbations to improve the well-posedness of ordinary (or partial) differential equations (ODE/PDE) (see [21] for a recent review). This approach has, however, certain limitations:

a) It requires to make very specific assumptions on the kind of randomness which is allowed in any specific problem.

2020 Mathematics Subject Classification: Primary 60H50; Secondary 37C20.
Keywords: Regularisation by noise, prevalence.

b) It introduces into the picture considerations which are not quite germane to the initial formulation. For example measurability (or adaptedness) with respect to Ω of solutions as soon as we need to look at them in the sense of stochastic processes (i.e., seen as random variables) and weaker notions of uniqueness which are not easy to compare to the deterministic setting.

With respect to point a) one can use other assumptions to justify specific choices. Within the class of time-dependent continuous random processes, for example, Brownian motion has suitable features of universality and Markovianity, making it a natural choice. Furthermore, a large set of theoretical tools is available to analyse the effect of Brownian perturbations to deterministic evolutions and this topic has a long and extensive literature [8, 16, 17, 23, 37, 58, 59]. Other classes of random perturbations, like fractional Brownian motion (fBm) have been more recently analysed, or even more exotic variants (e.g., α-stable and log regular processes) [1, 6, 13, 38, 45, 49]. Let us finally mention the remarkable results from [10] concerning rates of convergence of numerical schemes for (1.1).

As for the technical limitations in point b), a possible solution is to modify the probabilistic setting in order to derive path-wise statements:

(i) Davie and Flandoli introduced a stronger concept of uniqueness called *path-by-path uniqueness* [18, 21, 22] in the Brownian setting; see also [54] regarding its difference from pathwise uniqueness.

(ii) Catellier and one of the authors [12] studied almost sure regularisation properties of fractional Brownian motion (fBm) and applied them to show strong well-posedness results for (1.1) when w is a sample path of an fBm.

In this work we take a conceptually different approach and consider the regularisation by noise problem from the point of view of *generic* perturbations, in particular without reference to any (specific) probabilistic setting.

We will say that a property holds for *almost every* path w if it holds for a *prevalent* set of paths. Prevalence [46] is a notion of "Lebesgue measure zero sets" in infinite dimensional complete metric vector spaces. Such sets cannot be naively defined due to the fact that there cannot exist σ-additive, translation invariant measures in infinite dimensional spaces. It was first introduced by Christensen in [14] in the context of abelian Polish groups and later rediscovered independently by Hunt, Sauer and Yorke in [35] for complete metric vector spaces.

Prevalence has been used in different contexts in order to study the properties of generic functions belonging to spaces of suitable regularity. For instance, it was proved in [34] that almost every continuous function is nowhere differentiable, while in [26, 27] the multi-fractal nature of generic Sobolev functions was shown. Recently, prevalence has also attracted a lot of attention in the study of dimension of graphs and images of continuous functions, see among others [7, 25].

A key advantage of prevalence, with respect to other notions of genericity, is that it allows the use of probabilistic methods in the proof. However, the statements are fully non-probabilistic and the kind of problems one encounters in formulating prevalence results are quite distinct from those of a purely probabilistic setting, extensively investigated in the probabilistic literature.

Armed with this "noiseless" notion of "almost every path", we can already state informally one of the results of the paper as follows:

> *Let $b \in C([0,T]; H^{-s}(\mathbb{R}^d))$ be fixed, $s > 0$ arbitrarily large. Then* almost every *perturbation $w \in C([0,T];\mathbb{R}^d)$ has* infinite *regularisation effect on the ODE associated to b, namely it renders the ODE (1.1) well-posed and with a smooth flow.*

In order to proceed and precise the above claims, we will need a suitable notion of solution to (1.1) which makes sense for distributional fields b. The key observation in this direction comes from the work [12], which started the study of analytic properties of paths which affects the regularisation of ODEs. It introduces an *averaging operator T* as a tool to study the regularisation properties of a path w. This is the operator acting on time-dependent vector fields b and paths w as as

$$(w,b) \mapsto (T^w b)(t,x) = \int_0^t b(s, x + w(s))\,\mathrm{d}s, \quad x \in \mathbb{R}^d, t \geqslant 0.$$

It is a linear operator in b, so that one can fix w and consider the operator $T^w : b \mapsto T^w b$ as above; in this case we say that T^w is the averaging operator associated to w. Alternatively, one can fix b and vary w, $w \mapsto T^w b$; to stress the latter case, we say that $T^w b$ is an *averaged field*.

Averaging is connected to an alternative formulation of the ODE via the theory of non-linear Young integration. Assume for the moment $b: [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ smooth and consider the ODE (1.1) in integral form

$$x_t = x_0 + \int_0^t b(s, x_s)\,\mathrm{d}s + w_t, \quad t \in [0,T], \tag{1.2}$$

with $w \in C([0,T];\mathbb{R}^d)$. Then, this equation admits a unique solution of the form $x = w + C^1$, in the sense that the difference $x - w$ is a C^1 path, regardless the regularity of w. Applying the change of variables $\theta := x - w$, we get the new integral equation

$$\theta_t = \theta_0 + \int_0^t b(s, \theta_s + w_s)\,\mathrm{d}s, \quad t \in [0,T]. \tag{1.3}$$

Since both b and θ are continuous, the last integral can be approximated via Riemann–Stieltjes type sums as follows:

$$\begin{aligned} \int_0^t b(s, \theta_s + w_s)\,\mathrm{d}s &= \lim_{|\Pi| \to 0} \sum_i \int_{t_i}^{t_{i+1}} b(s, \theta_{t_i} + w_s)\,\mathrm{d}s \\ &= \lim_{|\Pi| \to 0} \sum_i T^w b_{t_i,t_{i+1}}(\theta_{t_i}), \end{aligned} \tag{1.4}$$

where the limit taken over all possible partitions $\Pi = \{t_0, \dots, t_n\}$ with $0 = t_0 < t_1 < \cdots < t_n = t$ with mesh $|\Pi| = \sup_i |t_{i+1} - t_i|$ converging to 0 and where for a function $A = A(t,x)$ we adopt the compact notation $A_{s,t}(x) := A(t,x) - A(s,x)$. The right-hand side of equation (1.4) depends now on the averaged field $T^w b$. The key observation of [12] is that, under suitable space-time regularity conditions on $T^w b$, it is possible to show convergence of the above Riemann–Stieltjes type sums to a unique limit even when b is not continuous anymore, thus allowing to *define* the integral on the left-hand side of (1.4) as their limit. This limit is called in [12] a *non-linear Young integral* and denoted as

$$\int_0^t T^w b(\mathrm{d}s, \theta_s).$$

Eq. (1.3) takes then the form of an integral equation involving non-linear Young integrals:

$$\theta_t = \theta_0 + \int_0^t T^w b(\mathrm{d}s, \theta_s).$$

The analysis of such equations (existence, uniqueness, regularity of the flow) for irregular b depends essentially on the regularity properties of the averaged field $T^w b$ and a substantial part of the present paper will be dedicated to analyse them in detail. For example we will prove that:

> *Let $b \in C([0,T]; H^{-s}(\mathbb{R}^d))$ be fixed, $s > 0$ arbitrarily large. Then* almost every *perturbation $w \in C([0,T]; \mathbb{R}^d)$ has* infinite *regularisation effect on b, namely $T^w b \in C([0,T]; C^\infty)$.*

A quantitative version of the statements above, which collects some of the main results of this paper, is the following one.

Theorem 1.1. *Let $b \in B^\alpha_{\infty,\infty}$ be a compactly supported distribution, and let $\alpha \in (-\infty, 1)$ and $\delta \in (0,1)$.*

- (i) *If $\delta < (2-2\alpha)^{-1}$, then for a.e. $w \in C^\delta_t$ it holds that $T^w b \in C^\gamma_t C^1_x$, and the ODE (1.2) has a meaningful interpretation; moreover, for any initial $x_0 \in \mathbb{R}^d$ there exists a solution to the ODE.*
- (ii) *If $\delta < (2-2\alpha)^{-1}$ and we fix $x_0 \in \mathbb{R}^d$, then for a.e. $w \in C^\delta_t$ there exists a unique solution to the ODE with initial condition x_0.*
- (iii) *If $\delta < (4-2\alpha)^{-1}$, then for a.e. $w \in C^\delta_t$ the ODE is well posed and it admits a locally C^1 flow.*
- (iv) *If $\delta < (2n-2\alpha)^{-1}$, then for a.e. $w \in C^\delta_t$ the flow is locally C^{n-1}.*
- (v) *Finally, for a.e. $w \in C^0$ the ODE admits a smooth flow.*

Remark 1.2. In this theorem we could allow time dependent $b \in L^\infty_t B^\alpha_{\infty,\infty}$ provided $\delta < 1/2$. This is due to some technical limitations in the proof technique.

Let us point out that this results is the first general statement which supports the heuristics "the rougher the noise, the better the regularisation" observed in the probabilistic literature since e.g., [12] but so far never discussed abstracting from a particular probabilistic model of the perturbation.

We conclude this introduction by discussing possible extensions are relations with related work. The averaging operator T^w is, in many respect, a key tool introduced in [12] to study analytically the regularisation properties of perturbations in dynamical problems. In this paper we refrain to investigate more thoroughly this operator from the point of view of prevalence since this will be the main objective of the companion paper [29]. There we continue the study of the prevalent properties of path which are associated to the regularisation by noise phenomenon by concentrating on the notion of ρ-irregularity of a path, as introduced in [12], and the related notion of occupation measure, obtaining as a by-product information on the prevalent properties of T^w.

The setting we propose in this paper opens up a completely new research subject with many natural problems; one prominent among them is to investigate the *zero noise limit*, that is, the limit as $\varepsilon \to 0$ for solutions to the equation $\dot{x}_\varepsilon = b(x_\varepsilon) + \varepsilon w$. Already in the probabilistic setting this limit is not well understood, especially from the path-wise perspective, and the dependence of the limit on the law assumed for w is not clear.

On a more technical level, several improvement of our results could be possible. For example, it would be interesting to obtain estimates for the averaging in L^p-based spaces with $p \in [1, 2)$, see Remark 3.19 below, and the related discussion in Appendix A.3. In particular let us note that the natural Conjecture 1.2 from [12] is still partially open; after the first draft of this work appeared, Nicolas Perkowski presented us a proof that answers negatively the conjecture in the case $H < 1/d$ for general $d \geqslant 3$ and $H \leqslant 1/2$ for $d = 2$.

While we were finalizing the present paper, two related preprints appeared. Harang and Perkowski [32] study the flow of the ODE (1.1) perturbed with a Gaussian process very similar to that considered in [1] but from the pathwise point of view of [12]. Along the way they give proofs of some results on the flow of Young differential equations alternative to those we give below. In [2] Amine, Mansouri and Proske study with techniques very different from ours, the path-by-path uniqueness for transport equations driven by fBm with Hurst index $H < 1/2$ and with bounded vector-fields. It is to be noted that while both works obtain interesting results, they still consider very specific probabilistic models. Therefore they are both far from the novel point of view we propose here and in the companion paper [29] and from the specific results it generates.

Let us finally mention the very recent work [31] in which Gerencser provides instances of regularisation by noise for $w \in C^\delta$ with $\delta > 1$.

Structure of the paper. We start by introducing the concept prevalence and its basic properties. Section 3 is devoted to the study of prevalence statements for averaged fields. Fractional Brownian motion (fBm) enters into the picture as a suitable transverse measure for prevalence. Thanks to a functional Itô–Tanaka type formula, we deduce regularity estimates for distributions averaged by fBm, which are strong enough to lead to prevalence statements. Section 4 is devoted to the application of the results from the previous section to perturbed ODEs via the theory of nonlinear Young integrals. After recalling and expanding the results from [12], we provide conditions (in terms of the regularity of $T^w b$) under the ODE admits a flow with prescribed regularity. Combined with Section 3, this allows to prove Theorem 1.1. Finally, we consider the case of perturbed transport type PDEs, for which it is again possible to establish well-posedness under suitable regularity conditions on $T^w b$. We choose to put in the Appendix reminders of standard facts and certain technical results.

Notation. We will use the notation $a \lesssim b$ to mean that there exists a positive constant c such that $a \leqslant cb$; we use the index $a \lesssim_x b$ to highlight the dependence $c = c(x)$. $a \sim b$ if and only if $a \lesssim b$ and $b \lesssim a$, similarly for $a \sim_x b$.

We will always work on a finite time interval $[0, T]$ unless stated otherwise. Whenever useful, we adopt the convention that f_t stands for $f(t)$ for a function f indexed on $t \in [0, T]$, but depending on the context we will use both notations; similarly for the increments of $f_{s,t} = f_t - f_s$.

For $x \in \mathbb{R}^d$, $|x|$ denotes the Euclidean norm, and $x \cdot y$ the scalar product. For any $R > 0$, B_R stands for $B(0, R) = \{x \in \mathbb{R}^d : |x| \leqslant R\}$.

We denote by $\mathcal{S}(\mathbb{R}^d; \mathbb{R}^m)$ and $\mathcal{S}'(\mathbb{R}^d; \mathbb{R}^m)$, respectively, the spaces of vector-valued Schwarz functions and tempered distributions on $\mathbb{R}^d$; similarly, $C_c^\infty(\mathbb{R}^d; \mathbb{R}^m)$ is the set of vector-valued smooth compactly supported functions.

Given a separable Banach space E, we denote by $L^q(0, T; E) = L_t^q E$ the Bochner–Lebesgue space of E-valued measurable functions $f: [0, T] \to E$ such that

$$\|f\|_{L^q(0,T;E)} = \Big(\int_0^T \|f_t\|_E^q \,\mathrm{d}t\Big)^{1/q} < \infty,$$

with the essential supremum in the limit case $q = \infty$; and $C^\alpha([0, T]; E) = C_t^\alpha E$ is the space of E-valued α-Hölder continuous functions, for $\alpha \in (0, 1)$, i.e., $f: [0, T] \to E$ such that

$$\|f\|_{C^\alpha E} := \|f\|_{C^0 E} + [\![f]\!]_{C^\alpha E} = \sup_{t\in[0,T]} \|f_t\|_E + \sup_{s\neq t\in[0,T]} \frac{\|f_{s,t}\|_E}{|t-s|^\alpha} < \infty.$$

A similar definition holds for $\mathrm{Lip}([0, T]; E) = \mathrm{Lip}_t\, E$. More generally, for a given modulus of continuity ω (possibly defined only in a neighbourhood of 0), we denote by $C^\omega([0, T]; E) = C^\omega E$ the set of all E-valued continuous functions with modulus of continuity ω, $\|f\|_{C^\omega E}$ and $[\![f]\!]_{C^\omega E}$ defined as above.

Whenever $E = \mathbb{R}^d$, we will refer to $w \in C_t^\alpha = C^\alpha([0, T]; \mathbb{R}^d)$ as a *path* and in this case we allow $\alpha \in [0, \infty)$ with the convention that $w \in C_t^\alpha$ it is has continuous derivatives up to order $\lfloor\alpha\rfloor$ and $D^{\lfloor\alpha\rfloor}\varphi$ is $\{\alpha\}$-Hölder continuous, where $\lfloor\alpha\rfloor$ and $\{\alpha\}$ denote respectively integer and fractional part.

We denote by $B_{p,q}^s(\mathbb{R}^d; \mathbb{R}^m)$, $L^{s,p}(\mathbb{R}^d; \mathbb{R}^m)$ and $F_{p,q}^s(\mathbb{R}^d; \mathbb{R}^m)$, respectively, the vector-valued Besov, Bessel and Triebel–Lizorkin spaces (see Appendix A.2), and by $L^p(\mathbb{R}^d; \mathbb{R}^m)$, the standard Lebesgue spaces. Whenever it does not create confusion, we will just write $B_{p,q}^s$, $L^{s,p}$, $F_{p,q}^s$ and L^p for short. For $\alpha \in \mathbb{R} \setminus \mathbb{N}_0$, $C^\alpha(\mathbb{R}^d; \mathbb{R}^m) = C_x^\alpha = B_{\infty,\infty}^\alpha$; instead for $\alpha \in \mathbb{N}_0$, $C^n(\mathbb{R}^d; \mathbb{R}^m) = C_x^n$ denotes the Banach space of all continuous functions with continuous derivatives up to order n, endowed with the norm

$$\|f\|_{C^\alpha} = \sup_{x\in\mathbb{R}^d} |f(x)| + \sum_{\beta\in\mathbb{N}_0^n : |\beta|=n} \sup_{x\in\mathbb{R}^d} |D^\beta f(x)|.$$

Let us stress in particular that by saying that $f \in C_x^n$, we are implying that we have a uniform bound on the whole $\mathbb{R}^d$ for its derivatives. If instead we want to say that f has continuous derivatives up to order n, we will write $f \in C_{\mathrm{loc}}^n$. We will adopt short-hand notations of the form $L_t^q L_x^p = L^q(0, T; L^p(\mathbb{R}^d; \mathbb{R}^m))$, $C_t^\alpha C_x^\beta = C^\alpha([0, T]; C^\beta(\mathbb{R}^d; \mathbb{R}^m))$.

Whenever a certain stochastic process $X = (X_t)_{t\geqslant 0}$ is considered, even when it is not specified, we assume the existence of an abstract underlying filtered probability space

$(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geqslant 0}, \mathbb{P})$ such that $\mathcal{F}$ and $\mathcal{F}_t$ satisfy the usual assumptions and X_t is adapted to $\mathcal{F}_t$. If $\mathcal{F}_t$ is said to be the natural filtration generated by X, then it is tacitly implied that it is actually its right continuous, normal augmentation. We denote by $\mathbb{E}$ integration (equiv. expectation) with respect to the probability $\mathbb{P}$.

2. Prevalence

Here we follow the setting and the terminology given in [35] even if, for our purposes, we will be interested only in the case of a Banach space E.

Definition 2.1. Let E be a complete metric vector space. A Borel set $A \subset E$ is said to be *shy* if there exists a measure μ such that

(i) there exists a compact set $K \subset E$ such that $0 < \mu(K) < \infty$,

(ii) for every $v \in E$, $\mu(v + A) = 0$.

In this case, the measure μ is said to be transverse to A. More generally, a subset of E is shy if it is contained in a shy Borel set. The complement of a shy set is called a prevalent set.

Sometimes it is said more informally that the measure μ "witnesses" the prevalence of A^c.

It follows immediately from part (i) of the definition that, if needed, one can assume μ to be a compactly supported probability measure on E. If E is separable, then any probability measure on E is tight and therefore (i) is automatically satisfied.

The following properties hold for prevalence (all proofs can be found in [35]):

(1) If E is finite dimensional, then a set A is shy if and only if it has zero Lebesgue measure.

(2) If A is shy, then so is $v + A$ for any $v \in E$.

(3) Prevalent sets are dense.

(4) If $\dim(E) = +\infty$, then compact subsets of E are shy.

(5) Countable union of shy sets is shy; conversely, countable intersection of prevalent sets is prevalent.

From now, whenever we say that a statement holds for a.e. $v \in E$, we mean that the set of elements of E for which the statement holds is a prevalent set. Property (1) states that this convention is consistent with the finite dimensional case.

In the context of a function space E, it is natural to consider as probability measure the law induced by an E-valued stochastic process. Namely, given a stochastic process W defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, taking values in a separable Banach space E, in order to show that a property $\mathcal{P}$ holds for a.e. $f \in E$, it suffices to show that

$$\mathbb{P}\,(f + W \text{ satisfies property } \mathcal{P}) = 1, \quad \forall f \in E.$$

Clearly, we are assuming that the set $A = \{w \in E : w \text{ satisfies property } \mathcal{P}\}$ is Borel measurable and if E is not separable, then we need to require in addition that the law of W is tight, so as to satisfy point (i) of Definition 2.1.

As a consequence of properties (4) and (5), the set of all possible realisations of a probability measure on a separable Banach space is a shy set, as it is contained in a countable union of compact sets (this is true more in general for any tight measure on a Banach space). This highlights the difference between a statement of the form

"Property $\mathcal{P}$ holds for a.e. f"

and, for instance,

"Property $\mathcal{P}$ holds for all Brownian trajectories",

where this last statement corresponds to μ (Property $\mathcal{P}$ holds) $= 1$, μ being the Wiener measure on $C([0, 1])$. Indeed, the second statement does not provide any information regarding whether the property might be prevalent or not. Intuitively, the elements satisfying a prevalence statement are "many more" than just the realisations of the Wiener measure.

3. Averaging operators

We introduce in detail the averaging operator $(w, b) \mapsto T^w b$ and analyse its prevalent properties in various functional spaces. Fractional Brownian motion is used as a convenient tranverse measure to detect prevalent regularisation properties of paths.

3.1. Definition of averaging operator and basic properties

In this section we provide the definition of the averaging operator T^w for measurable $w\colon [0, T] \to \mathbb{R}^d$, together with some basic properties which will be fundamental for later sections and our first main prevalence result. Our definition is rather abstract and works for a general class of Banach spaces E, but keep in mind that for our purposes E will always be either a Bessel space $L^{s,p}$ or a Besov space $B^s_{p,q}$ with $p \in [2, \infty)$. Also, we consider for simplicity the scalar-valued case, i.e., $E \subseteq \mathcal{S}'(\mathbb{R}^d)$. Everything generalises immediately to the vector-valued case $\mathcal{S}'(\mathbb{R}^d; \mathbb{R}^m)$ reasoning component by component.

Assume that E is a separable Banach space that continuously embeds into $\mathcal{S}'(\mathbb{R}^d)$ (so that there is also a dual embedding $\mathcal{S}(\mathbb{R}^d) \hookrightarrow E^*$) such that translations $\tau^v\colon f \mapsto \tau^v f = f(\cdot + v)$ act continuously on it and leave the norm invariant: $\|\tau^v f\|_E = \|f\|_E$ for all $v \in \mathbb{R}^d$ and $f \in E$. Assume moreover that the map $v \mapsto \tau^v$ is continuous in the sense that if $v_n \to v$, then $\tau^{v_n} f \to \tau^v f$ for all $f \in E$.

Definition 3.1. Let $w\colon [0, T] \to \mathbb{R}^d$ be a measurable function, and let E be as above. Then we define the averaging operator T^w as the continuous linear map from $L^1(0, T; E)$ to $C^0([0, T]; E)$ given by

$$T^w_t b = T^w b(t) := \int_0^t \tau^{w_s} b_s \, \mathrm{d}s \quad \forall t \in [0, T].$$

We will refer to $T^w b$ as an averaged function to stress that b is fixed, while w might be varying.

The definition is meaningful, since by the continuity properties of $v \mapsto \tau^v$, the map $s \mapsto \tau^{w_s} b(s)$ is still measurable and by the invariance of $\|\cdot\|_E$ under translations, $\|b\|_{L^1 E} = \|\tau^{w_\cdot} b\|_{L^1 E}$. Continuity of $T^w b$ and the bound $\|T^w b\|_{C^0 E} \leqslant \|b\|_{L^1 E}$ follow from standard properties of the Bochner integral, as well as the linearity of the map $b \mapsto T^w b$. Similarly, it is easy to see that, in the case b enjoys higher integrability, T^w can also be defined as a linear bounded operator from $L^q_t E$ to $C^{1-1/q}_t E$. Furthermore, if w and $\tilde{w}$ are such that $w_t = \tilde{w}_t$ for Lebesgue-a.e. $t \in [0,T]$, then $T^w b$ and $T^{\tilde{w}} b$ coincide for all b, so that T^w can be defined for w in an equivalence class.

Lemma 3.2. *Let $w^n \to w$ in $L^1_t \mathbb{R}^d$ and let $b \in L^q_t E$. Then $T^{w^n} b \to T^w b$ in $C^{1-1/q}_t E$.*

Proof. We can assume in addition that $w^n_t \to w_t$ for Lebesgue-a.e. t; the general case follows from applying the reasoning to any possible subsequence that can be extracted from $\{T^{w_n} b\}_n$. Since $\tau^{w^n_t} b_t \to \tau^{w_t} b_t$ for Lebesgue-a.e. t and $\|\tau^{w^n_t} b_t - \tau^{w_t} b_t\|^q \lesssim \|b_t\|^q \in L^1$, it follows from dominated convergence that

$$\|T^{w^n} b - T^w b\|_{C^{1-1/q} E} \lesssim \int_0^T \|\tau^{w^n_t} b_t - \tau^{w_t} b_t\|^q \to 0, \quad \text{as } n \to \infty,$$

which gives the conclusion. ∎

The advantage of the above definition of T^w is that it is intrinsic and does not depend on any approximation procedure by mollifiers. However, a possibly more intuitive description of $T^w b$ can be given by duality. Recall that, in the sense of distributions, $(\tau^v)^* = \tau^{-v}$, so that for any $\varphi \in \mathcal{S}(\mathbb{R}^d) \hookrightarrow E^*$ it holds

$$\langle T^w_t b, \varphi \rangle = \int_0^t \langle b_s, \varphi(\cdot - w_s) \rangle \, \mathrm{d}s,$$

where the pairing is integrable since $|\langle b_s, \varphi(\cdot - w_s)\rangle| \lesssim_\varphi \|b_s\|_E$. The above relation holds for all $\varphi \in \mathcal{S}(\mathbb{R}^d)$ and therefore uniquely identifies $T^w b(t)$ as an element of $\mathcal{S}'(\mathbb{R}^d)$, for all $t \in [0,T]$. The advantage now is that the map $(t,x) \mapsto \varphi(x - w_t)$ can be regarded as an element of $L^\infty(0,T; \mathcal{S}(\mathbb{R}^d))$, to which standard operations on $\mathcal{S}(\mathbb{R}^d)$ such as differentiation and convolution can be applied.

Lemma 3.3. *Let w and b be as above. Then*:

(i) *Averaging and spatial differentiation commute, i.e., $\partial_i T^w b = T^w \partial_i b$ for all $i = 1, \dots, d$.*

(ii) *Averaging and spatial convolution commute, i.e., for any $K \in C^\infty_c(\mathbb{R}^d)$ it holds*

$$K * (T^w b) = T^w (K * b) = (T^w K) * b.$$

Proof. Both statements follow easily from the duality formulation. For any $\varphi \in \mathcal{S}(\mathbb{R}^d)$ and $t \in [0,T]$, it holds

$$\begin{aligned}\langle \partial_i T^w b(t), \varphi \rangle &= -\langle T^w b(t), \partial_i \varphi \rangle = -\int_0^t \langle b_r, \partial_i \varphi(\cdot - w_r) \rangle \, \mathrm{d}r \\ &= \int_0^t \langle \partial_i b_r, \varphi(\cdot - w_r) \rangle \, \mathrm{d}r = \langle (T^w \partial_i b)(t), \varphi \rangle.\end{aligned}$$

If $K \in C_c^\infty(\mathbb{R}^d)$, then denoting by $\tilde{K}$ its reflection, by duality it holds

$$\langle K * T^w b(t), \varphi\rangle = \langle T^w b(t), \tilde{K} * \varphi\rangle = \int_0^t \langle b_r, \tau^{-w_r}(\tilde{K} * \varphi)\rangle \,\mathrm{d}r = \int_0^t \langle b_r, \tilde{K} * (\tau^{-w_r}\varphi)\rangle \,\mathrm{d}r$$
$$= \int_0^t \langle K * b(r), \tau^{-w_r}\varphi\rangle \,\mathrm{d}r = \langle T^w(K * b)(t), \varphi\rangle.$$

A similar computation shows the other part of the identity. ■

Remark 3.4. Let us point out that if $w \in L^\infty$, then the averaging operator has finite speed of propagation and so behaves well under localisation. Indeed, if $b \in L_t^1 E$ is such that $\operatorname{supp} b_t \subset B_R$ for all $t \in [0, T]$, then $\operatorname{supp} T^w b(t) \subset B_{R+\|w\|_\infty}$ for all $t \in [0, T]$; and similarly, if b and $\tilde{b}$ are such that their restrictions to B_R coincide for all t, then $T^w b$ and $T^w \tilde{b}$ will still coincide on $B_{R-\|w\|_\infty}$.

In view of the applications in Section 4, our main goal is to establish conditions under which $T^w b \in C_t^\gamma F$, where $\gamma > 1/2$ and F is another Banach space which enjoys better regularity properties than the original space E: typically $F = C_x^\beta$ for suitable values of β. For this reason, we are going to assume from now on that $b \in L_t^q E$ for some $q > 2$. The idea behind this restriction is that sometimes averaging allows to trade off time regularity for space regularity (think of the analogy with parabolic regularity theory) and therefore in order to have $T^w b \in C_t^\gamma F$, knowing a priori only that $T^w b \in C_t^{1-1/q} E$, we need to require at least

$$1 - \frac{1}{q} > \gamma > \frac{1}{2} \quad \Rightarrow \quad q > 2.$$

Remark 3.5. Despite our use of the terminology "regularisation by averaging", what we mean is really that we *fix* a drift b and we want to establish that for a.e. path w the averaged function $T^w b$ has nice regularity properties. This is *different* from trying to establish that the averaging operator T^w as a linear operator from $L_t^q E$ to $C_t^\gamma F$ is bounded, which is clearly false due to the time dependence of the drifts we consider. Indeed, given any $b \in E$, defining $\tilde{b}_t = \tau^{-w_t} b$, by definition of averaging we obtain $T_{s,t}^w \tilde{b} = (t-s) b$, which shows that for such choice of $\tilde{b}$, $T^w \tilde{b}$ cannot have better spatial regularity than $\tilde{b}$. The situation is more interesting if one defines T^w for time independent drifts only. Prevalence statements for that case will be analysed in the companion paper [29].

In order to show prevalence of regularisation by averaging, we first need to show that such a property indeed defines Borel sets in suitable spaces of paths. To this end, we require F to be another Banach space which embeds into $\mathcal{S}'(\mathbb{R}^d)$ and which enjoys the following *Fatou type property*: if $\{x_n\}_n$ is a bounded sequence in F such that x_n converge to x in the sense of distributions, then $x \in F$ and $\|x\|_F \leqslant \liminf \|x_n\|_F$.

In the next lemma we allow any $\bar{\gamma} \in (0, 1)$, but our primary focus will be $\bar{\gamma} = 1/2$.

Lemma 3.6. *Let F be as above, and let $b \in L_t^q E$ for some $q > 2$. Then for any $\bar{\gamma} \in (0, 1)$, the set*

$$\mathcal{A}^{\bar{\gamma}} = \left\{ w : [0, T] \to \mathbb{R}^d \text{ such that } T^w b \in C_t^\gamma F \text{ for some } \gamma > \bar{\gamma} \right\}$$

is Borel measurable with respect to the following topologies: L^p with $p \in [1, \infty]$, and C^α with $\alpha \geqslant 0$.

Proof. We can write $\mathcal{A}^{\bar\gamma}$ as a countable union of sets as follows:

$$\mathcal{A}^{\bar\gamma} = \bigcup_{m,n\in\mathbb{N}} \mathcal{A}^{\bar\gamma}_{m,n} := \bigcup_{m,n\in\mathbb{N}} \big\{w : [0,T] \to \mathbb{R}^d \text{ such that } \|T^w b\|_{C^{\bar\gamma+1/m}F} \leqslant n\big\};$$

in order to show the statement, it suffices to show that for every m, n the set $\mathcal{A}^{\bar\gamma}_{m,n}$ is closed in the above topologies. It suffices to show that it is closed in the L^1-topology, which is weaker than any of the others considered. Let w^k be a sequence in $\mathcal{A}^{\bar\gamma}_{m,n}$ such that $w^k \to w$ in L^1. Then by Lemma 3.2 we know that $T^{w^k} b \to T^w b$ in $C([0,T];E)$ and so that for any $s < t$, $T^{w^k}_{s,t} b \to T^w_{s,t} b$ in E and in $\mathcal{S}'(\mathbb{R}^d)$. On the other hand, by definition of $\mathcal{A}^{\bar\gamma}_{m,n}$ it holds

$$\sup_k \frac{\|T^{w^k}_{s,t} b\|_F}{|t-s|^{\bar\gamma+1/m}} \leqslant n,$$

which implies by the Fatou property of F that $T^w_{s,t} b \in F$ and

$$\frac{\|T^w_{s,t} b\|_F}{|t-s|^{\bar\gamma+1/m}} \leqslant n.$$

As the reasoning holds for any $s < t$, it follows that $T^w b \in \mathcal{A}^{\bar\gamma}_{m,n}$ as well. ■

Remark 3.7. Any weakly-$*$ compact Banach space F which embeds in $\mathcal{S}'(\mathbb{R}^d)$ satisfies the Fatou property. In the following we will always work with L^p_x-based function spaces with $p \in [2,\infty]$, so the property holds automatically. Let us also point out that the proof actually works more generally for conditions of the form $T^w b \in C^\omega_t F$, where ω is a prescribed modulus of continuity.

We are now ready to provide a first prevalence statement.

Theorem 3.8. *Let $b \in L^\alpha_t L^{s,p}_x$ (or $b \in L^\alpha_t B^s_{p,q}$) for some $\alpha > 2$, $s \in \mathbb{R}$ and $p, q \in [2,\infty)$. Let $\delta \in [0,1)$ and $\beta \in \mathbb{R}$ satisfy*

$$\beta < s + \frac{1}{\delta}\Big(\frac{1}{2} - \frac{1}{\alpha}\Big) - \frac{d}{p}, \tag{3.1}$$

where d is the space dimension, i.e., $L^{s,p}_x = L^{s,p}(\mathbb{R}^d;\mathbb{R}^m)$, and we adopt the convention that (3.1) is satisfied for any β if $\delta = 0$. Then for almost every $\varphi \in C^\delta_t$, $T^\varphi b \in C^\gamma_t C^\beta_x$ for some $\gamma > 1/2$.

Proof of Theorem 3.8. By Lemma 3.6, the set

$$\mathcal{A} = \big\{ w \in C^\delta_t : T^w b \in C^\gamma_t C^\beta_x \text{ for some } \gamma > 1/2 \big\}$$

is Borel in C^δ_t. For simplicity we will adopt the notation $b \in L^\alpha_t E$, as the reasoning is the same for $E = L^{s,p}_x$ or $E = B^s_{p,q}$. In order to prove the statement, it remains to find a suitable tight probability distribution μ on C^δ_t such that for any $\varphi \in C^\delta_t$ it holds

$$\mu(\varphi + \mathcal{A}) = \mu\big(w \in C^\delta_t : T^{\varphi+w} b \in C^\gamma_t C^\beta_x \text{ for some } \gamma > 1/2\big) = 1. \tag{3.2}$$

Thanks to the translation invariance of $\|\cdot\|_E$, we can reduce the above problem to an easier one. Indeed, setting $\tilde{b}_t := \tau^{\varphi_t} b_t$ for all $t \in [0,T]$, $\tilde{b} \in L^\alpha_t E$ and it holds $T^{\varphi+w} b = T^w \tilde{b}$. In particular, in order to show that (3.2) holds for fixed $b \in L^\alpha_t E$ and for all $\varphi \in C^\delta_t$, it actually suffices to find μ such that

$$\mu\big(w \in C^\delta_t : T^w \tilde{b} \in C^\gamma_t C^\beta_x \text{ for some } \gamma > 1/2\big) = 1 \quad \text{for all } \tilde{b} \in L^\alpha_t E. \tag{3.3}$$

Considering equation (3.3) for the choice $E = L^{s,p}$ (respectively, $E = B^s_{p,q}$), it suffices to show that for all β satisfying (3.1) there exists a tight measure $\mu_{\beta,\delta}$ on C^δ_t such that

$$\mu_{\beta,\delta}\big(w \in C^\delta_t : T^w b \in C^\gamma_t C^\beta_x \text{ for some } \gamma > 1/2\big) = 1 \quad \text{for all } b \in L^\alpha_t E. \tag{3.4}$$

The rest of the section will be devoted to the identification of such a measure. In particular, using Theorem 3.12 (respectively, Theorem 3.16) combined with Remark 3.21 below, we can choose $\mu_{\beta,\delta} = \mu^H$ to be the law of a fractional Brownian motion of parameter $H \in (0,1)$ such that $H > \delta$ and

$$\beta < s + \frac{1}{H}\Big(\frac{1}{2} - \frac{1}{\alpha}\Big) - \frac{d}{p}. \qquad \blacksquare$$

We conclude this section with a lemma on approximation by mollifications which will be very useful in Section 4.

Lemma 3.9. *Let $b \in L^q_t E$ such that $T^w b \in C^\gamma_t C^\beta_x$ for some $\gamma \in (0,1]$, $\beta \in (0,\infty)$, and let $\{\rho^\varepsilon\}_{\varepsilon>0}$ be a family of standard spatial mollifiers. Let $b^\varepsilon := \rho^\varepsilon * b$. Then, for any $\delta > 0$, it holds $T^w b^\varepsilon \to T^w b$ locally in $C^{\gamma-\delta}_t C^{\beta-\delta}_x$; namely, for any $R > 0$, $T^w b^\varepsilon|_{[0,T]\times B_R} \to T^w b|_{[0,T]\times B_R}$ in $C^{\gamma-\delta}([0,T]; C^{\beta-\delta}(B_R))$.*

Proof. It follows immediately from the property $(T^w b)^\varepsilon = T^w b^\varepsilon$ that

$$\|T^w b^\varepsilon\|_{C^\gamma_t C^\beta_x} \leqslant \|T^w b\|_{C^\gamma_t C^\beta_x} \quad \forall \varepsilon > 0,$$

and moreover that $T^w b^\varepsilon(t) \to T^w b(t)$ in $\mathcal{S}'(\mathbb{R}^d)$ as $\varepsilon \to 0$. For any $R > 0$ and $\delta > 0$, thanks to the above uniform bound, we can extract by Ascoli–Arzelà a (not relabelled) subsequence such that $T^w b^\varepsilon|_{[0,T]\times B_R}$ converges in $C^{\gamma-\delta}([0,T]; C^{\beta-\delta}(B_R))$ to a suitable limit; by the above convergence in probability, the limit must necessarily coincide with $T^w b|_{[0,T]\times B_R}$, and since the reasoning holds for any subsequence we can extract, the whole sequence must converge to $T^w b|_{[0,T]\times B_R}$. $\blacksquare$

3.2. Fractional Brownian motion and the Itô–Tanaka formula

In view of concluding the proof of Theorem 3.8, we give here the essential details on the fractional Brownian motion (fBm), whose law will be used as a transverse measure for prevalence.

In the literature, it is more common the use of *probes*, that is, finite dimensional transverse measures in order to establish prevalence properties. The only other work we are aware of using general stochastic processes in this context is [7]. However, see also [47] and the references therein for the study of properties of fractional Brownian motion with deterministic drift.

The material on fractional Brownian motion presented here is classical and taken from [44] and [48]. A one dimensional fBm $(W^H_t)_{t\geqslant 0}$ of Hurst parameter $H \in (0,1)$ is a mean zero continuous Gaussian process with covariance

$$\mathbb{E}[W^H_t W^H_s] = \frac{1}{2}(|t|^{2H} + |s|^{2H} - |t-s|^{2H}).$$

When $H = 1/2$, it coincides with the standard Brownian motion, and for $H \neq 1/2$ it is not a semi-martingale nor a Markov process. However it shares many properties with Brownian motion, such as stationarity, reflexivity and self-similarity. The trajectories of fBm are $\mathbb{P}$-a.s. δ-Hölder continuous for any $\delta < H$ and nowhere δ-Hölder continuous for any $\delta \geqslant H$; it follows from Ascoli–Arzelà that its law μ^H is tight on C^δ_t for any $\delta < H$.

A d-dimensional fBm W^H of Hurst parameter $H \in (0,1)$ is an $\mathbb{R}^d$-valued Gaussian process with components given by independent one dimensional fBms; we state for simplicity in the rest of the section all the results for $d = 1$, but they generalise immediately to higher dimension reasoning component by component.

A very useful property of fBm is that it admits representations in terms of stochastic integrals. Given a two-sided Brownian motion $\{B_t\}_{t\in\mathbb{R}}$, a fBm of parameter $H \neq 1/2$ can be constructed by

$$W^H_t = c_H \int_{-\infty}^t \left[(t-r)_+^{H-1/2} - (-r)_+^{H-1/2}\right] \mathrm{d}B_r \tag{3.5}$$

where $c_H = \Gamma(H+1/2)^{-1}$ is a suitable renormalising constant. Such a representation is usually called *non canonical*, as the filtration $\mathcal{F}_t = \sigma(B_s : s \leqslant t)$ is strictly larger than the one generated by W^H; it is useful as it immediately shows that, for any pair $0 \leqslant s < t$, the variable W^H_t decomposes into the sum of two mean zero Gaussian variables,

$$W^H_t = W^{1,H}_{s,t} + W^{2,H}_{s,t},$$

where

$$W^{1,H}_{s,t} = c_H \int_s^t (t-r)^{H-1/2}\,\mathrm{d}B_r, \quad W^{2,H}_{s,t} = c_H \int_{-\infty}^s [(t-r)_+^{H-1/2} - (-r)_+^{H-1/2}]\,\mathrm{d}B_r,$$

with $W^{2,H}_{s,t}$ being $\mathcal{F}_s$-measurable and $W^{1,H}_{s,t}$ being independent of $\mathcal{F}_s$ and with variance

$$\mathrm{Var}(W^{1,H}_{s,t}) = \tilde{c}_H |t-s|^{2H},$$

where $\tilde{c}_H = c_H^2/(2H)$. In particular this implies that

$$\mathrm{Var}\left(W^H_t | \sigma(W^H_r, r \leqslant s)\right) \geqslant \mathrm{Var}(W^H_t|\mathcal{F}_s) = \mathrm{Var}(W^{1,H}_{s,t}) = \tilde{c}_H |t-s|^{2H} \tag{3.6}$$

which is a local nondeterminism property. Loosely speaking, it means that for any $s < t$, the increment $W^H_t - W^H_s$ contains a part which is independent of the the history of the path $W^H_\cdot$ up to time s and therefore makes the path $W^H_\cdot$ “intrinsically chaotic”. The local nondeterminism property was first formulated by Berman in [9] in a different context; it plays a major role in the proofs of this section and indeed the prevalence statement can be alternatively proved by using the laws of other locally nondeterministic Gaussian processes, see Remark 3.20.

We are going to prove an Itô–Tanaka type formula for averaged functionals, in the same spirit of the one considered in [15]. We first need to recall the Clark–Ocone formula, see [44]. Given a two-sided standard Brownian motion B on a space $(\Omega, \mathcal{F}, \mathbb{P})$, $\mathcal{F}_t = \sigma(B_s, s \leqslant t)$, and given a Malliavin differentiable random variable A with Malliavin derivative $D.A$, the Clark–Ocone formula states that

$$A = \mathbb{E}[A] + \int_{-\infty}^{+\infty} \mathbb{E}[D_r A | \mathcal{F}_r] \, \mathrm{d}B_r. \tag{3.7}$$

From (3.7) it follows immediately that, for any $s \in \mathbb{R}$, we have the more general identity

$$A = \mathbb{E}[A|\mathcal{F}_s] + \int_s^{+\infty} \mathbb{E}[D_r A|\mathcal{F}_r] \, \mathrm{d}B_r.$$

We do not provide here the general definition of Malliavin derivative of a Brownian variable, which can be found in [44]; we only provide it in the following specific case, which is the one of our interest: given a smooth function f and a variable $X = \int_{-\infty}^{+\infty} K_s \, \mathrm{d}B_s$, the Malliavin derivative of $A := f(X)$ is given by

$$D_t A = \nabla f(X) \cdot K_t. \tag{3.8}$$

In the next statement, P_t denotes the heat kernel, i.e., $P_t f = p_t * f$, where

$$p_t(x) = (2\pi t)^{-d/2} \, e^{-\frac{|x|^2}{2t}}.$$

Lemma 3.10. *Let $b\colon [0,T] \times \mathbb{R}^d \to \mathbb{R}$ be a smooth, compactly supported function. Then for any fixed $0 \leqslant s \leqslant t \leqslant T$, $H \in (0,1)$ and $x \in \mathbb{R}^d$, the following identity holds with probability* 1:

$$\begin{aligned} \int_s^t b(r, x + W_r^H) \, \mathrm{d}r &= \int_s^t P_{\tilde{c}_H |r-s|^{2H}} \, b(r, x + W_{s,r}^{2,H}) \, \mathrm{d}r \\ &\quad + \int_s^t \int_u^t P_{\tilde{c}_H |r-u|^{2H}} \, \nabla b(r, x + W_{u,r}^{2,H}) \, c_H |r-u|^{H-1/2} \, \mathrm{d}r \, \cdot \mathrm{d}B_u. \end{aligned} \tag{3.9}$$

Proof. For $H = 1/2$ the above formula is well known and coincides with a standard application of the Itô–Tanaka trick together with a representation formula for solutions of the heat equation, see for instance the discussion in [15]; so we can assume $H \neq 1/2$. Let us fix $x \in \mathbb{R}^d$. Since b is smooth, for fixed r we can apply the Clark–Ocone formula to $b(r, x + W_r^H)$ to obtain

$$\begin{aligned} b(r, x + W_r^H) &= \mathbb{E}[b(r, x + W_r^H)|\mathcal{F}_s] + \int_s^r \mathbb{E}[\nabla b(r, x + W_r^H)|\mathcal{F}_u] \, c_H (r-u)^{H-1/2} \cdot \mathrm{d}B_u \\ &= P_{\tilde{c}_H |r-s|^{2H}} b(r, x + W_{s,r}^{2,H}) \\ &\quad + \int_s^r P_{\tilde{c}_H |r-u|^{2H}} \, \nabla b(r, x + W_{u,r}^{2,H}) \, c_H |r-u|^{H-1/2} \cdot \mathrm{d}B_u \,, \end{aligned}$$

where we have used both the representation of W^H in terms of a stochastic integral and the decomposition $W_r^H = W_{u,r}^{1,H} + W_{u,r}^{2,H}$, with $W_{u,r}^{1,H}$ independent of $\mathcal{F}_u$. Integrating over $[s,t]$ and applying the stochastic Fubini theorem (which is allowed since we are assuming b smooth and compactly supported), we obtain

$$\begin{aligned}\int_s^t b(r,x+W_r^H)\,\mathrm{d}t &= \int_s^t P_{\tilde{c}_H|r-s|^{2H}} b(r,x+W_{s,r}^{2,H})\,\mathrm{d}r \\ &\quad + c_H \int_s^t \int_s^r P_{\tilde{c}_H|t-s|^{2H}} \nabla b(t,x+W_{u,r}^{2,H})|r-u|^{H-1/2}\cdot \mathrm{d}B_u \mathrm{d}r \\ &= \int_s^t P_{\tilde{c}_H|r-s|^{2H}} b(r,x+W_{s,r}^{2,H})\,\mathrm{d}r \\ &\quad + c_H \int_s^t \int_u^t P_{\tilde{c}_H|r-u|^{2H}} \nabla b(r,x+W_{u,r}^{2,H})|r-u|^{H-1/2}\,\mathrm{d}r\cdot \mathrm{d}B_u,\end{aligned}$$

which gives the conclusion. ∎

The previous result can be strengthened by considering for instance $b \in C_b^1$ instead of smooth, or showing that we can find a set of probability 1 on which the identity holds for all $0 \leqslant s \leqslant t \leqslant T$; we do not do it here since it is not needed for our purposes. Instead, we need to strengthen the result to the following functional equality.

Theorem 3.11. *Let $b\colon [0,T]\times\mathbb{R}^d \to \mathbb{R}$ be a smooth, compactly supported function. Then for any fixed $0 \leqslant s \leqslant t \leqslant T$, $H \in (0,1)$, with probability 1 it holds*

$$\begin{aligned}T^{W^H} b_{s,t} &= \int_s^t P_{\tilde{c}_H|r-s|^{2H}} b\left(r,\cdot+W_{s,r}^{2,H}\right)\,\mathrm{d}r \\ &\quad + c_H \int_s^t \int_u^t P_{\tilde{c}_H|r-u|^{2H}} \nabla b\left(r,\cdot+W_{u,r}^{2,H}\right)|r-u|^{H-1/2}\,\mathrm{d}r\cdot \mathrm{d}B_u,\end{aligned} \tag{3.10}$$

where the first integral must be interpreted as a Bochner integral, and the second one as a functional stochastic integral.

We postpone the proof of this result to Appendix A.3, as it is quite technical and requires some knowledge of stochastic integration in UMD spaces. Up to technical details, it is mostly a rewriting of the statement already contained in Lemma 3.10 without further insights.

3.3. Regularity estimates in Bessel and Besov spaces

We provide here the regularity estimates for $T^w b$ when w is sampled as a fBm of parameter H, in view of establishing (3.4).

The main ingredients of the proof are the use of the functional Itô–Tanaka formula given in (3.10) together with Burkholder's inequality (Theorem A.13 below), heat kernel and interpolation estimates from Lemmata A.10 and A.9. We refer the reader to Appendices A.2 and A.3 for more information on these tools. Let us point out that the strategy of proof is fairly general and in principle could work also in other classes of spaces, up to the requirement that the above tools are still available. However, in order to apply Burkholder's inequality, we need to restrict to scales of L^p-based spaces with $p \geqslant 2$. See Appendix A.3 for a deeper discussion of this point.

Although our main aim is to establish prevalence results, our results are also new in the probabilistic setting, and therefore we will try to give their sharpest versions. In particular, we will always achieve exponential integrability whenever it is possible.

Theorem 3.12. *Let W^H be a fBm of parameter H and let $b \in L^q_t L^{s,p}_x$ for some $p, q \in [2, \infty)$. Then for any $\rho > 0$ satisfying*

$$H\rho + \frac{1}{q} < \frac{1}{2}, \tag{3.11}$$

$T^{W^H} b \in C^\gamma_t L^{s+\rho,p}_x$ for some $\gamma > 1/2$ with probability 1; moreover, there exist positive constants λ and K, independent of b, such that

$$\mathbb{E}\Big[\exp\Big(\lambda \frac{\|T^{W^H} b\|^2_{C^\gamma L^{s+\rho,p}}}{\|b\|^2_{L^q L^{s,p}}}\Big)\Big] \leqslant K. \tag{3.12}$$

Proof. Without loss of generality, we can assume $s = 0$. Indeed, if $b \in L^q_t L^{s,p}_x$, then $b = G^s \tilde{b}$, where $\tilde{b} \in L^q_t L^p_x$ and $\|b\|_{L^q L^{s,p}} = \|\tilde{b}\|_{L^q L^p}$; once the statement is shown for $\tilde{b}$, we can use the commutating property of averaging operators $T^{W^H} b = T^{W^H}(G^s \tilde{b}) = G^s (T^{W^H} \tilde{b})$ to obtain the analogue statement for b as well.

Let us first assume b to be a smooth function. By the Itô–Tanaka formula,

$$\begin{aligned}
\int_s^t b(r, \cdot + W^H_r)\,\mathrm{d}r = \int_s^t & P_{\tilde{c}_H |r-s|^{2H}}\, b(r, \cdot + W^{2,H}_{s,r})\,\mathrm{d}r \\
& + c_H \int_s^t \int_u^t P_{\tilde{c}_H |r-u|^{2H}} \nabla b(r, \cdot + W^{2,H}_{u,r}) |r-u|^{H-1/2}\,\mathrm{d}r \cdot \mathrm{d}B_u \\
& =: I^{(1)}_{s,t} + I^{(2)}_{s,t}.
\end{aligned}$$

From now on, for simplicity, we will drop the constants c_H and $\tilde{c}_H$, as they do not play any significant role in the following calculations. For the first term, we can apply the deterministic estimate:

$$\begin{aligned}
\|I^{(1)}_{s,t}\|_{L^{\rho,p}} &= \Big\| \int_s^t P_{|r-s|^{2H}}\, b\left(r, \cdot + W^{2,H}_{s,r}\right)\,\mathrm{d}r \Big\|_{L^{\rho,p}} \leqslant \int_s^t \big\| P_{|r-s|^{2H}}\, b_r \big\|_{L^{\rho,p}}\,\mathrm{d}r \\
&\lesssim \int_s^t |r-s|^{-\rho H}\, \|b_r\|_{L^p}\,\mathrm{d}r \leqslant \|b\|_{L^q L^p} \Big| \int_s^t |r-s|^{-\rho H q'}\,\mathrm{d}r \Big|^{1/q'} \\
&\lesssim \|b\|_{L^q L^p}\, |t-s|^{1-1/q-\rho H},
\end{aligned}$$

where we used the heat kernel estimates for Bessel spaces, see Lemma A.10, and the fact that the $L^{\rho,p}$-norm of b_r is not affected by a translation of $W^{2,H}_{r,s}$. Observe that $\rho H q' < 1$ is granted by condition (3.11). Moreover, (3.11) implies that $1 - 1/q - \rho H > 1/2$ and therefore we deduce that there exists $\gamma > 1/2$ such that, uniformly in $\omega \in \Omega$,

$$\|I^{(1)}\|_{C^\gamma L^{s,p}} \lesssim \|b\|_{L^q L^p}. \tag{3.13}$$

For the second term, applying Burkholder's inequality (A.12) (which is allowed since $L^{\rho,p}$ with $p \geqslant 2$ is a martingale type 2 space), we obtain

$$\mathbb{E}[\| I^{(2)}_{s,t} \|^{2k}_{L^{\rho,p}}] \\ \leqslant (Ck)^k\ \mathbb{E}\Big[\Big(\int_s^t \Big\| \int_u^t P_{|r-u|^{2H}} \nabla b(r, \cdot + W^{2,H}_{u,r}) |r-u|^{H-1/2}\, \mathrm{d}r \Big\|^2_{L^{\rho,p}} \mathrm{d}s \Big)^k \Big].$$

We can then estimate the inner integral by deterministic estimates similar to the ones above:

$$\begin{aligned} &\Big\| \int_u^t P_{|r-u|^{2H}} \nabla b(r, \cdot + W^{2,H}_{u,r})\, |r-u|^{H-1/2}\, \mathrm{d}r \Big\|_{L^{\rho,p}} \\ &\leqslant \int_u^t \| P_{|r-u|^{2H}} \nabla b_r \|_{L^{\rho,p}} |r-u|^{H-1/2}\, \mathrm{d}r \lesssim \int_u^t |r-u|^{-H(\rho+1)+H-1/2}\, \|b_r\|_{L^p}\, \mathrm{d}r \\ &\leqslant \|b\|_{L^q L^p} \Big(\int_u^t |r-u|^{-(H\rho+1/2)q'}\, \mathrm{d}r \Big)^{1/q'} \lesssim \|b\|_{L^q L^p} |t-u|^{1/2-1/q-H\rho}, \end{aligned}$$

where again we used the fact that $(H\rho + 1/2)q' < 1$, thanks to (3.11). Setting $\varepsilon := 1 - 2/q - 2H\rho$, inserting the estimate inside the one for $I^{(2)}_{s,t}$ we obtain that, for a suitable $C' > 0$, it holds

$$\mathbb{E}\big[\, \| I^{(2)}_{s,t} \|^{2k}_{L^{\rho,p}} \big] \leqslant (C'k)^k \|b\|^{2k}_{L^q L^p}\, |t-s|^{k(1+\varepsilon)}.$$

But then we have

$$\begin{aligned} \mathbb{E}\Big[\exp\Big(\lambda\, \frac{\| I^{(2)}_{s,t} \|^2_{L^{\rho,p}}}{|t-s|^{1+\varepsilon} \|b\|^2_{L^q L^p}} \Big) \Big] &= \sum_k \frac{\lambda^k}{k!}\, \mathbb{E}\Big[\frac{\| I^{(2)}_{s,t} \|^{2k}_{L^{\rho,p}}}{|t-s|^{k(1+\varepsilon)} \|b\|^{2k}_{L^q L^p}} \Big] \\ &\leqslant \sum_k \frac{(\lambda C')^k k^k}{k!} \lesssim \sum_k (\lambda C' e)^k < \infty \end{aligned}$$

as soon as $\lambda < (C'e)^{-1}$. It follows from Lemma A.1 that $I^{(2)} \in C^{1/2+\varepsilon'}_t L^{\alpha,p}_x$ for any $\varepsilon' < \varepsilon$ and that there exists another $\lambda > 0$ (not relabelled for simplicity) such that

$$\mathbb{E}\Big[\exp\Big(\lambda\, \frac{\| I^{(2)} \|^2_{C^{1/2+\varepsilon'} L^{\rho,p}}}{\|b\|^2_{L^q L^p}} \Big) \Big] \leqslant K \tag{3.14}$$

for a constant K independent of b. This, together with (3.13), proves the claim for smooth b.

Now let b be a generic element of $L^q_t L^p_x$; let us consider the case $q < \infty$ first. We can then find a sequence b_n of smooth functions such that $\|b - b_n\|_{L^q L^p} \to 0$ as $n \to \infty$; we know that in this case $\|T^{W^H}(b_n - b)\|_{C^0 L^p} \to 0$, uniformly on $\omega \in \Omega$. On the other hand, it follows from (3.12), applied to $b_n - b_m$, that for any k it holds

$$\mathbb{E}\big[\, \| T^{W^H}(b_n - b_m) \|^{2k}_{C^\gamma L^{\rho,p}} \big] \lesssim_k \|b_n - b_m\|^{2k}_{L^q L^p},$$

which implies that the sequence $T^{W^H} b_n$ is Cauchy in $L^{2k}(\Omega, \mathbb{P}; C_t^\gamma L_x^{\rho,p})$, hence it admits a limit. But then the limit must coincide with $T^{W^H} b \in L^{2k}(\Omega, \mathbb{P}; C_t^\gamma L_x^{\rho,p})$. Applying Fatou's lemma we deduce

$$\mathbb{E}\left[\exp\left(\lambda \frac{\|T^{W^H} b\|^2_{C^\gamma L^{\rho,p}}}{\|b\|^2_{L^q L^p}}\right)\right] \leqslant \liminf_{n\to\infty} \mathbb{E}\left[\exp\left(\lambda \frac{\|T^{W^H} b_n\|^2_{C^\gamma L^{\rho,p}}}{\|b_n\|^2_{L^q L^p}}\right)\right] \leqslant K,$$

which gives the conclusion. In the case $q = \infty$, since $b \in L_t^{q'} L_x^p$ for every $q' < \infty$, for any fixed ρ we can find q' big enough such that (3.11) still holds and apply the result for such q'. ∎

Remark 3.13. Theorem 3.12 immediately implies that, under assumption (3.11), the random averaging operator $T^{W^H}: b \mapsto T^w b$ is a linear bounded map from $L_t^q L_x^{s,p}$ into $L^k(\Omega, \mathbb{P}; C_t^\gamma L_x^{s+\rho,p})$, for any $k \in \mathbb{N}$. Observe the difference with Remark 3.5.

We can actually even improve the regularity result of Theorem 3.12.

Corollary 3.14. *Let $b \in L_t^q L_x^{s,p}$ with $p \in [2, \infty)$, $\rho > 0$, and assume (3.11) holds. Then there exist $\gamma > 1/2$ and a function $K(\lambda)$ independent of b such that*

$$\mathbb{E}\left[\exp\left(\lambda \frac{\|T^{W^H} b\|^2_{C^\gamma L^{s+\rho,p}}}{\|b\|^2_{L^q L^{s,p}}}\right)\right] \leqslant K(\lambda) < \infty \quad \forall \lambda \in \mathbb{R}.$$

Proof. As before, we can assume without loss of generality $s = 0$. If ρ satisfies (3.11), then there exists $\varepsilon > 0$ such that also $\rho + \varepsilon$ satisfies (3.11); it then follows from Lemma A.9 that

$$\|T^{W^H} b\|_{C^\gamma L^{\rho,p}} \lesssim \|T^{W^H} b\|^{1-\theta}_{C^\gamma L^p} \|T^{W^H} b\|^{\theta}_{C^\gamma L^{\rho+\varepsilon,p}} \leqslant \|b\|^{1-\theta}_{L^q L^p} \|T^{W^H} b\|^{\theta}_{C^\gamma L^{\rho+\varepsilon,p}},$$

where $\theta = \varepsilon/(s + \varepsilon)$ and we used the fact that $q > 2$ due to condition (3.11). It follows that

$$\frac{\|T^{W^H} b\|^{2/\theta}_{C^\gamma L^{\rho,p}}}{\|b\|^{2/\theta}_{L^q L^p}} \lesssim \frac{\|T^{W^H} b\|^2_{C^\gamma L^{\rho+\varepsilon,p}}}{\|b\|^2_{L^q L^p}},$$

where $1/\theta = (s + \varepsilon)/\varepsilon =: \beta$. Applying Theorem 3.12 to $\rho + \varepsilon$, we obtain that there exist $\bar{\lambda}$ and $\bar{K}$, independent of b, such that

$$\mathbb{E}\left[\exp\left(\bar{\lambda} \frac{\|T^{W^H} b\|^{2\beta}_{C^\gamma L^{\rho,p}}}{\|b\|^{2\beta}_{L^q L^p}}\right)\right] \leqslant \mathbb{E}\left[\exp\left(C_\varepsilon \bar{\lambda} \frac{\|T^{W^H} b\|^2_{C^\gamma L^{\rho+\varepsilon,p}}}{\|b\|^2_{L^q L^p}}\right)\right] \leqslant \bar{K}.$$

Since $\beta > 1$, the conclusion follows with the constant $K(\lambda)$ given by the optimal deterministic constant such that $\exp(\lambda x^2) \leqslant K(\lambda) \exp(\bar{\lambda} x^{2\beta})/\bar{K}$ for all $x \geqslant 0$. ∎

In the limiting case in which (3.11) becomes an equality, slightly more careful estimates still allow to obtain a regularity result in space at the cost of lower time regularity.

Theorem 3.15. *Let $b \in L^q_t L^{s,p}_x$ with $p \in [2,\infty)$, $q \in (2,\infty)$, and let $\rho > 0$ satisfy*

$$H\rho + \frac{1}{q} = \frac{1}{2}\cdot \tag{3.15}$$

Then $T^{W^H} b \in C^0_t L^{s+\rho,p}_x$ with probability 1 and there exist positive constants λ and K, independent of b, such that

$$\mathbb{E}\Big[\exp\Big(\lambda\,\frac{\|T^{W^H} b\|^2_{C^0 L^{s+\rho,p}}}{\|b\|^2_{L^q L^{s,p}}}\Big)\Big] < K.$$

Proof. As before, we assume $s = 0$, b smooth, and decompose $T^{W^H} b = I^{(1)} + I^{(2)}$. Going through the same calculations for $I^{(1)}$, we obtain

$$\|I^{(2)}_{s,t}\|_{L^{\alpha,p}} \lesssim \|b\|_{L^q_t L^p_x}\,|t-s|^{1-1/q-\alpha H} = \|b\|_{L^q_t L^p_x}\,|t-s|^{1/2},$$

where the estimate is uniform in $\omega \in \Omega$; it follows immediately that

$$\mathbb{E}\big[\exp\big(\lambda\|I^{(2)}\|^2_{C_0 L^{\rho,p}}\big)\big] < \infty,$$

and therefore we only need to focus on $I^{(2)}$. By Burkholder's inequality, we have

$$\begin{aligned}
&\mathbb{E}\big[\|I^{(2)}\|^{2k}_{C^0 L^{\alpha,p}}\big]\\
&\quad\leqslant (Ck)^k\,\mathbb{E}\Big[\Big(\int_0^T \Big\|\int_u^T P_{|r-u|^{2H}}\nabla b(r,\cdot + W^{2,H}_{u,r})|r-u|^{H-1/2}\,\mathrm{d}r\Big\|^2_{L^{\alpha,p}}\,\mathrm{d}s\Big)^k\Big],
\end{aligned}$$

and as before we want to estimate the integral inside in a deterministic manner. Going through similar calculations we obtain

$$\begin{aligned}
&\int_0^T \Big\|\int_u^T P_{|r-u|^{2H}}\nabla b(r,\cdot + W^{2,H}_{u,r})|r-u|^{H-1/2}\,\mathrm{d}r\Big\|^2_{L^{\alpha,p}}\,\mathrm{d}u\\
&\qquad\qquad\lesssim \int_0^T \Big(\int_u^T |r-u|^{-H\alpha-1/2}\,\|b_r\|_{L^p}\,\mathrm{d}r\Big)^2\,\mathrm{d}u;
\end{aligned}$$

thanks to the assumptions, we can now apply the Hardy–Littlewood–Sobolev inequality to obtain

$$\Big(\int_0^T \Big(\int_u^T |r-u|^{-H\alpha-1/2}\,\|b(r)\|_{L^p}\,\mathrm{d}r\Big)^2\,\mathrm{d}u\Big)^{1/2} \lesssim \|b\|_{L^q_t L^p_x},$$

which implies

$$\mathbb{E}\big[\|I^{(2)}\|^{2k}_{C^0 L^{\alpha,p}}\big] \leqslant (C'k)^k\,\|b\|^{2k}_{L^q_t L^p_x}.$$

The conclusion then follows by expanding the exponential and choosing λ sufficiently small as before. ∎

Going through the exact same calculations as above, an analogue result can be obtained in the case of Besov spaces $B^s_{p,q}$ with $p,q \in [2,\infty)$. In order to avoid unnecessary repetitions, we omit the proof.

Theorem 3.16. *Let W^H be a fBm of parameter H and let $b \in L^\alpha_t B^s_{p,q}$ for some $p, q \in [2, \infty)$. Then for any $\rho > 0$ satisfying*

$$H\rho + \frac{1}{\alpha} < \frac{1}{2}, \tag{3.16}$$

$T^{W^H} b \in C^\gamma_t B^{s+\rho}_{p,q}$ for some $\gamma > 1/2$ with probability 1; moreover, there exists a positive function $K(\lambda)$, independent of b, such that

$$\mathbb{E}\Big[\exp\Big(\lambda \frac{\|T^{W^H} b\|^2_{C^\gamma B^{s+\rho}_{p,q}}}{\|b\|^2_{L^\alpha B^s_{p,q}}}\Big)\Big] \leqslant K(\lambda) < \infty \quad \forall \lambda \in \mathbb{R}. \tag{3.17}$$

If equality holds in (3.16), then there exist positive constant $\tilde{\lambda}$ and $\tilde{K}$, independent of b, such that

$$\mathbb{E}\Big[\exp\Big(\tilde{\lambda} \frac{\|T^{W^H} b\|^2_{C^0 B^{s+\rho}_{p,q}}}{\|b\|^2_{L^\alpha B^s_{p,q}}}\Big)\Big] < \tilde{K}.$$

We end this section with several remarks discussing various technical point and extensions, and which can be skipped on a first reading.

Remark 3.17. Heuristically, condition (3.16) can be seen as a time-space weighted regularity condition, where time counts as $1/H$ times space (which is in agreement with parabolic regularity in the case $H = 1/2$ of Brownian motion). Indeed, we know that the averaging operator T^w maps $L^\alpha B^s_{p,q}$ into $W^{1,\alpha} B^s_{p,q}$; if we assume that regularity can be distributed between time and space, it should also map $L^\alpha B^s_{p,q}$ into $W^{\theta,\alpha} B^{s+(1-\theta)/H}_{p,q}$ for any $\theta \in (0,1)$. In order to achieve $1/2 + \varepsilon$ regularity in time, it is then required $\theta - 1/\alpha > 1/2$, which implies that the regularity gain in space is at most

$$\frac{1-\theta}{H} < \frac{1}{H}\Big(\frac{1}{2} - \frac{1}{\alpha}\Big),$$

which matches exactly condition (3.16) for ρ.

Remark 3.18. The restriction to work with $B^s_{p,q}$ with $q \in [2, \infty)$ is not particularly relevant since by Besov embedding if $b \in L^\alpha_t B^s_{p,q}$, then it also belongs to $L^\alpha_t B^s_{p,q'}$ for any $q' > q$ and to $L^\alpha_t B^{s-\varepsilon}_{p,q'}$ for any $q' < q$ and $\varepsilon > 0$, so that we can first embed it for a choice $q' \in [2, \infty)$ and then apply the estimate there. Also the restriction $p \neq \infty$ can be overcome, for instance by first localising it as $\tilde{b}$ in a ball B_R and then embedding it into some $p < \infty$; by the properties of averaging, we know that $T^{W^H} b = T^{W^H} \tilde{b}$ in $B_{R-\|W^H\|_\infty}$ and we can choose R big enough such that $\mathbb{P}(R - \|W^H\|_\infty < R/2)$ is very small, to deduce local estimates for $T^{W^H} b$ which hold with high probability. Alternatively, estimates for averaging in Besov–Hölder spaces have been given by a different technique in [12], Section 4.1. However, for simplicity, when dealing with $b \in L^\alpha_t B^s_{\infty,\infty}$, we will always assume that b has compact support in space, uniformly in time, so that we can embed it in $L^\alpha_t B^s_{p,p}$ for any $p < \infty$ and then apply estimates there.

Remark 3.19. The restriction to work with L^p-based spaces with $p \geqslant 2$ is more restrictive and it would be of fundamental importance to weaken it, especially reaching the case $p = 1$; this was already pointed out in Conjecture 1.2 from [12]. The reason is that, by the properties of averaging, we know that for any $K \in C_c^\infty$ and time independent b it holds $K * T^w b = T^w(K * b) = (T^w K) * b$; if we were able to show that $T^w K \in C_t^\gamma W_x^{\rho,1}$ with an estimate that only depends on the L^1-norm of K, then we could automatically deduce regularity estimates of the form $K * T^w b \in C_t^\gamma L^{\rho,p}$ with $b \in L^p$ for any $p \in [1,\infty]$. We could then consider a family of mollifiers obtained by rescaling K (which all have the same L^1-norm, so the same estimate in $C_t^\gamma W_x^{\rho,1}$) to get estimates for the map $b \mapsto T^w b$ in any L^p based space with $p \in [1,\infty]$ (as above, only time independent b considered).

Remark 3.20. A closer look at the proofs shows that both the Itô–Tanaka formula from Theorem 3.11 and the regularity estimates from Theorems 3.12 and 3.16 can be generalised to Gaussian processes X different from fBm and of the form

$$X_t = \int_0^t K(t,s)\,\mathrm{d}B_s,$$

for some deterministic matrix-valued function K, such that for some $H \in (0,1)$ it holds

$$\operatorname{Var}(X_t|\mathcal{F}_s) \gtrsim |t-s|^{2H} \quad \forall s < t, \tag{3.18}$$

where $\mathcal{F}_t = \sigma(B_s : s \leqslant t)$. Condition (3.18) is a type of strong local nondeterminism (SLND) and these type of processes satisfy many interesting properties, which are studied in detail in [29].

Remark 3.21. It follows immediately from the above results and from Bessel (respectively Besov) embeddings (see Appendix A.2) that if $b \in L_t^\alpha L_x^{s,p}$ (respectively $b \in L_t^\alpha B_{p,q}^s$) for some $\alpha > 2$, $p, q \in [2,\infty)$, then for any β such that

$$\beta < s + \frac{1}{H}\Big(\frac{1}{2} - \frac{1}{\alpha}\Big) - \frac{d}{p}, \tag{3.19}$$

there exists $\gamma > 1/2$ such that $T^{W^H} b \in C_t^\gamma C_x^\beta$ with full probability. For instance, in the case $s = 0$, i.e., $b \in L_t^\alpha L_x^p$, in order to require $T^{W^H} b \in C_t^\gamma C_x^0$ it is enough

$$\frac{1}{\alpha} + H\frac{d}{p} < \frac{1}{2},$$

while in order to require $T^{W^H} b \in C_t^\gamma C_x^1$ it suffices

$$\frac{1}{\alpha} + H\frac{d}{p} < \frac{1}{2} - H.$$

If $b \in L_t^\alpha B_{\infty,\infty}^s$ is compactly supported in space, uniformly in time, then it holds $T^{W^H} b \in C_t^\gamma C_x^n$ if

$$H < \frac{1}{n-s}\Big(\frac{1}{2} - \frac{1}{\alpha}\Big).$$

Remark 3.22. Finally, let us compare our results for $T^{W^H} b$ with existing literature; with the exception of the case $H = 1/2$, in which classical stochastic calculus provides more refined information, the only references we are aware of are the aforementioned [12, 38]. The technique applied in [12] allows to deal only with time independent b; however, introducing suitable weighted spaces, it does not require b to belong to $B^s_{p,p}$ for some $p < \infty$. The results from Section 7 of [38], where $b \in L^q_t B^s_{\infty,\infty}$ is considered, are in line with those from Remark 3.21; still, the techniques used therein, based on moment estimates and the Garsia–Rodemich–Rumsey lemma, do not provide global regularity estimates for $T^{W^H} b$ (only local ones) nor the exponential integrability (3.12). Both such features will be fundamental in the solution theory presented the next section: global estimates avoid finite time blow-up of solutions, exponential integrability allows the use of Girsanov's theorem. Finally, let us point out that both references only provide estimates for $T^{W^H}_{s,t} b$ in $B^s_{\infty,\infty}$, not covering other scales $B^s_{p,q}$ with $p, q < \infty$.

4. Application to perturbed ODEs

Now we are going to transfer the prevalence results for the averaged vector-field to prevalence of well-posedness to perturbed ODEs including regularity of the flow. The key technical tool to achieve this connection is a simple theory of nonlinear Young equations which we recall and adapt to our specific setting.

4.1. Perturbed ODEs as nonlinear Young differential equations

In this section we provide a summary of the results contained in [12] on nonlinear Young differential equations (YDEs). Sometimes we will provide slightly different statements which fit better our context, and in order to facilitate the understanding we will provide self-contained proofs whenever possible.

Let us fix some notation first. Given $A \in C^\gamma_t C^\nu_x = C^\gamma([0,T]; C^\nu(\mathbb{R}^d; \mathbb{R}^d))$ for $\gamma, \nu \in (0,1)$, we denote by the norm $\|A\|_{C^\gamma C^\nu}$ and the semi-norm $[\![A]\!]_{C^\gamma C^\nu}$ respectively the quantities

$$\|A\|_{C^\gamma C^\nu} = \sup_{t\in[0,T]} \|A_t(\cdot)\|_{C^\nu} + \sup_{s\neq t} \frac{\|A_{s,t}(\cdot)\|_{C^\nu}}{|t-s|^\gamma}$$

and

$$[\![A]\!]_{C^\gamma C^\nu} = \sup_{s\neq t} \frac{[\![A_{s,t}(\cdot)]\!]_{C^\nu}}{|t-s|^\gamma} = \sup_{s\neq t, x\neq y} \frac{|A(t,x) - A(t,y) - A(s,x) + A(s,y)|}{|t-s|^\gamma |x-y|^\nu}.$$

One of the main results of [12] is the rigorous construction of the nonlinear Young integral.

Theorem 4.1. *Let $\gamma, \rho, \nu \in (0,1)$ be such that $\gamma + \nu\rho > 1$, let $A \in C^\gamma_t C^\nu_x$ and let $\theta \in C^\rho_t$. Then for any $[s,t] \subset [0,T]$ and for any sequence of partitions of $[s,t]$ with mesh converging to zero, the following limit exists and is independent of the chosen sequence of partitions:*

$$\int_s^t A(\mathrm{d}u, \theta_u) := \lim_{|\Pi|\to 0} \sum_i A_{t_i, t_{i+1}}(\theta_{t_i}).$$

The limit is usually referred as a nonlinear Young integral. Furthermore:

(1) *For all* $s \leqslant r \leqslant t$ *it holds* $\int_s^r A(\mathrm{d}u, \theta_u) + \int_r^t A(\mathrm{d}u, \theta_u) = \int_s^t A(\mathrm{d}u, \theta_u)$.

(2) *If* $\partial_t A$ *is continuous, then* $\int_s^t A(\mathrm{d}u, \theta_u) = \int_s^t \partial_t A(u, \theta_u)\,\mathrm{d}u$.

(3) *There exists a universal constant* $C = C(\gamma, \rho, \nu)$ *such that*

$$\left| \int_s^t A(\mathrm{d}u, \theta_u) - A_{s,t}(\theta_s) \right| \leqslant C|t-s|^{\gamma+\nu\rho} \llbracket A \rrbracket_{C^\gamma C^\nu} \llbracket \theta \rrbracket_{C^\rho}.$$

(4) $(A, \theta) \mapsto \int_0^\cdot A(\mathrm{d}u, \theta_u)$ *is continuous as a function from* $C_t^\gamma C_x^\nu \times C_t^\rho \to C_t^\gamma$, *is linear in* A *and there exists a constant* $\tilde{C} = \tilde{C}(\gamma, \rho, \nu, T)$ *such that*

$$\left\| \int_0^\cdot A(\mathrm{d}u, \theta_u) \right\|_{C^\gamma} \leqslant \tilde{C} \|A\|_{C^\gamma C^\nu} (1 + \llbracket \theta \rrbracket_{C^\rho}).$$

The statement is a (less general) version of Theorem 2.4 from [12]; we omit the proof, but let us mention that an elementary proof based on the sewing lemma has been also given in [33]. The statement above can be localised, i.e., it is enough to require $A \in C_t^\gamma C_{\mathrm{loc}}^\nu$, and in this case all the estimates depend on the $C_t^\gamma C_x^\nu$-norm (respectively semi-norm) of A restricted to $[0, T] \times B_{\|\theta\|_\infty}$.

With this tool at hand, we can provide an alternative definition of solutions to the perturbed ODE which is meaningful even when b is distributional in space. Since we want to apply the results from Section 3, from now on when we say that b is distributional we are always going to implicitly assume that there exists $q > 2$ such that $b \in L_t^q E$, where E is a suitable space of distributions as those described in Section 3.1.

Definition 4.2. Let b be a distributional drift such that $T^w b \in C_t^\gamma C_x^\nu$ for some $\gamma, \nu \in (0, 1]$ such that $\gamma(1+\nu) > 1$. Given $x_0 \in \mathbb{R}^d$, we say that x is a solution to the ODE

$$x_t = x_0 + \int_0^t b(s, x_s)\,\mathrm{d}s + w_t \quad \forall t \in [0, T] \tag{4.1}$$

if and only if $x \in w + C^\gamma$ and $\theta = x - w$ solves the non-linear Young differential equation

$$\theta_t = \theta_0 + \int_0^t T^w b(\mathrm{d}s, \theta_s) \quad \forall t \in [0, T]. \tag{4.2}$$

Observe that the condition $\gamma(1+\nu) > 1$ immediately implies $\gamma > 1/2$, in line with standard Young differential equations; in the case of continuous b, it follows from the discussion in the introduction that the condition $x \in w + C^\gamma$ is trivially satisfied and so the two formulations (4.1) and (4.2) are equivalent, (4.1) being interpreted as the classical integral equation.

Remark 4.3. From now on we will mostly focus on solving (4.2) with $T^w b = A$ being regarded as an abstract element in a class $C_t^\gamma C_x^\nu$; however, whenever b is spatially bounded, the ODE formulation for θ is still useful, as it provides additional regularity estimates for θ compared to the ones given by the Young integral formulation: for instance, if $b \in L_{t,x}^\infty$, then any solution θ of the integral equation is automatically Lipschitz with $\llbracket \theta \rrbracket_{\mathrm{Lip}} \leqslant \|b\|_{L^\infty}$, while point (3) of Theorem 4.1 only provides estimate for $\|\theta\|_{C^\gamma}$, where $\gamma < 1$ (usually we will take γ as small as possible, namely $\gamma \sim 1/2$).

Theorem 4.4. *Let $\gamma > 1/2$ and $\nu \in (0,1)$ be such that $\gamma(1+\nu) > 1$, and assume that $T^w b \in C^\gamma_t C^\nu_x$. Then for any $\theta_0 \in \mathbb{R}^d$ there exists a solution $\theta \in C^\gamma$ to (4.2), defined on the whole interval $[0,T]$; furthermore, there exists a constant $C = C(\gamma,\nu,T)$ such that any solution to (4.2) satisfies*

$$\llbracket \theta \rrbracket_{C^\gamma} \leqslant C \left(1 + \|T^w b\|^2_{C^\gamma_t C^\nu_x}\right) \tag{4.3}$$

as well as

$$\|\theta\|_{C^0} \leqslant C \left(1 + |\theta_0| + \|T^w b\|^2_{C^\gamma_t C^\nu_x}\right). \tag{4.4}$$

Proof. The existence of solutions is granted under milder conditions on $T^w b$ by Theorem 2.9 from [12], so here we only show the a-priori estimates. Let $\theta \in C^\gamma$ be a solution and for any $\Delta > 0$ define the semi-norm

$$\llbracket \theta \rrbracket_{\gamma,\Delta} := \sup_{\substack{s \neq t \\ 0<|s-t|\leqslant \Delta}} \frac{|\theta_{s,t}|}{|t-s|^\gamma}.$$

Let Δ be a parameter to be fixed later; for any $s < t$ such that $|s-t| \leqslant \Delta$ it holds

$$\begin{aligned}
|\theta_{s,t}| &= \left| \int_s^t T^w b(\mathrm{d}r, \theta_r) \right| \\
&\leqslant |T^w b_{s,t}(\theta_s)| + C|t-s|^{\gamma(1+\nu)} \|T^w b\|_{C^\gamma C^\nu} \llbracket \theta \rrbracket^\nu_{\gamma,\Delta} \\
&\leqslant |t-s|^\gamma (\|T^w b\|_{C^\gamma C^\nu} + C\Delta^{\gamma\nu} \|T^w b\|_{C^\gamma C^\nu} \llbracket \theta \rrbracket^\nu_{\gamma,\Delta}) \\
&\leqslant |t-s|^\gamma (\|T^w b\|_{C^\gamma C^\nu} + C\Delta^{\gamma\nu} \|T^w b\|_{C^\gamma C^\nu} + C\Delta^{\gamma,\nu} \|T^w b\|_{C^\gamma C^\nu} \llbracket \theta \rrbracket_{\gamma,\Delta}),
\end{aligned}$$

where in the last passage to used the trivial inequality $a^\nu \leqslant 1 + a$ for all $a \geqslant 0$ and $\nu \in (0,1]$. Dividing both sides by $|t-s|^\gamma$ and taking the supremum s,t such that $|s-t| \leqslant \Delta$ we get

$$\llbracket \theta \rrbracket_{\gamma,\Delta} \leqslant \|T^w b\|_{C^\gamma C^\nu}(1 + C\Delta^{\gamma\nu}) + C\Delta^{\gamma\nu} \|T^w b\|_{C^\gamma C^\nu} \llbracket \theta \rrbracket_{\gamma,\Delta}.$$

Choosing Δ small enough such that $C\Delta^{\gamma\nu}\|T^w b\|_{C^\gamma C^\nu} \leqslant 1/2$, we obtain

$$\llbracket \theta \rrbracket_{\gamma,\Delta} \leqslant 2\|T^w b\|_{C^\gamma C^\nu}(1 + C\Delta^{\gamma\nu}) \lesssim 1 + \|T^w b\|_{C^\gamma C^\nu}.$$

If we can take $\Delta = T$, this provides an estimate for $\llbracket \theta \rrbracket_{C^\gamma}$, which together with $\|\theta\|_{C^0} \leqslant |\theta_0| + T^\gamma \llbracket \theta \rrbracket_{C^\gamma}$ gives the conclusion. If this is not the case, we can choose Δ as above such that in addition $C\Delta^{\gamma\nu}\|T^w b\| \geqslant 1/4$, and then by the simple inequality (see for instance Exercise 4.24 from [28])

$$\llbracket \theta \rrbracket_{C^\gamma} \lesssim \llbracket \theta \rrbracket_{\gamma,\Delta}(1 + \Delta^{\gamma-1}),$$

it follows that

$$\begin{aligned}
\llbracket \theta \rrbracket_{C^\gamma} &\lesssim (1 + \|T^w b\|_{C^\gamma C^\nu})(1 + \Delta^{\gamma-1}) \\
&\lesssim (1 + \|T^w b\|_{C^\gamma C^\nu})(1 + \|T^w b\|^{(1-\gamma)/(\gamma\nu)}_{C^\gamma C^\nu}) \lesssim 1 + \|T^w b\|^2_{C^\gamma},
\end{aligned}$$

where in the last inequality we used the fact that $\gamma(1+\nu) > 1$ implies $(1-\gamma)/(\gamma\nu) < 1$. Conclusion again follows by the standard inequality $\|\theta\|_{C^\gamma} \lesssim |\theta_0| + T^\gamma \llbracket \theta \rrbracket_{C^\gamma}$. ∎

Given that in general we consider γ to be very close to $1/2$, in order to have existence in general we need ν to be arbitrarily close to 1, thus we will usually require directly $T^w b \in C_t^\gamma \operatorname{Lip}_x$ (with the quantities $\|T^w b\|_{C^\gamma \operatorname{Lip}}$ and $[\![T^w b]\!]_{C^\gamma \operatorname{Lip}}$ defined as above).

To establish uniqueness of solutions, we need the following lemma of independent interest.

Lemma 4.5. *Let $\gamma, \nu, \rho \in (0,1]$ be such that $\gamma + \nu\rho > 1$, and let $A \in C_t^\gamma C_x^{1+\nu}$. Then for any θ^1 and $\theta^2 \in C^\rho$ it holds*

$$\int_0^t A(\mathrm{d}s, \theta_s^1) - \int_0^t A(\mathrm{d}s, \theta_s^2) = \int_0^t (\theta_s^1 - \theta_s^2) \cdot \mathrm{d}V_s,$$

where $V \in C^\gamma([0,T]; \mathscr{L}(\mathbb{R}^d; \mathbb{R}^d))$ is given by

$$V_\cdot = \int_0^1 \int_0^\cdot \nabla A(\mathrm{d}s, \theta_s^2 + x(\theta_s^1 - \theta_s^2))\, \mathrm{d}x.$$

The integral is meaningful in the Bochner sense and

$$\|V\|_{C^\gamma \mathscr{L}} \lesssim \|A\|_{C^\gamma C^{1+\nu}}(1 + [\![\theta^1]\!]_{C^\rho} + [\![\theta^2]\!]_{C^\rho}).$$

Proof. Suppose first that in addition $\partial_t A \in C_t^0 C_x^2$. Then, by Taylor's expansion,

$$\begin{aligned}
\int_0^t A(\mathrm{d}s, \theta_s^1) - \int_0^t A(\mathrm{d}s, \theta_s^2) &= \int_0^t [\partial_t A(s, \theta_s^1) - \partial_t A(s, \theta_s^2)]\, \mathrm{d}s \\
&= \int_0^t (\theta_s^1 - \theta_s^2) \cdot \int_0^1 \partial_t \nabla A(s, \theta_s^2 + x(\theta_s^1 - \theta_s^2))\, \mathrm{d}x\, \mathrm{d}s \\
&= \int_0^t (\theta_s^1 - \theta_s^2) \cdot \frac{\mathrm{d}}{\mathrm{d}s}\Big(\int_0^s \int_0^1 \partial_t \nabla A(u, \theta_u^2 + x(\theta_u^1 - \theta_u^2))\, \mathrm{d}x\, \mathrm{d}u\Big)\, \mathrm{d}s \\
&= \int_0^t (\theta_s^1 - \theta_s^2) \cdot \mathrm{d}\Big(\int_0^1 \int_0^s \partial_t \nabla A(u, \theta_u^2 + x(\theta_u^1 - \theta_u^2))\, \mathrm{d}u\, \mathrm{d}x\Big) \\
&= \int_0^t (\theta_s^1 - \theta_s^2) \cdot \mathrm{d}\Big(\int_0^1 \int_0^s \nabla A(\mathrm{d}u, \theta_u^2 + x(\theta_u^1 - \theta_u^2))\, \mathrm{d}x\Big) =: \int_0^t (\theta_s^1 - \theta_s^2) \cdot \mathrm{d}V_s,
\end{aligned}$$

where all manipulations in this case are allowed by the properties of the Young integral and the fact that we are assuming A regular; by hypothesis $\nabla A \in C_t^\gamma C_x^\nu$ and $\theta^i \in C_t^\rho$ with $\gamma + \nu\rho > 1$, so the interpretation of the integrals as nonlinear Young integrals is legit. Observe that the map $A \mapsto V(A)$ is linear by construction and we have the estimate

$$\begin{aligned}
\|V\|_{C^\gamma \mathscr{L}} &= \Big\| \int_0^1 \int_0^\cdot \nabla A(\mathrm{d}u, \theta_u^2 + x(\theta_u^1 - \theta_u^2))\, \mathrm{d}x \Big\|_{C^\gamma \mathscr{L}} \\
&\leqslant \int_0^1 \Big\| \int_0^\cdot \nabla A(\mathrm{d}u, \theta_u^2 + x(\theta_u^1 - \theta_u^2)) \Big\|\, \mathrm{d}x \\
&\lesssim \int_0^1 \|\nabla A\|_{C^\gamma C^\nu}(1 + [\![\theta^1]\!]_{C^\rho} + [\![\theta^2]\!]_{C^\rho})\, \mathrm{d}x,
\end{aligned}$$

which gives the conclusion in this case. The general case follows by approximation, considering a sequence of regular $A^n \to A$ locally in $C_t^{\gamma-\delta} C_x^{1+\nu-\delta}$, on a ball of radius $R > \|\theta^i\|_\infty$, for δ small enough such that $\gamma - \delta + (\nu - \delta)\rho > 1$. ∎

With the above lemma at hand, we can provide a comparison principle, which estimates the difference between solutions. It comes in two versions, which apply to different scenarios.

Theorem 4.6 (Comparison principle, version 1). *Let $\gamma > 1/2$ and assume that b^1, b^2 are distributional drifts such that $T^w b^i \in C_t^\gamma C_x^2$ with $\|T^w b^i\|_{C^\gamma C^2} \leqslant R$. Let $\theta^i \in C^\gamma$, $i = 1, 2$, be solutions respectively of the YDEs*

$$\theta_t^i = \theta_0^i + \int_0^t T^w b^i(\mathrm{d}s, \theta_s^i) \quad \forall t \in [0, T].$$

Then there exists a constant $C = C(\gamma, T, R)$ such that

$$\|\theta_\cdot^1 - \theta_\cdot^2\|_{C^\gamma} \leqslant C(|\theta_0^1 - \theta_0^2| + \|T^w b^1 - T^w b^2\|_{C_t^\gamma \mathrm{Lip}}). \tag{4.5}$$

Similarly, let b^i be such that $b^i \in L_{t,x}^\infty$ and $T^w b^i \in C_t^\gamma C_x^{3/2}$ with $\|b^i\|_{L^\infty} \leqslant R$, $\|T^w b^i\|_{C^\gamma C^{3/2}} \leqslant R$ for $i = 1, 2$ and let θ^i be Lipschitz solutions of the YDEs. Then there exists $\tilde{C} = \tilde{C}(\gamma, T, R)$ such that

$$\|\theta_\cdot^1 - \theta_\cdot^2\|_{C^\gamma} \leqslant \tilde{C}(|\theta_0^1 - \theta_0^2| + \|T^w b^1 - T^w b^2\|_{C_t^\gamma \mathrm{Lip}}). \tag{4.6}$$

Proof. We show in detail the derivation of (4.5) and briefly sketch the one of (4.6), as the structure of the proof is the same. By the assumptions and Lemma 4.5 applied to $A = T^w b^1$, which is allowed for the choice $\nu = 1$, $\rho = \gamma$, the difference $v = \theta^1 - \theta^2$ satisfies

$$\begin{aligned} v_t &= v_0 + \Big[\int_0^t T^w b^1(\mathrm{d}s, \theta_s^1) - \int_0^t T^w b^1(\mathrm{d}s, \theta_s^2) \Big] + \int_0^t (T^w b^1 - T^w b^2)(\mathrm{d}s, \theta_s^2) \\ &= v_0 + \int_0^t v_s \cdot \mathrm{d}V_s + \psi_t. \end{aligned}$$

This is a linear Young differential equation, for which standard estimates are available; Theorem 4.4, Lemma 4.5 and properties of nonlinear Young integral provide

$$[\![\theta^i]\!]_{C^\gamma} \lesssim 1, \quad \|V\|_{C^\gamma \mathcal{L}} \lesssim \|T^w b^1\|_{C^\gamma C^2}(1 + [\![\theta^1]\!]_{C^\gamma} + [\![\theta^2]\!]_{C^\gamma}) \lesssim 1,$$
$$[\![\psi]\!]_{C^\gamma} \lesssim \|T^w b^1 - T^w b^2\|_{C_t^\gamma \mathrm{Lip}}$$

where the constants appearing all depend on γ, T and R. Combining this estimates with Lemma A.2 from Appendix A.1 yields the conclusion.

The proof in the second case is analogue, but we have the additional estimate $[\![\theta^i]\!]_{\mathrm{Lip}} \leqslant \|b^i\|_{L^\infty}$ coming from the ODE integral interpretation of the YDE; so we can apply as above Lemma 4.5 to $T^w b^1$, this time for the choice $\nu = 1/2$, $\rho = 1$. ∎

Remark 4.7. It follows immediately from the above result that if $T^w b \in C_t^\gamma C_x^2$ or $b \in L_{t,x}^\infty$ and $T^w b \in C_t^\gamma C_x^{3/2}$, then for any $\theta_0 \in \mathbb{R}^d$ there exists a unique solution to the YDE (4.2) and moreover the solution map $\theta_0 \mapsto \theta_\cdot$ is Lipschitz continuous with respect to θ_0. The solution constructed this way is also stable under approximation of $T^w b$ by other drifts $T^w \tilde{b}$, which can be combined with Lemma 3.9, as we can take $\tilde{b} = b^\varepsilon = \rho^\varepsilon * b$ for some spatial mollifier ρ^ε.

The above version of the comparison principle is of straightforward application, as it only requires good regularity estimates on $T^w b$. The next version is instead slightly more subtle and can be regarded as a conditional comparison principle, as it allows to deduce estimates under less regularity on $T^w b$ imposing the existence of a solution with suitable properties; however, the existence of such solutions is not granted a priori by the deterministic theory and in order to construct them probabilistic tools will be needed, specifically the Girsanov transform.

Theorem 4.8 (Comparison principle, version 2). *Let $\gamma > 1/2$ and assume that b^1, b^2 are distributional drifts such that $T^w b^i \in C^\gamma_t \operatorname{Lip}_x$ with $\|T^w b^i\|_{C^\gamma \operatorname{Lip}} \leqslant R$. Let $\theta^i \in C^\gamma$, $i = 1, 2$, be solutions respectively of the YDEs*

$$\theta^i_t = \theta^i_0 + \int_0^t T^w b^i(\mathrm{d}s, \theta^i_s) \quad \forall t \in [0, T],$$

and assume that θ^1 is such that $T^{w+\theta^1} b \in C^\gamma_t \operatorname{Lip}_x$ with $\|T^{w+\theta^1} b^1\|_{C^\gamma \operatorname{Lip}} \leqslant R$. Then there exists a constant $C = C(\gamma, T)$ such that

$$\|\theta^1_\cdot - \theta^2_\cdot\|_{C^\gamma} \leqslant C \exp(C R^{1/\gamma})(|\theta^1_0 - \theta^2_0| + \|T^w b^1 - T^w b^2\|_{C^\gamma_t \operatorname{Lip}}). \tag{4.7}$$

The proof requires the following technical lemma.

Lemma 4.9. *Let w, θ be such that $T^w b$, $T^{w+\theta} b \in C^\gamma_t \operatorname{Lip}_x$, and let $\theta \in C^{1/2}_t$ for some $\gamma > 1/2$. Then for any $\tilde{\theta} \in C^{1/2}_t$ it holds*

$$\int_0^\cdot T^{w+\theta} b(\mathrm{d}s, \tilde{\theta}_s) = \int_0^\cdot T^w b(\mathrm{d}s, \tilde{\theta}_s + \theta_s). \tag{4.8}$$

Proof. If b is jointly continuous in (t, x), then the result is straightforward by the equivalence between the Young integral formulation and the standard integral formulation. Next, if b satisfies the hypothesis and in addition $b \in L^1_t C^\alpha_x$ for some $\alpha > 0$, then for any $t > 0$ and for any sequence of partitions Π_n of $[0, t]$ such that $|\Pi_n| \to 0$ it holds

$$\begin{aligned}
\Big|\int_0^t & T^{w+\theta} b(\mathrm{d}s, \tilde{\theta}_s) - \int_0^t T^w b(\mathrm{d}s, \theta_s + \tilde{\theta}_s)\Big| \\
&= \lim_{n\to\infty} \Big|\sum_i \int_{t_i}^{t_{i+1}} b(s, w_s + \theta_s + \tilde{\theta}_{t_i}) - b(s, w_s + \theta_{t_i} + \tilde{\theta}_{t_i})\, \mathrm{d}s\Big| \\
&\leqslant \lim_{n\to\infty} \sum_i \int_{t_i}^{t_{i+1}} \|b(s)\|_{C^\alpha_x} |\theta_s - \theta_{t_i}|^\alpha\, \mathrm{d}s \\
&\leqslant \|\theta\|_{C^{1/2}} \lim_{n\to\infty} |\Pi_n|^{\alpha/2} \sum_i \int_{t_i}^{t_{i+1}} \|b(s)\|_{C^\alpha_x}\, \mathrm{d}s = 0,
\end{aligned}$$

which proves the statement in this case. For a general b, consider $b^\varepsilon = \rho^\varepsilon * b$, where ρ^ε is a sequence of spatial mollifiers; for b^ε, by the previous step, identity (4.8) is true and by Lemma 3.9, $T^w b^\varepsilon \to T^w b$ locally in $C^{\gamma-\delta}_t C^{1-\delta}_x$, similarly for $T^{w+\theta} b^\varepsilon \to T^{w+\theta} b$. Choosing δ small such that $\gamma - \delta + (1 - \delta)\gamma > 1$ and using the continuity of the Young integral, we obtain the conclusion in the general case. ■

Proof of Theorem 4.8. The idea of the proof is the same as that of Theorem 4.6, and it is based on finding a Young differential equation for $v = \theta^2 - \theta^1$, only we now need to exploit the additional information on $T^{w+\theta^1} b^1$. By the assumptions combined with Lemma 4.9, v satisfies

$$\begin{aligned} v_t &= v_0 + \Big[\int_0^t T^w b^1(\mathrm{d}s, \theta^2_s) - \int_0^t T^w b^1(\mathrm{d}s, \theta^1_s)\Big] + \int_0^t (T^w b^2 - T^w b^1)(\mathrm{d}s, \theta^2_s) \\ &= v_0 + \int_0^t T^{w+\theta^1}(\mathrm{d}s, v_s) - \int_0^t T^{w+\theta^1}(\mathrm{d}s, 0) + \psi_t = v_0 + \int_0^t A(\mathrm{d}s, v_s) + \psi_t, \end{aligned}$$

where ψ_t is defined in the usual way and $A(t,x) = T^{w+\theta^1}(t,x) - T^{w+\theta^1}(t,0)$, so that $A \in C^\gamma_t \operatorname{Lip}_x$ with $[\![A]\!]_{C^\gamma \operatorname{Lip}} = [\![T^{w+\theta^1} b]\!]_{C^\gamma \operatorname{Lip}}$ and $A(t,0) = 0$ for all $t \in [0,T]$. We can then apply the estimates from Lemma A.3 from Appendix A.1 to deduce

$$\|v\|_{C^\gamma} \lesssim \exp(C[\![A]\!]^{1/\gamma}_{C^\gamma})\,(|v_0| + [\![\psi]\!]_{C^\gamma})$$

for some constant $C = C(\gamma, T)$, which together with the estimate

$$[\![\psi]\!]_{C^\gamma} \lesssim \|T^w b^2 - T^w b^1\|_{C^\gamma \operatorname{Lip}}(1 + [\![\theta^1]\!]_{C^\gamma} + [\![\theta^2]\!]_{C^\gamma}) \lesssim \|T^w b^2 - T^w b^1\|_{C^\gamma \operatorname{Lip}}(1 + R^2)$$

yields the conclusion. ■

Remark 4.10. It follows immediately from Theorem 4.8 that, if there exists a solution θ to the YDE associated to $T^w b$ with initial data θ_0 such that $T^{w+\theta} \in C^\gamma_t \operatorname{Lip}_x$, then this is necessarily the unique solution with initial data θ_0 and it is stable under perturbation. This provides a nice "duality principle": existence of solutions is granted if $T^w b \in C^\gamma_t \operatorname{Lip}_x$, uniqueness instead if there exists a solution with similar averaging properties. In the case b is continuous, so that, by Peano's theorem, existence of a solution $x = w + \theta \in w + \operatorname{Lip}$ is automatic, the statement can be rephrased as the fact that uniqueness for the Cauchy problem associated to x_0 holds under the condition $T^x b \in C^\gamma_t \operatorname{Lip}_x$ for some $\gamma > 1/2$.

Remark 4.11. For the sake of simplicity we considered from the start $T^w b \in C^\gamma_t C^\beta_x$ in order to develop a global theory in space, but many results from Section 4 can be localised, thanks to Remark 3.4, in a similar fashion to what is done in Section 2.3 of [12]. For instance, local existence holds for $T^w b \in C^\gamma_t \operatorname{Lip}_{\mathrm{loc}}$, while local existence and uniqueness holds for $T^w b \in C^\gamma_t C^2_{\mathrm{loc}}$; in the second version of the comparison principle, if there exists a solution x defined on $[0, T^*)$ such that $T^x b \in C^\gamma_t \operatorname{Lip}_{\mathrm{loc}}$, then it is the unique solution on $[0, T^*)$. Analogue considerations hold for the results from Section 4.3 on the regularity of the flow.

4.2. Prevalence for the Cauchy problem

In this section we focus on establishing conditions under which, for a given drift b and a given initial datum $x_0 \in \mathbb{R}^d$, for almost every $\varphi \in C^\delta_t$ the Cauchy problem (from now on referred to as (CP_{x_0}))

$$x_t = x_0 + \int_0^t b(s, x_s) + \varphi_t \tag{4.9}$$

is well-posed, for suitable values of δ. Here by well-posedness for (CP_{x_0}) we mean the following: φ is such that $T^{\varphi}b \in C_t^{\gamma} \operatorname{Lip}_x$ for some $\gamma > 1/2$, so that it makes sense to talk about solutions to (4.9) in the sense of Definition 4.2, and there exists a unique such solution in the class $x \in \varphi + C^{\gamma}$. The main results we are going to prove are the following.

Theorem 4.12. *Let $b \in C_x^{\alpha}$ for some $\alpha \in (-\infty, 1)$, b being compactly supported, and let $x_0 \in \mathbb{R}^d$ be fixed. Let $\delta \in [0, 1)$ satisfy*

$$\delta < \frac{1}{2(1-\alpha)}.$$

Then for almost every $\varphi \in C_t^{\delta}$ the Cauchy problem (CP_{x_0}) is well-posed.

Theorem 4.13. *Let $b \in C_x^{\alpha}$ for some $\alpha \in (-\infty, 1)$, b being compactly supported, and let $x_0 \in \mathbb{R}^d$ be fixed. Let μ^H denote the law of fBm of parameter H and suppose that*

$$\alpha > 1 - \frac{1}{2H}.$$

Then path-by-path uniqueness holds for (CP_{x_0}) and w sampled according to μ^H. Moreover, there exists $\gamma > 1/2$ which only depends on α such that

$$\mu^H \left(w : T^w b \in C_t^{\gamma} \operatorname{Lip}_x, \exists \text{ a solution } x \in w + C^{\gamma} \text{ s.t. } T^x b \in C_t^{\gamma} \operatorname{Lip}_x\right) = 1.$$

In the second statement we have used the terminology "path-by-path uniqueness" as it appears frequently in regularisation by noise results, see [21], but in the framework introduced above it just amounts to stating that there exists $\gamma > 1/2$ such that

$$\mu^H\big(w : T^w b \in C_t^{\gamma} \operatorname{Lip}_x \text{ and } (\mathrm{CP}_{x_0}) \text{ is well-posed}\big) = 1.$$

The section is organised as follows: we first prove Theorem 4.12 in Section 4.2.1 relying on the validity of Theorem 4.13; then we pass to the proof of the latter, which is based on an application of Theorem 4.8 in combination with the Girsanov transform for fBm, which is introduced in Section 4.2.2. The proof of Theorem 4.13 is completed in Section 4.2.3, along with several other results of the same nature. We leave the details to the following subsections, but let us point out already here that we will exploit crucially the general principle

$$\boxed{T^{W^H} b \in C_t^{\gamma} \operatorname{Lip}_x} + \boxed{\text{Girsanov}} \implies \boxed{\text{path-by-path uniqueness}}.$$

Such a principle is not new and was crucially exploited in [18] and [12]. However, we believe it is the first time it is properly formalised as in Lemma 4.21, and its general structure allows to apply it in other situations.

4.2.1. Proof of Theorem 4.12. We need a few preparations first. Recall that in order to establish prevalence of well-posedness for (CP_{x_0}) in C_t^{δ}, we need to find a set $\mathcal{A} \subset C_t^{\delta}$ and a tight probability μ on C_t^{δ} such that: i) $\mathcal{A}$ is Borel with respect to the topology of C_t^{δ}; ii) for all $w \in \mathcal{A}$, (CP_{x_0}) is well-posed; iii) for all $\varphi \in C_t^{\delta}$, $\mu(\varphi + \mathcal{A}) = 1$.

A good candidate for the set $\mathcal{A}$ is given by Theorem 4.8 as follows: for $\gamma > 1/2$, define

$$\mathcal{A}_\gamma = \left\{w \in C_t^{\delta} : T^w b \in C_t^{\gamma} \operatorname{Lip}_x, \exists \text{ a solution } x \text{ s.t. } T^x b \in C_t^{\gamma} \operatorname{Lip}_x\right\}. \tag{4.10}$$

For such an $\mathcal{A}_\gamma$, it is now rather clear by the statement of Theorem 4.13 that we plan to use as a measure μ^H for suitable choice of H. But we first need to check that condition i) holds, which is the aim of the following lemma.

Lemma 4.14. *Let $\gamma > 1/2$. Then the set $\mathcal{A}_\gamma$ is Borel measurable in the topology of C^δ for any $\delta \geqslant 0$.*

Proof. The idea of the proof is the usual one: we write the set $\mathcal{A}_\gamma$ as the countable union

$$\begin{aligned}\mathcal{A}_\gamma &= \bigcup_{N\geqslant 1} \mathcal{A}_N \\ &:= \bigcup_{N\geqslant 1} \big\{w \in E : \|T^w b\|_{C_t^\gamma \operatorname{Lip}_x} \leqslant N, \exists \text{ a solution } x \text{ s.t. } \|T^x b\|_{C_t^\gamma \operatorname{Lip}_x} \leqslant N\big\}.\end{aligned}$$

In order to conclude it is then sufficient to show that, for each N, the set $\mathcal{A}_N$ is closed under the topology of C^δ. We can restrict ourselves to the case C^0, since any other convergence we consider is stronger than this one.

Let w^n be a sequence of elements of $\mathcal{A}_N$ such that $w^n \to w$. Then by Lemma 3.6 we know that $T^w b \in C_t^\gamma \operatorname{Lip}_x$ with the bound $\|T^w b\|_{C_t^\gamma \operatorname{Lip}_x} \leqslant N$. For each n, denote by $x^n = \theta^n + w^n$ the associated solution of (CP_{x_0}) such that $\|T^{x^n} b\|_{C_t^\gamma \operatorname{Lip}_x} \leqslant N$; by the a priori estimates from Theorem 4.4, together with $\|T^{w^n} b\|_{C_t^\gamma \operatorname{Lip}_x} \leqslant N$, we deduce that $\|\theta^n\|_{C^\gamma}$ are uniformly bounded. We can therefore (up to subsequence) consider $\theta^n \to \theta$ in $C^{\gamma-\varepsilon}$ for suitable $\varepsilon > 0$. Since $w^n + \theta^n \to w + \theta$ in C^0, again it must hold $\|T^{w+\theta} b\|_{C_t^\gamma \operatorname{Lip}_x} \leqslant N$.

To complete the proof, it remains to show that $x = \theta + w$ is a solution of the (CP_{x_0}) associated to $T^w b$. Since the sequence $T^{w^n} b \to T^w b$ in the sense of distributions and it is uniformly bounded in $C_t^\gamma \operatorname{Lip}_x$, reasoning as in the proof of Lemma 3.9 we deduce that also local convergence in $C_t^{\gamma-\varepsilon} C_x^{1-\varepsilon}$ holds, for any $\varepsilon > 0$. Choosing $\varepsilon > 0$ small enough such that $\gamma - \varepsilon + (1-\varepsilon)/(\gamma-\varepsilon) > 1$, by continuity of the nonlinear Young integral it holds $\int_0^\cdot T^{w^n} b(\mathrm{d}s, \theta_s^n) \to \int_0^\cdot T^w b(\mathrm{d}s, \theta_s)$ in $C^{\gamma-\varepsilon}$. Taking the limit as $n \to \infty$ of

$$\theta_t^n = x_0 - w_0^n + \int_0^t T^{w^n} b(\mathrm{d}s, \theta_s^n)$$

we deduce that x is a solution with respect to $T^w b$ of (CP_{x_0}), which concludes the proof. ∎

Proof of Theorem 4.12. To prove the statement, it suffices to show that we can find $\gamma > 1/2$ and $H > \delta$ such that $\mu^H(\varphi + \mathcal{A}_\gamma) = 1$ for all $\varphi \in C_t^\delta$. Choose $\varepsilon > 0$ small enough so that

$$H := \delta + \varepsilon < \frac{1}{2(1-\alpha)}. \tag{4.11}$$

We need to find $\gamma > 1/2$ such that for any fixed $\varphi \in C_t^\delta$,

$$\mu^H\big(w \in C_t^\delta : T^{w+\varphi} b \in C_t^\gamma \operatorname{Lip}_x, \exists\, x \text{ solution to } (\mathrm{CP}_{x_0}) \text{ s.t. } T^x b \in C_t^\gamma \operatorname{Lip}_x\big) = 1.$$

By definition of the averaging operator, $T^{w+\varphi} b = T^w \tilde{b}$ for $\tilde{b}(t,\cdot) = b(t, \cdot + \varphi_t)$; moreover, $x \in (w+\varphi) + C_t^\gamma$ solves (CP_{x_0}) if and only if $\tilde{x} := x - \varphi \in w + C_t^\gamma$ is again a solution to another Cauchy problem of the same type. Indeed, by definition of solution, $\theta = x - (w + \varphi) = \tilde{x} - w$ must solve

$$\theta_t = x_0 - (w_0 + \varphi_0) + \int_0^t T^{w+\varphi} b(\mathrm{d}s, \theta_s) = \tilde{x}_0 - w_0 + \int_0^t T^w \tilde{b}(\mathrm{d}s, \theta_s)$$

where $\tilde{x}_0 = x_0 - \varphi_0$, so that $\tilde{x}$ is a solution to the Cauchy problem associated to $\tilde{x}_0$, $\tilde{b}$ and w. Moreover, by properties of averaging operators it holds $T^{\tilde{x}}\tilde{b} = T^x b$.

By the translation invariance of the C^α_x-norm, it holds $\tilde{b} \in C^\alpha_x$, $\|\tilde{b}\|_{C^\alpha} = \|b\|_{C^\alpha}$; moreover, $\tilde{b}$ has still compact support in space, uniformly in time. Since condition (4.11) implies $\alpha > 1 - (2H)^{-1}$, we can apply Theorem 4.13 for the choice $\tilde{x}_0$, $T^w\tilde{b}$ to find $\gamma > 1/2$ (independent of φ) such that

$$\begin{aligned} 1 &= \mu^H\big(w \in C^\delta_t : T^w\tilde{b} \in C^\gamma_t \operatorname{Lip}_x \exists\, \tilde{x} \text{ solution to } (\mathrm{CP}_{\tilde{x}_0}) \text{ s.t. } T^{\tilde{x}}\tilde{b} \in C^\gamma_t \operatorname{Lip}_x\big) \\ &= \mu^H\big(w \in C^\delta_t : T^{w+\varphi}b \in C^\gamma_t \operatorname{Lip}_x \ \exists\, x \text{ solution to } (\mathrm{CP}_{x_0}) \text{ s.t. } T^x b \in C^\gamma_t \operatorname{Lip}_x\big), \end{aligned}$$

which gives the conclusion. ■

Remark 4.15. For simplicity we have preferred to give the statement of Theorem 4.12 as above, but it will be clear from the contents of Section 4.2.3 that similar prevalence statements can be formulated under other hypothesis on b and δ simply by going through the same proof and applying in the end either Theorem 4.23 or Corollary 4.24.

4.2.2. Girsanov's transform. Before introducing Girsanov's theorem, we need to recall another representation formula for fBm, different from the one given in Section 3.2, which can be found in [44], [48]. The representation is based on fractional calculus, which we also quickly introduce and for which we refer the interested reader to [50].

Given $f \in L^1(0,T)$ and $\alpha > 0$, the fractional integral of order α of f is defined as

$$(I^\alpha f)_\cdot = \frac{1}{\Gamma(\alpha)} \int_0^\cdot (t-s)^{\alpha-1} f_s \,\mathrm{d}s, \tag{4.12}$$

where Γ denotes the Gamma function. For $\alpha \in (0,1)$ and $p > 1$, the map I^α is an injective bounded operator on L^p and we denote by $I^\alpha(L^p)$ the image of L^p under the I^α, which is a Banach space endowed with the norm $\|f\|_{I^\alpha(L^p)} := \|g\|_{L^p}$ if $f = I^\alpha g$. On this domain, I^α admits an inverse, which is the fractional derivative of order α, given by

$$(D^\alpha f)_t = \frac{1}{\Gamma(1-\alpha)} \frac{\mathrm{d}}{\mathrm{d}x} \int_0^t \frac{f_s}{(t-s)^\alpha} \mathrm{d}s = \frac{1}{\Gamma(1-\alpha)} \Big(\frac{f_t}{t^\alpha} + \alpha \int_0^t \frac{f_t - f_s}{(t-s)^{\alpha+1}} \mathrm{d}s \Big).$$

With this notation in mind, a fBm of Hurst parameter $H \in (0,1)$ can be constructed starting from a standard Brownian motion B on the interval $[0,T]$ by setting $W^H = K_H(\mathrm{d}B)$, where the operator K_H is defined as

$$K_H f = \begin{cases} I^1 s^{H-1/2} I^{H-1/2} s^{1/2-H} h & \text{if } H \geqslant 1/2, \\ I^{2H} s^{1/2-H} I^{1/2-H} s^{H-1/2} h & \text{if } H \leqslant 1/2, \end{cases}$$

where the notation s^β denotes the multiplication operator with the map $s \mapsto s^\beta$. It can be shown that this definition of W^H is meaningful and that the operator K_H corresponds to a Volterra kernel $K_H(t,s)$, so that the above representation is equivalent to

$$W_t^H = \int_0^t K_H(t,s)\,\mathrm{d}B_s. \tag{4.13}$$

The explicit expression for K_H in the case $H > 1/2$ is given by

$$K_H(t,s) = c_H\, s^{1/2-H} \int_s^t (u-s)^{H-3/2} u^{H-1/2}\,\mathrm{d}u. \tag{4.14}$$

In the case $H < 1/2$ it is more complicated; we omit it as we will not need it. It can be shown that the operator K_H can be inverted, which implies that the processes B and W^H generate the same filtration, which makes it a *canonical representation*; moreover, this implies that given any fBm W^H on a probability space, it is possible to construct the associated B by setting $B_\cdot = \int_0^\cdot (K_H^{-1} W^H)_s\,\mathrm{d}s$. The inverse operator K_H^{-1} is given by

$$K_H^{-1} f = \begin{cases} s^{H-1/2} D^{H-1/2} s^{1/2-H} f' & \text{if } H > 1/2, \\ s^{1/2-H} D^{1/2-H} s^{H-1/2} D^{2H} f & \text{if } H < 1/2. \end{cases} \tag{4.15}$$

We will use the following terminology: given a filtered space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geqslant 0}, \mathbb{P})$, we say that a process W^H is an $\mathcal{F}_t$-fBm if it is a fBm under $\mathbb{P}$ and the associated B is an $\mathcal{F}_t$-Bm in the usual sense.

Theorem 4.16 (Girsanov). *Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geqslant 0}, \mathbb{P})$ be a filtered probability space, let W^H be an $\mathcal{F}_t$-fBm of parameter $H \in (0,1)$ and let h be an $\mathcal{F}_t$-adapted process with continuous trajectories such that $h_0 = 0$. Let B be the Bm associated to W^H, namely such that $W^H = K_H \mathrm{d}B$. Suppose that $K_H^{-1} h \in L_t^2$ with probability 1 and that*

$$\mathbb{E}\Big[\frac{\mathrm{d}\mathbb{P}}{\mathrm{d}\mathbb{Q}}\Big] = 1, \tag{4.16}$$

where the variable $\mathrm{d}\mathbb{P}/\mathrm{d}\mathbb{Q}$ is given by

$$\frac{\mathrm{d}\mathbb{P}}{\mathrm{d}\mathbb{Q}} = \exp\Big(-\int_0^T (K_H^{-1}h)_s\,\mathrm{d}B_s - \frac{1}{2}\int_0^T |(K_H^{-1}h)_s|^2\,\mathrm{d}s\Big). \tag{4.17}$$

Then the shifted process $\tilde{W}^H := W^H + h$ is an $\mathcal{F}_t$-fBm with parameter H under the probability $\mathbb{Q}$. A sufficient condition in order for (4.16) to hold is given by Novikov's condition

$$\mathbb{E}\Big[\exp\Big(\frac{1}{2}\int_0^T |(K_H^{-1}h)_s|^2\,\mathrm{d}s\Big)\Big] < \infty. \tag{4.18}$$

The result is taken from [45], Theorem 2, with the exception of the final part which is just the classical Novikov condition; in the original statement from [45], the process h is taken of the form $h_\cdot = \int_0^\cdot u_s\,\mathrm{d}s$, but this does not play any role in the proof, which indeed holds also in the case h is not of bounded variation.

In order to apply Theorem 4.16 in cases of interest, we first need to establish conditions under which (4.18) holds, which requires a good control of $\|K_H^{-1}h\|_{L^2}$ in terms of h.

Since K_H^{-1} is defined in terms of fractional derivatives, the following fact will be quite useful: if $f \in C^\beta$ and $f_0 = 0$, then $D^\alpha f$ is well defined for any $\alpha < \beta$ and moreover $D^\alpha f \in C^\gamma$ for any $\gamma < \beta - \alpha$ together with the estimate

$$\| D^\alpha f \|_{C^\gamma} \lesssim_{\gamma,\alpha} \| f \|_{C^\beta}. \tag{4.19}$$

For a self-contained proof of this fact, see Theorem 2.8 in [48] (on a finite interval $[0, T]$, the space $\mathbb{H}^{\beta,0}$ considered therein corresponds to the functions $f \in C_t^\beta$ such that $f_0 = 0$).

Lemma 4.17. *Let $\alpha \in (0, 1/2)$ and let $h \in C_t^\beta$ for some $\beta > \alpha$, with $h_0 = 0$. Then $s^\alpha D^\alpha s^{-\alpha} h \in L_t^2$ and there exists a constant $C = C(\alpha, \beta, T)$ such that*

$$\| s^\alpha D^\alpha s^{-\alpha} h \|_{L^2} \leqslant C \| h \|_{C^\beta}. \tag{4.20}$$

In particular, for any $H \in (0, 1)$, if $h \in C_t^\beta$ for some $\beta > H + 1/2$, $h_0 = 0$, then $K_H^{-1} \in L_t^2$ and there exists a constant $C = C(H, \beta, T)$ such that

$$\| K_H^{-1} h \|_{L^2} \leqslant C \| h \|_{C^\beta}. \tag{4.21}$$

Proof. We have

$$(s^\alpha D^\alpha s^{-\alpha} h)(t) = \Gamma(1-\alpha)^{-1} \Big[h_t + \alpha t^\alpha \int_0^t \frac{t^{-\alpha} h_t - s^{-\alpha} h_s}{(t-s)^{\alpha+1}} \, \mathrm{d}s \Big].$$

Since $h \in C^\beta$, it clearly also belongs to L^2, so we only need to control the term

$$\begin{aligned}
t^\alpha \Big| \int_0^t \frac{t^{-\alpha} h_t - s^{-\alpha} h_s}{(t-s)^{\alpha+1}} \, \mathrm{d}s \Big| &\leqslant t^\alpha \int_0^t \frac{t^{-\alpha} |h_t - h_s| + (s^{-\alpha} - t^{-\alpha}) |h_s|}{(t-s)^{\alpha+1}} \, \mathrm{d}s \\
&\leqslant \| h \|_{C^\beta} \, t^\alpha \int_0^t \frac{t^{-\alpha} (t-s)^\beta + (s^{-\alpha} - t^{-\alpha})}{(t-s)^{\alpha+1}} \, \mathrm{d}s \\
&\lesssim_T \| h \|_{C^\beta} \, t^{-\alpha} \Big[\int_0^1 \frac{1}{(1-u)^{1+\alpha-\beta}} \, \mathrm{d}u + \int_0^1 u^{-\alpha} \frac{(1-u^\alpha)}{(1-u)^{1+\alpha}} \, \mathrm{d}u \Big] \lesssim_T \| h \|_{C^\beta} \, t^{-\alpha}.
\end{aligned}$$

Since $\alpha \in (0, 1/2)$, $t^{-\alpha} \in L_t^2$ and so we deduce that the overall expression belongs to L_t^2, as well as estimate (4.20). Regarding the second statement, the case $H = 1/2$ is straightforward since $K_H^{-1} h = h'$. In the case $H > 1/2$, by the formula for K_H^{-1} combined with estimates (4.19) and (4.20) for the choice $\alpha = H - 1/2$, choosing $\varepsilon > 0$ sufficiently small we have

$$\| K_H^{-1} h \|_{L^2} \lesssim \| h' \|_{C^{H-1/2+\varepsilon}} \lesssim \| h' \|_{C^\beta};$$

the case $H < 1/2$ is analogous. ∎

Remark 4.18. We have given an explicit proof of Lemma 4.17, but a similar (stronger) type of result can be achieved by more abstract arguments. Indeed it follows from the proof of Theorem 5.4 from [48] that $\| s^\alpha D^\alpha (s^{-\alpha} h) \|_{L^2} \sim \| D^\alpha h \|_{L^2}$ and similarly $\| K_H^{-1} h \|_{L^2} \sim \| D^{H+1/2} h \|_{L^2}$; we have already seen that if $h \in C^\beta$ with $\beta > \alpha$ and $h_0 = 0$, then $D^\alpha h$ is a continuous function, so its L^2-norm is trivially finite. The inclusion $C^\beta \subset I^\alpha(L^2)$ is strict and therefore the hypothesis of Lemma 4.17 are non optimal, but they are rather useful when dealing with functions h not of bounded variation.

We can now state a general result on the applicability of Girsanov's transform together with a good control on the density defining $\mathbb{Q}$.

Theorem 4.19. *Let* $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geqslant 0}, \mathbb{P})$ *be a filtered probability space, let* W^H *be an* $\mathcal{F}_t$*-fBm of parameter* $H \in (0,1)$ *and let* h *be an* $\mathcal{F}_t$*-adapted process with trajectories in* C^β_t, $\beta > H + 1/2$, *such that* $h_0 = 0$ *and*

$$\mathbb{E}[\exp(\lambda \|h\|^2_{C^\beta})] < \infty \quad \forall \lambda \in \mathbb{R}. \tag{4.22}$$

Then the Girsanov transform for $\tilde{W}^H = h + W^H$ *is applicable, i.e.,* $\tilde{W}^H$ *is an* $\mathcal{F}_t$*-fBm of parameter* H *under the probability measure* $\mathbb{Q}$ *given by* (4.17). *Moreover, the measures* $\mathbb{Q}$ *and* $\mathbb{P}$ *are equivalent and it holds*

$$\mathbb{E}_{\mathbb{P}}\Big[\Big(\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}}\Big)^n + \Big(\frac{\mathrm{d}\mathbb{P}}{\mathrm{d}\mathbb{Q}}\Big)^n\Big] < \infty \quad \forall n \in \mathbb{N}.$$

Proof. By hypothesis (4.22) and Lemma 4.17 it follows immediately that

$$\mathbb{E}[\exp(\lambda \|K^{-1}h\|^2_{L^2})] < \infty \quad \forall \lambda \in \mathbb{R}.$$

Therefore Novikov's criterion is satisfied and Girsanov's transform is applicable. The proof of second part of the statement follows from classical arguments, but we include it for the sake of completeness. Let us prove integrability of the moments: for any $\alpha \geqslant 1$ it holds

$$\begin{aligned}
\mathbb{E}_{\mathbb{P}}\Big[\Big(\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}}\Big)^\alpha\Big] &= \mathbb{E}_{\mathbb{P}}\Big[\exp\Big(\alpha \int_0^T (K_H^{-1}h)\cdot \mathrm{d}B - \alpha^2 \|K_H^{-1}h\|^2_{L^2} + \Big(\alpha^2 - \frac{\alpha}{2}\Big)\|K_H^{-1}h\|^2_{L^2}\Big)\Big] \\
&\leqslant \mathbb{E}_{\mathbb{P}}\Big[\exp\Big(2\alpha \int_0^T (K_H^{-1}h)\cdot \mathrm{d}B - 2\alpha^2 \|K_H^{-1}h\|^2_{L^2}\Big)\Big]^{1/2} \mathbb{E}_{\mathbb{P}}[\exp((2\alpha^2-\alpha)\|K_H^{-1}h\|^2_{L^2})]^{1/2} \\
&= \mathbb{E}_{\mathbb{P}}[\exp((2\alpha^2-\alpha)\|K_H^{-1}h\|^2_{L^2})]^{1/2} < \infty,
\end{aligned}$$

where in the last passage we used the fact that the integrand in the first term is again a probability density by Novikov's criterion, this time applied to the process $\tilde{h} = 2\alpha h$. Now in order to show that the measures $\mathbb{Q}$ and $\mathbb{P}$ are equivalent, we need to show that the inverse density $\mathrm{d}\mathbb{P}/\mathrm{d}\mathbb{Q}$ is integrable with respect to $\mathbb{Q}$. Again by Girsanov, since we have $W^H = \tilde{W}^H - h$, the inverse density is given by

$$\frac{\mathrm{d}\mathbb{P}}{\mathrm{d}\mathbb{Q}} = \exp\Big(\int_0^T (K_H^{-1}h)(s)\cdot \mathrm{d}\tilde{B}_s - \frac{1}{2}\int_0^T |(K_H^{-1}h)(s)|^2\,\mathrm{d}s\Big),$$

where $\tilde{B}$ denotes the standard Bm associated to $\tilde{W}^H$ such that $\tilde{W}^H_t = \int_0^t K_H(t,s)\,\mathrm{d}\tilde{B}_s$. Since we have

$$\begin{aligned}
\mathbb{E}_{\mathbb{Q}}\Big[\exp\Big(\frac{1}{2}\int_0^T |(K_H^{-1}h)(s)|^2\,\mathrm{d}s\Big)\Big] &= \mathbb{E}_{\mathbb{P}}\Big[\exp\Big(\frac{1}{2}\int_0^T |(K_H^{-1}h)(s)|^2\,\mathrm{d}s\Big)\frac{\mathrm{d}\mathbb{P}}{\mathrm{d}\mathbb{Q}}\Big] \\
&\leqslant \mathbb{E}_{\mathbb{P}}\Big[\exp\Big(\int_0^T |(K_H^{-1}h)(s)|^2\,\mathrm{d}s\Big)\Big]^{1/2}\mathbb{E}_{\mathbb{P}}\Big[\Big(\frac{\mathrm{d}\mathbb{P}}{\mathrm{d}\mathbb{Q}}\Big)^2\Big]^{1/2} < \infty,
\end{aligned}$$

we can conclude, again by applying Novikov, that $\mathrm{d}\mathbb{P}/\mathrm{d}\mathbb{Q}$ is integrable with respect to $\mathbb{Q}$. Reasoning as before, it can be shown that $\mathrm{d}\mathbb{P}/\mathrm{d}\mathbb{Q}$ admits moments of any order with respect to $\mathbb{Q}$, which gives the conclusion. ∎

4.2.3. Path-by-path uniqueness for SDEs driven by additive fBm. Girsanov's theorem allows to construct a probabilistically weak solution of (CP_{x_0}), which we define in the following way.

Definition 4.20. We say that $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geqslant 0}, \mathbb{Q}, W^H_\cdot, X_\cdot)$ is a weak solution of the Cauchy problem

$$X_t = x_0 + \int_0^t b(s, X_s)\,\mathrm{d}s + W^H_t \tag{4.23}$$

if $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geqslant 0}, \mathbb{Q})$ is a filtered probability space, W^H is an $\mathcal{F}_t$-fBm of parameter H under the probability $\mathbb{Q}$ and $\mathbb{Q}$-a.s. the following holds: there exists $\gamma > 1/2$ such that $T^{W^H} b \in C^\gamma_t \operatorname{Lip}_x$, $X \in W^H + C^\gamma$ and X is a solution of (4.23) in the sense of Definition 4.2.

We have given a non classical notion of weak solution, which is well suited when dealing with a distributional b; depending on the context, this is not the only possible definition, see for instance [3] and [24] for different choices.

We are now ready to provide a general principle to establish path-by-path uniqueness.

Lemma 4.21. *Let W^H be an $\mathcal{F}_t$-fBm of parameter H on $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geqslant 0}, \mathbb{P})$ and let $x_0 \in \mathbb{R}^d$. Suppose that*

(1) *b is a distributional drift such that, for some $\gamma > 1/2$, $T^{W^H} b \in C^\gamma_t \operatorname{Lip}_x$ $\mathbb{P}$-a.s.,*

(2) *Girsanov is applicable to $W^H - h$ for $h_\cdot = \int_0^\cdot b(s, x_0 + W^H) = T^{W^H}(\cdot, x_0)$.*

Then path-by-path uniqueness for (CP_{x_0}) holds.

Proof. Consider $\gamma > 1/2$ as in the assumption and the set $\mathcal{A}_\gamma$ defined as in (4.10); by Theorem 4.8, in order to conclude it is enough to show that $\mu^H(\mathcal{A}_\gamma) = 1$. By hypothesis, the first half of the statement defining $\mathcal{A}_\gamma$ is already satisfied on a set of full probability, so we only need to concentrate on the second half. By the definition of h, the process $X = x_0 + W^H$ satisfies

$$X_t = x_0 + \int_0^t T^{W^H}(\mathrm{d}s, x_0) + [W^H_t - h_t] =: x_0 + \int_0^t T^X(\mathrm{d}s, 0) + \tilde{W}^H_t; \tag{4.24}$$

by hypothesis, Girsanov's theorem is applicable, so we can construct a new probability measure $\mathbb{Q}$ which is absolutely continuous with respect to $\mathbb{P}$ such that $\tilde{W}^H$ is an $\mathcal{F}_t$-fBm under $\mathbb{Q}$. Observe that $\mathbb{P}$-a.s. $T^X b = \tau^{x_0} T^{W^H} b \in C^\gamma_t \operatorname{Lip}_x$ and so $\mathbb{P}$-a.s. the difference $X_\cdot - \tilde{W}^H_\cdot = x_0 + T^X(\cdot, 0) \in C^\gamma$ (if $T^X b \in C^\gamma_t \operatorname{Lip}_x$, then it also belongs to $C^0_x C^\gamma_t$); then by Lemma 4.9, on a set of full measure $\mathbb{P}$ equation (4.24) is equivalent to

$$X_t = x_0 + \int_0^t T^W(\mathrm{d}s, X_s - \tilde{W}^H_s) + \tilde{W}^H_t,$$

and so X is $\mathbb{P}$-a.s. a solution to (CP_{x_0}) in the sense of Definition 4.2. Since $\mathbb{Q} \ll \mathbb{P}$, all the above statements also hold on a set of $\mathbb{Q}$-full measure. But then since $\tilde{W}^H$ has law μ^H

under $\mathbb{Q}$, we obtain

$$\begin{aligned}\mu^H(\mathcal{A}_\gamma) &= \mathbb{Q}\big(T^{\tilde W^H} b \in C_t^\gamma \operatorname{Lip}_x,\ \exists \text{ a solution } x \in \tilde W^H + C^\gamma \text{ s.t. } T^x b \in C_t^\gamma \operatorname{Lip}_x\big)\\ &\geqslant \mathbb{Q}\big(T^{\tilde W^H} b \in C_t^\gamma \operatorname{Lip}_x,\ X \text{ is a solution to } (\mathrm{CP}_{x_0}) \text{ s.t. } T^X b \in C_t^\gamma \operatorname{Lip}_x\big) = 1,\end{aligned}$$

which gives the conclusion. ∎

Remark 4.22. We cannot apply directly the Yamada–Watanabe theorem to deduce existence of a strong solution under the assumptions of Lemma 4.21, because our path-by-path uniqueness statement holds only in the class $w + C^\gamma$ and not in the class of all possible continuous paths (although in the case of continuous b the two classes coincide). There is however a more direct way to show that the path-by-path unique solution X is adapted to the filtration generated by W^H. Consider a sequence $\varepsilon_n \downarrow 0$ and $b^n := \rho^{\varepsilon_n} * b$, where as usual $\{\rho^\varepsilon\}_{\varepsilon>0}$ is a sequence of spatial mollifiers, and consider X^n solution to

$$\mathrm{d}X_t^n = b^n(t, X_t^n)\,\mathrm{d}t + \mathrm{d}W^H;$$

by classical theory X^n is unique and adapted to the filtration generated by W^H. Then by Theorem 4.8 (possibly combined with Lemma 3.9), $\mathbb{P}$-a.s. $X_\cdot^n \to X_\cdot$ in C^γ, which implies that X is adapted as well and thus a strong solution.

All the results obtained so far are of abstract nature. Now we are going to show how to apply them to establish path-by-path uniqueness for (CP_{x_0}) in our context. In particular, Theorem 4.13 is a direct consequence of the following more general result.

Theorem 4.23. *Let b be a given drift, $H \in (0,1)$. Assume one of the following*:

- *if $H > 1/2$, then there exist $\alpha > 1 - 1/(2H) > 0$ and $\beta > H - 1/2 > 0$ such that $b \in C_t^0 C_x^\alpha$ and*

$$|b(t,x) - b(s,y)| \leqslant C(|x-y|^\alpha + |t-s|^\beta) \quad \textit{for all } s,t \in [0,T],\ x,y \in \mathbb{R}^d;$$

- *if $H \leqslant 1/2$, then $b \in L_t^\infty C_x^\alpha$ for $\alpha > 1 - 1/(2H)$, such that b has compact support, uniformly in time*; *here $\alpha < 0$ is allowed.*

Then, for any $x_0 \in \mathbb{R}^d$, path-by-path uniqueness holds for (CP_{x_0}).

Proof. In both cases, in order to conclude, we need to show that we can apply Lemma 4.21 to the process $h_\cdot = \int_0^\cdot b(s, x_0 + W_s^H)\,\mathrm{d}s$; in order to do so, we will check that the conditions of Theorem 4.19 are satisfied. Up to shifting b, we can assume without loss of generality $x_0 = 0$.

Let $H > 1/2$. Then by the hypothesis $b \in C_t^0 C_x^\alpha$ and Theorem 3.16, we know that $T^{W^H} \in C_t^\gamma C_x^{1+\varepsilon}$ (at least locally) for some $\gamma > 1/2$ and $\varepsilon > 0$; the process h belongs to $C_t^{H+1/2+\varepsilon}$ if and only if the map $t \mapsto b(t, W_t^H) \in C_t^{H-1/2+\varepsilon}$. Recall that for any $\gamma < H$, $W^H \in C_t^\gamma$; then by the hypothesis it holds

$$|b(t,W_t^H) - b(s,W_s^H)| \leqslant C(|t-s|^\beta + |W_t^H - W_s^H|^\alpha) \leqslant C(|t-s|^\beta + [\![W^H]\!]_\gamma^\alpha |t-s|^{\alpha\gamma})$$

and so we can find $\varepsilon > 0$ small enough such that $\gamma = H - \varepsilon$ and

$$[\![b(\cdot, W_\cdot^H)]\!]_{C^{H+1/2+\varepsilon}} \lesssim 1 + [\![W^H]\!]_{C^{H-\varepsilon}}^\alpha.$$

As the exponent $\alpha < 1$, by Fernique's theorem we deduce that

$$\mathbb{E}[\exp(\lambda \|h\|^2_{C^{H+1/2-\varepsilon}})] \lesssim \mathbb{E}[\exp(\lambda C [\![W^H]\!]^{2\alpha}_{C^{H-\varepsilon}})] < \infty \quad \forall \lambda \in \mathbb{R}.$$

Consider now the case $H < 1/2$. By Theorem 3.16 (as the support of b is compact uniformly in time, we have the embedding $C^\alpha \hookrightarrow B^\alpha_{p,p}$ for any $p < \infty$), we know that

$$T^{W^H} b \in C^\gamma_t C^{\alpha+1/2H-\varepsilon}_x \hookrightarrow C^\gamma_t C^{1+\varepsilon}_x$$

for some $\gamma > 0$ and $\varepsilon > 0$ sufficiently small, therefore the process $h_t = T^{W^H} b(t,0)$ is a well defined element of C^γ_t. We now want to show that it actually belongs to $C^{H+1/2+\varepsilon}_t$; we can do so by interpolation, using the fact that T^{W^H} has higher spatial regularity. Indeed, by properties of the averaging operator $T^{W^H} b \in \mathrm{Lip}_t\, C^\alpha_x$, and so for any $\theta \in (0,1)$ it holds

$$\|h_\cdot\|_{C^{1-\theta/2}} \leqslant \|T^{W^H} b\|_{C^{1-\theta/2}_t C^\beta_x} \leqslant \|T^{W^H} b\|^{1-\theta}_{\mathrm{Lip}_t C^\alpha_x} \|T^{W^H} b\|^\theta_{C^{1/2}_t C^{\alpha+1/(2H)-\varepsilon}_x},$$

where $\beta = (1-\theta)\alpha + \theta(\alpha + (2H)^{-1} - \varepsilon)$ and thanks to the hypothesis we can choose $\theta \in (0,1)$ such that

$$\begin{cases} \beta = \alpha + \theta/(2H) - \varepsilon\theta > 0 \\ 1 - \theta/2 > H + 1/2 \end{cases} \iff \alpha - \varepsilon\theta > -\frac{\theta}{2H} > 1 - \frac{1}{2H}.$$

For this choice of θ therefore we obtain

$$\|h\|^2_{C^{H+1/2+}} \lesssim \|T^{W^H} b\|^{2\theta}_{C^{1/2}_t C^{\alpha+1/(2H)-\varepsilon}_x}$$

and since the exponent $2\theta < 2$, and we have exponential integrability for the term on the right-hand side by Theorem 3.16, we get the conclusion. ∎

In the regime $H > 1/2$, the hypothesis required on b is the same as in [45], although therein path-wise uniqueness is shown only in the case $d = 1$, while here we obtain path-by-path uniqueness in any dimension. In the case $H = 1/2$, we can allow $b \in L^\infty_t C^\alpha_x$ for any $\alpha > 0$; this result is comparable to the one from [18], in which sharper estimates allow to reach $b \in L^\infty_{t,x}$, see also [52,53] for further extensions. Observe that in the regime $H < 1/2$ we can allow b to be only distributional; in this case, we recover the results from [12]. Unfortunately, the original proof from [12] is wrong, due to an incorrect version of the formula defining K_H^{-1} (see the formula for H^n just before Lemma 4.8 therein), which is why we have decided to give an alternative proof rather than directly invoking the results from [12].

The driving principle given by Lemma 4.21 is fairly general and can be applied under different hypothesis on b, especially when we combine it with Theorems 3.12 and 3.16.

Corollary 4.24. *Let $H < 1/2$ and $b \in L^q_t B^\alpha_{p,p}$ with $(q,p) \in [2,\infty)^2$, $\alpha < 0$ such that*

$$\frac{1}{q} + H\Big(\frac{d}{p} - \alpha\Big) < \frac{1}{2} - H. \tag{4.25}$$

Then for any $x_0 \in \mathbb{R}^d$, *path-by-path uniqueness for* (CP_{x_0}) *under* μ^H *holds. A similar statement holds for* $b \in L^q_t L^p_x$ *with*

$$\frac{1}{q} + H\frac{d}{p} < \frac{1}{2} - H. \tag{4.26}$$

Proof. It follows from hypothesis (4.25), combined with Theorem 3.16 and the Besov embeddings $B^{\alpha+s}_{p,p} \hookrightarrow C^{\alpha+s-d/p}$, that we can choose s satisfying (3.16) such that $T^{W^H} b \in C^\gamma_t C^1_x$ for some $\gamma > 1/2$. As before, we can now assume $x_0 = 0$ and it remains to show that the process $h_t = T^{W^H} b(t,0) \in C^\beta_t$ for some $\beta > H + 1/2$ and satisfies integrability conditions like those of Theorem 4.19. By the properties of the averaging operator, on a set of full probability it holds

$$T^{W^H} b \in C^{1-1/q}_t B^\alpha_{p,p} \cap C^{1/2}_t B^{\alpha+s}_{p,p}$$

for any s such that $Hs + 1/q < 1/q$. By interpolation, for any $\theta \in (0,1)$, it holds

$$T^{W^H} b \in C_t^{(1-\frac{1}{q})(1-\theta)+\frac{\theta}{2}} B^{(1-\theta\alpha)+\theta(\alpha+s)}_{p,p} \hookrightarrow C_t^{(1-\frac{1}{q})(1-\theta)+\frac{\theta}{2}2} C^{(1-\theta)\alpha+\theta(\alpha+s)-\frac{d}{p}}.$$

In order to deduce that $T^{W^H} b(\cdot,0) \in C^\beta_t$ with $\beta > H + 1/2$, we need to find parameters $s > 0$ and $\theta \in (0,1)$ such that

$$\begin{cases} Hs + 1/q < 1/2, \\ (1-1/q)(1-\theta) + \theta/2 > H + 1/2, \\ (1-\theta)\alpha + \theta(\alpha+s) - d/p > 0. \end{cases}$$

A few algebraic manipulations show that the above system is equivalent to condition (4.25); from interpolation we then obtain, for $\beta = (1-1/q)(1-\theta) + \theta/2$ as above,

$$\|h\|_{C^\beta} \lesssim \|T^{W^H} b\|^\theta_{C^{1/2}_t B^{\alpha+s}_{p,p}}$$

and since the parameter $\theta \in (0,1)$, we deduce that $\|h\|_{C^\beta}$ satisfies (4.22).

In the case $b \in L^q_t L^p_x$, using the embedding $L^p_x \hookrightarrow B^{-\varepsilon}_{p,p}$ for any $\varepsilon > 0$ (see Appendix A.2) and applying the previous result for ε sufficiently small we get the conclusion. ■

In the case $b \in L^q_t L^p_x$, it was already shown in [38] that pathwise uniqueness holds. Here we have strengthened the result to path-by-path uniqueness. The case $b \in L^q_t B^\alpha_{p,p}$ with $\alpha < 0$, to the best of our knowledge, has not been considered in the literature so far. Condition (4.25) actually holds also in the regime $\alpha > 0$, but this is not particularly interesting as one can use fractional Sobolev embeddings (see [19]) to deduce $L^q_t B^\alpha_{p,p} \hookrightarrow L^q_t L^{p^*}_x$ with

$$\frac{1}{p^*} = \frac{1}{p} - \frac{\alpha}{d}$$

and then reduce it to the case (4.26).

Remark 4.25. The guiding principle of Lemma 4.21 is fairly general, but there are situations in which it is possible to establish path-by-path uniqueness even if Girsanov's theorem is not applicable (or at least we are currently not able to find suitable estimates in order to apply it). Consider for instance the case of $H > 1/2$ and $b \in L^\infty_t C^\alpha_x$ for $\alpha \in (0,1)$ such that

$$\alpha > \frac{3}{2} - \frac{1}{2H};$$

observe that the condition is non trivial for every $H \in (1/2)$. Then by Theorem 3.16 (possibly after a localisation procedure) $T^{W^H} b \in C^\gamma_t \,\mathrm{Lip}_x$ (at least locally) and so by Theorem 4.6 path-by-path uniqueness holds for the whole ODE. However, lack of continuity in time of b prevents us from applying Girsanov.

4.3. Regularity of the flow

4.3.1. Variational formula for flow of diffeomorphisms. It follows from Theorem 4.6 that, if b and $T^w b$ satisfy the regularity assumptions, the solution map $(\theta_0, t) \mapsto \theta_t$ is Lipschitz in space, uniformly in time (more precisely, it follows from (4.5) and (4.6) that it is $C^\gamma_t \,\mathrm{Lip}_{\mathrm{loc}}$). However we cannot yet talk about a flow, as we have not shown the invertibility of the solution map, nor the flow property; this is accomplished by the following two lemmas.

Lemma 4.26. *Let $T^w b \in C^\gamma_t C^\beta_x$ and $\theta \in C^\alpha_t$ such that $\gamma + \beta\alpha > 1$. Then setting $\tilde{w}_t = w_{T-t}$, $\tilde{b}_t = b_{T-t}$, it holds*

$$\int_0^t T^w b(\mathrm{d}s, \theta_s) = -\int_{T-t}^T T^{\tilde{w}} \tilde{b}(\mathrm{d}s, \theta_{T-s}). \tag{4.27}$$

In particular, if θ is a solution of the YDE

$$\theta_t = \theta_0 + \int_0^t T^w b(\mathrm{d}s, \theta_s),$$

then $\tilde{\theta}_t = \theta_{T-t}$ satisfies the time-reversed YDE

$$\tilde{\theta}_t = \tilde{\theta}_0 + \int_0^t T^{\tilde{w}} \tilde{b}(\mathrm{d}s, \tilde{\theta}_s).$$

Proof. Let Π be a partition of $[0,t]$ given by $0 = t_0 < t_1 < \cdots < t_n = t$ and define $\tilde{t}_i = T - t_i$, which defines a partition of $[T-t, T]$ (up to the fact that it is decreasing with respect to i); it holds

$$\sum_i T^w b_{t_i,t_{i+1}}(\theta_{t_i}) = \sum_i T^w b_{T-\tilde{t}_i, T-\tilde{t}_{i+1}}(\theta_{T-\tilde{t}_i}) = -\sum_i T^w b_{T-\tilde{t}_{i+1}, T-\tilde{t}_i}(\theta_{T-\tilde{t}_{i+1}}) + J,$$

where the remainder term J satisfies

$$\begin{aligned} |J| &= \Big| \sum_i [T^w b_{T-\tilde{t}_{i+1}, T-\tilde{t}_i}(\theta_{T-\tilde{t}_i+1}) - T^w b_{T-\tilde{t}_{i+1}, T-\tilde{t}_i}(\theta_{T-\tilde{t}_i})] \Big| \\ &\leqslant \|T^w b\|_{C^\gamma_t C^\beta_x} \|\theta\|_{C^\alpha_t} \sum_i |t_{i+1} - t_i|^{\alpha+\beta\gamma} \lesssim \|\Pi\|^{\alpha+\beta\gamma-1} = o(\|\Pi\|). \end{aligned}$$

By basic properties of the averaging operator we have $T^w b_{T-t,T-s}(x) = T^{\tilde w}\tilde b_{s,t}(x)$, and so overall we obtain

$$\sum_i T^w b_{t_i,t_{i+1}}(\theta_{t_i}) = -\sum_i T^{\tilde w}\tilde b_{\tilde t_i,\tilde t_{i+1}}(\theta_{T-\tilde t_{i+1}}) + o(\|\Pi\|)$$

Taking a sequence of partitions Π_N such that $\|\Pi_N\|\to 0$ and taking the limits on both sides we obtain the first statement. Regarding the second statement, if θ is a solution of the YDE, then by (4.27) for any $t\in[0,T]$ it holds

$$\theta_{T-t}-\theta_T = -\int_{T-t}^T T^w b(\mathrm{d}s,\theta_s) = \int_0^t T^{\tilde w}\tilde b(\mathrm{d}s,\theta_{t-s})$$

which implies the conclusion. ∎

Similar arguments also provide the following lemma, whose proof is therefore omitted.

Lemma 4.27. *Let $T^w b\in C^\gamma_t C^\nu_x$ with $\gamma(1+\nu)>1$ and let θ be a solution of*

$$\theta_t = \theta_s + \int_s^t T^w b(\mathrm{d}r,\theta_r)\quad \forall\, t\in[s,T].$$

Then setting $\tilde\theta_t = \theta_{s+t}$, $\tilde w_t = w_{s+t}$ and $\tilde b_t = b_{s+t}$, it holds

$$\tilde\theta_t = \tilde\theta_0 + \int_0^t T^{\tilde w}\tilde b(\mathrm{d}r,\tilde\theta_r)\quad \forall t\in[0,T-s].$$

We are now ready to provide sufficient conditions for the existence of a Lipschitz flow.

Theorem 4.28. *Let b, $T^w b$ satisfy the assumptions of Theorem 4.6. Then the YDE admits a locally Lipschitz flow. Namely, setting $\Delta_T := \{(s,t)\in[0,T]^2 : s\leqslant t\}$, there exists a map $\Phi\colon \Delta_T\times\mathbb{R}^d\to\mathbb{R}^d$ with the following properties:*

- (i) *$\Phi(t,t,x)=x$ for all $t\in[0,T]$ and $x\in\mathbb{R}^d$;*
- (ii) *$\Phi(s,\cdot,x)\in C^\gamma([s,T];\mathbb{R}^d)$ for all $s\in[0,T]$ and $x\in\mathbb{R}^d$;*
- (iii) *for all $(s,t,x)\in\Delta_T\times\mathbb{R}^d$ it satisfies*

$$\Phi(s,t,x) = x + \int_s^t T^w b(\mathrm{d}r,\Phi(s,r,x));$$

- (iv) *for all $0\leqslant s\leqslant u\leqslant t\leqslant T$ and $x\in\mathbb{R}^d$, $\Phi(u,t,\Phi(s,u,x))=\Phi(s,t,x)$;*
- (v) *there exists C depending on γ, T and $\|T^w b\|_{C^\gamma_t C^2_x}$ (respectively, $\|T^w b\|_{C^\gamma_t C^{3/2}_x}$ $\vee\|b\|_{L^\infty_{t,x}}$) such that*

$$|\Phi(s,t,x)-\Phi(s,t,y)|\leqslant C|t-s|^\gamma|x-y|\quad \textit{for all } (s,t)\in\Delta_T,\ x,y\in\mathbb{R}^d;$$

moreover, $\Phi(s,t,\cdot)$ as a function from $\mathbb{R}^d$ to itself is invertible and the same inequality holds for its inverse, which we denote by $\psi(s,t,\cdot)=\Phi(s,t,\cdot)^{-1}$.

Proof. The proof is a straightforward application of Theorem 4.6 and Lemmata 4.26 and 4.27. In both cases of time reversal and translation we have $\|T^{\tilde{w}}\tilde{b}\|_{C^\gamma C^2} \lesssim \|T^w b\|_{C^\gamma C^2}$ (same for $\|\cdot\|_{C^\gamma C^{3/2}}$ and $\|\cdot\|_{L^\infty_{t,x}}$), so that uniqueness holds also for the reversed/translated YDE, with the same continuity estimates; this provides respectively invertibility of the solution map and flow property. ∎

Further estimates for Φ are available, since Theorem 4.6 actually implies that

$$\llbracket \Phi(s,\cdot,x)\rrbracket_{C^\gamma([s,T];\mathbb{R}^d)} \lesssim 1, \quad \llbracket \Phi(s,\cdot,x)-\Phi(s,\cdot,y)\rrbracket_{C^\gamma([s,T];\mathbb{R}^d)} \lesssim |x-y|$$

uniformly in $s \in [0,T]$, $x, y \in \mathbb{R}^d$.

Let us denote by Φ_t the map $\Phi_t(x) = \Phi(0,t,x)$; from now on we are only going to consider the map $\Phi_\cdot = \Phi_t(x)$, which by an abuse of notation and language, will be just denoted by Φ and referred to as the flow of the YDE. This is just to keep the notation simple and indeed all the proofs below can be easily adapted to the whole flow $\Phi(s,t,x)$.

We will keep using the incremental notation $\Phi_{s,t}(x) = \Phi_t(x) - \Phi_s(x)$; it follows from the above estimates that $\Phi \in C^\gamma_t \mathrm{Lip}_{\mathrm{loc}}$, since

$$|\Phi_{s,t}(x)-\Phi_{s,t}(y)| \leqslant |t-s|^\gamma \llbracket \Phi(s,\cdot,x)-\Phi(s,\cdot,y)\rrbracket_{C^\gamma} \lesssim |t-s|^\gamma |x-y|.$$

Similarly, we define $\psi_t(x) = \psi(0,t,x)$, so that $\psi_t = \Phi_t^{-1}$ as a map from $\mathbb{R}^d$ to itself.

We now state a specialised version of Theorem 4.6 which is quite useful for practical purposes, as it clearly identifies a way to approximate the flow associated to $T^w b$, which by the YDE formulation is well defined when b is only a distribution, by means of more regular flows, associated to drifts b^ε for which also the ODE interpretation is meaningful.

Lemma 4.29. *Let b, $T^w b$ satisfy the hypothesis of Theorem 4.6 and let $\{\rho^\varepsilon\}_{\varepsilon>0}$ be a family of spatial mollifiers, $b^\varepsilon = \rho^\varepsilon * b$. Then b^ε satisfies the hypothesis of Theorem 4.6 for any $\varepsilon > 0$. Denote by Φ^ε and Φ the flows associated respectively to b^ε and b. Then $\Phi^\varepsilon \to \Phi$ uniformly on compact sets; more precisely, for any $\tilde{\gamma} < \gamma$ and any fixed $R > 0$ it holds*

$$\lim_{\varepsilon\to 0} \sup_{x\in B_R} \|\Phi(\cdot,x)-\Phi^\varepsilon(\cdot,x)\|_{C^{\tilde{\gamma}}} = 0. \tag{4.28}$$

In the case $b \in L^\infty_{t,x}$, the above convergence actually holds for any $\tilde{\gamma} < 1$.

Proof. We only prove the statement in the case $T^w b \in C^\gamma_t C^2_x$, the other one being almost identical. By the properties of mollifiers it holds $T^w b^\varepsilon = (T^w b)^\varepsilon$, so that $\|T^w b^\varepsilon\|_{C^\gamma_t C^2_x} \leqslant \|T^w b\|_{C^\gamma_t C^2_x}$ for all $\varepsilon > 0$, thus the hypothesis of Theorem 4.6 are satisfied uniformly in $\varepsilon > 0$. Once we fix $R > 0$, by the a priori estimates from Theorem 4.4 we have a uniform bound of the form

$$\sup_{\varepsilon>0} \sup_{x\in B_R} \|\Phi^\varepsilon(\cdot,x)\|_{C^\gamma} \leqslant C < \infty;$$

in particular we can localise $T^w b$ and $T^w b^\varepsilon$ in such a way that they all have support contained in a sufficiently big ball (say for instance B_{2R}) in such a way that for $x \in B_R$, $\Phi(\cdot,x)$ and $\Phi^\varepsilon(\cdot,x)$ are not affected by it. Now take any $\tilde{\gamma} \in (1/2,\gamma)$, then by (4.5) in order to conclude it is enough to show that $T^w b^\varepsilon \to T^w b$ locally in $C^{\tilde{\gamma}}_t \mathrm{Lip}_x$; but this is an immediate consequence of Lemma 3.9. ∎

From now on we will adopt the following notation: whenever all the Young integrals involved are well defined, we write

$$\int_0^t f_s A(\mathrm{d}s,\theta_s) := \int_0^t f_s \,\mathrm{d}\Big(\int_0^\cdot A(\mathrm{d}r,\theta_r)\Big),$$

so that in particular, whenever φ is regular enough for $T^w\varphi$ to make sense both as a Young integral and a Lebesgue integral, it holds

$$\int_0^t f_s T^w\varphi(\mathrm{d}s,\theta_s) = \int_0^t f_s \varphi(s,\theta_s+w_s)\,\mathrm{d}s.$$

We are now ready to further improve the regularity of the flow Φ and provide a variational equation for $D_x\Phi$, as well as an expression for its Jacobian. In the case $A = T^w b \in C^\gamma_t C^2_x$ a similar result was proved in [33], Section 3.3; our derivation is of different nature and based on approximating b by more regular b^ε, for which standard ODE theory applies. The case $T^w b \in C^\gamma_t C^{3/2}_x$ appears to be new.

Theorem 4.30. *Let b, $T^w b$ satisfy the hypothesis of Theorem 4.6. Then Φ associated to b is a flow of diffeomorphisms and belongs to $C^\gamma_t C^1_{\mathrm{loc}}$; it satisfies the variational equation*

$$D_x\Phi_t(x) = I + \int_0^t D_x\Phi_s(x)\circ T^w D_x b(\mathrm{d}r,\Phi_r(x)) \tag{4.29}$$

which is meaningful as a YDE; here, $\circ$ denotes the matrix-type product given by $A\circ B = BA$.

The Jacobian $J\Phi_t(x) = \det(D_x\Phi_t(x))$ satisfies the identity

$$J\Phi_t(x) = \exp\Big(\int_0^t \operatorname{div} T^w b(\mathrm{d}s,\Phi_s(x))\Big) \tag{4.30}$$

and there exists a positive constant C depending on γ, T and $\|T^w b\|_{C^\gamma C^2}$ (respectively, $\|T^w b\|_{C^\gamma C^{3/2}}\vee\|b\|_{L^\infty}$) such that

$$C^{-1}\leqslant J\Phi_t(x)\leqslant C \quad \forall (t,x)\in[0,T]\times\mathbb{R}^d.$$

Proof. As before, to avoid repetitions we give a detailed proof only in the case $T^w b \in C^\gamma_t C^2_x$; we provide in the end the main differences of the proof in the case $b \in L^\infty_{t,x}$, $T^w b \in C^\gamma_t C^{3/2}_x$.

We divide the proof in several steps, but the main idea is the following: in the case of spatially smooth b, the result is just a reformulation of the standard ODE results; in the general case we can recover the result by reasoning by approximation with the help of Lemma 4.29.

Step 1: *Proof in the case of regular b.* Let us first assume in addition that $b \in L^q_t C^2_x$ for some $q > 2$; then in this case we know that the YDE formulation is equivalent to the ODE one, so that the flow Φ associated to b satisfies

$$\Phi_t(x) = x + \int_0^t b(s,\Phi_s(x)+w_s)\,\mathrm{d}s.$$

Moreover, by standard ODE theory we have the variational equation

$$\begin{aligned} D_x\Phi_t(x) &= I + \int_0^t D_x\Phi_s(x) \circ D_x b(s, \Phi_s(x) + w_s)\, \mathrm{d}s \\ &= I + \int_0^t D_x\Phi_s(x) \circ \frac{\mathrm{d}}{\mathrm{d}s}\Big(\int_0^s D_x b(r, \Phi_r(x) + w_r)\mathrm{d}r \Big) \\ &= I + \int_0^t D_x\Phi_s(x) \circ T^w D_x b(\mathrm{d}r, \Phi_r(x)). \end{aligned}$$

The term in the last line now makes perfectly sense as a Young integral, as the term

$$\int_0^\cdot T^w Db(\mathrm{d}r, \Phi_r(x))$$

is a well defined C^γ_t map for $\gamma > 1/2$, since $b \in L^q_t C^1_x$, proving the first part of the claim.

Step 2: *Approximation and characterisation of the limit as $\varepsilon \to 0$ of $D_x\Phi^\varepsilon$.* Consider a sequence $T^w b^\varepsilon$, Φ^ε defined by spatial mollification as in Lemma 4.29. By Step 1, for any $\varepsilon > 0$, $D_x\Phi^\varepsilon$ satisfies the variational equation, which for fixed x is a linear YDE in the unknown $D_x\Phi^\varepsilon(\cdot, x)$ with drift $\int_0^\cdot T^w D_x b^\varepsilon(\mathrm{d}r, \Phi^\varepsilon_r(x))$; thanks to the a priori bounds given by Theorem 4.4, which for fixed x are uniform in ε, we have the estimate

$$\Big\| \int_0^\cdot T^w Db^\varepsilon(\mathrm{d}r, \Phi^\varepsilon_r(x)) \Big\|_{C^\gamma} \lesssim \|T^w Db^\varepsilon\|_{C^\gamma \mathrm{Lip}_x} \|\Phi^\varepsilon(\cdot, x)\|_{C^\gamma} \lesssim \|T^w b\|_{C^\gamma C^2}$$

which implies by Proposition A.2 in Appendix A.1 that for fixed x we have the uniform estimate

$$\sup_{\varepsilon>0} \|D_x\Phi^\varepsilon(\cdot, x)\|_{C^\gamma} < \infty.$$

As in the proof of Lemma 4.29, for any $\delta > 0$ we have $T^w Db^\varepsilon \to T^w Db$ locally in $C^{\gamma-\delta}_t C^{2-\delta}_x$, as well as $\Phi^\varepsilon(\cdot, x) \to \Phi(\cdot, x)$ in $C^{\gamma-\delta}_t$. Thus choosing δ sufficiently small such that $(\gamma - \delta)(2 - \delta) > 1$, by the continuity of the nonlinear Young integral it holds

$$\int_0^\cdot T^w Db^\varepsilon(\mathrm{d}r, \Phi^\varepsilon_r(x)) \to \int_0^\cdot T^w Db(\mathrm{d}r, \Phi_r(x)) \quad \text{in } C^{\gamma-\delta}_t.$$

By the a priori estimates on $D_x\Phi^\varepsilon(\cdot, x)$, we can extract a subsequence converging to a limit in C^β for any $1/2 < \beta < \gamma$; let us denote this limit by $g(\cdot, x)$ (the notation will be clear in a second). By Step 1, $D_x\Phi^\varepsilon$ satisfy variational equations with drifts $\int_0^\cdot T^w D_x b^\varepsilon(\mathrm{d}r, \Phi^\varepsilon(r, x)) \to \int_0^\cdot T^w D_x b(\mathrm{d}r, \Phi(r, x))$, which implies that $g(\cdot, x)$ must satisfy the linear YDE

$$g(t, x) = I + \int_0^t g(s, x) \circ T^w D_x b(\mathrm{d}r, \Phi_r(x)).$$

But the solution to this linear equation unique, thus so is the limit of any subsequence we can extract, showing that the whole sequence $\{D_x\Phi^\varepsilon(\cdot, x)\}_{\varepsilon>0}$ converges to such $g(\cdot, x)$. The reasoning holds for any $x \in \mathbb{R}^d$.

Step 3: *Continuity of the map* $(t,x) \mapsto g(t,x)$. This step is very similar to the previous one, so we only sketch it. Continuity in t is clear, we only need to prove continuity in x; by the continuity of the flow, for any sequence $x_n \to x$ we have $\Phi(\cdot,x_n) \to \Phi(\cdot,x)$ in $C^{\gamma-\delta}$ for any $\delta > 0$ and since all x_n lie in a bounded ball, we have uniform estimate both on $\int_0^\cdot T^w Db(\mathrm{d}r, \Phi_r(x_n))$ and $g(\cdot,x_n)$. Therefore by the usual compactness argument we deduce that $g(\cdot,x_n)$ converge in $C^{\gamma-\delta}$ to the unique solution of the YDE associated to $\int_0^\cdot T^w Db(\mathrm{d}r, \Phi_r(x))$, namely $g(\cdot,x)$.

Step 4: *Flow of diffeomorphisms.* We know that for any $R > 0$, the flows Φ^ε are spatially Lipschitz in B_R, uniformly in $[0,T]$ and $\varepsilon > 0$, and that they converge uniformly on compact sets to Φ, while their spatial derivatives $D_x\Phi^\varepsilon$ converge to the continuous function $g(t,x)$. Therefore we deduce that $g(t,x) = D_x\Phi_t(x)$, thus showing that Φ is C^1 in space, uniformly in time; moreover by construction g is the unique solution to the variational equation (4.29). The reasoning applies to $\psi = \Phi^{-1}$ as well, as it can be represented through the flow associated to the time reversed drift $T^{\tilde{w}}\tilde{b}$, which enjoys the same regularity as $T^w b$.

Step 5: *Jacobian.* As before, let us first assume b spatially smooth. Then by standard ODE theory it holds

$$J\Phi_t(x) = \exp\Big(\int_0^t \operatorname{div} b(s, \Phi_s(x) + w_s)\,\mathrm{d}s\Big) = \exp\Big(\int_0^t \operatorname{div} T^w b(\mathrm{d}s, \Phi_s(x))\Big),$$

which gives equation (4.30) in this case. The general case is accomplished as above by an approximation procedure, using the continuity of Young integrals. Regarding the bound on $J\Phi$, by point (4) of Theorem 4.1 combined with the a priori estimates on Φ, we obtain

$$J\Phi_t(x) \lesssim \|\operatorname{div} T^w b\|_{C^\gamma \mathrm{Lip}} \left(1 + [\![\Phi(\cdot,x)]\!]_{C^\gamma}\right) \leqslant C$$

giving the upper bound; lower bound follows from $(J\Phi_t(x))^{-1} = J\psi_t(\Phi_t(x)) \leqslant C$.

Step 6: *Differences in the case* $b \in L^\infty_{t,x}$ *with* $T^w b \in C^\gamma_t C^{3/2}_x$. The proof in this case goes along the exact same lines, with only slightly different regularity estimates. In this case we know that $\Phi(\cdot,x)$ is Lipschitz with $[\![\Phi(\cdot,x)]\!]_{\mathrm{Lip}} \leqslant \|b\|_{L^\infty}$ for all $x \in \mathbb{R}^d$, and so the drift associated to the variational equation is controlled by

$$\Big\|\int_0^\cdot T^w Db^\varepsilon(\mathrm{d}r, \Phi^\varepsilon(r,x))\Big\|_{C^\gamma} \lesssim \|T^w Db^\varepsilon\|_{C^\gamma C^{1/2}} [\![\Phi^\varepsilon(\cdot,x)]\!]_{\mathrm{Lip}} \leqslant \|T^w b\|_{C^\gamma C^{3/2}} \|b\|_{L^\infty}.$$

Moreover, by Lemma 4.29, we now have $\Phi^\varepsilon(\cdot,x) \to \Phi(\cdot,x)$ in $C^{1-\delta}$ for all $\delta > 0$ and so all the reasonings related to compactness and continuity of Young integrals still work. A similar reasoning goes for equation (4.30) and the two-sided estimates for $J(t,x)$. ■

Remark 4.31. A closer look at the proof shows that the result can be further generalised to include the case of $T^w b \in C^\gamma_t C^\nu_x$ with $b \in L^p_t L^\infty_x$, under the conditions $\gamma > 1/2$ and $\nu \geqslant 1 + q/2$, q being the conjugate of p, i.e., $1/p + 1/q = 1$.

Remark 4.32. Recall that in the case of spatially smooth b, differentiating the relation $\psi_t(\Phi_t(x)) = x$ with respect to t, one obtains that ψ satisfies the PDE

$$\partial_t\psi_t(x) + D_x\psi_t(x) b(t, x + w_t) = 0 \quad \text{for all } (t,x) \in [0,T] \times \mathbb{R}^d. \tag{4.31}$$

Equation (4.31) still holds if $b \in C^0_{t,x}$ and $T^w b \in C^\gamma_t C^{3/2}_x$, since in this case Φ is locally $C^1_{t,x}$ and the same holds for ψ.

In the general case $T^w b \in C^\gamma_t C^2_x$, reasoning by approximation, if $\psi \in C^\gamma_t C^1_{\mathrm{loc}}$ then the equation is still satisfied in the following generalised sense:

$$\psi_t(x) - x = \int_0^t D_x\psi_s(x) T^w b(\mathrm{d}s, x) \quad \text{for all } (t,x) \in [0,T] \times \mathbb{R}^d, \tag{4.32}$$

where the right-hand side is a Young integral in time, for fixed $x \in \mathbb{R}^d$.

However, the regularity requirement $\psi \in C^\gamma_t C^1_{\mathrm{loc}}$ does not need to hold; in general the only information available is $\psi \in C^\gamma_t C^0_{\mathrm{loc}} \cap C^0_t C^1_{\mathrm{loc}}$. Indeed, by the group property

$$\psi(0,s,\cdot) \circ \psi(s,t,\cdot) = \psi(0,t,\cdot)$$

it holds

$$|\psi_t(x) - \psi_s(x)| = |\psi_s(\psi(s,t,x)) - \psi_s(x)| \leqslant [\![\psi_s]\!]_{\mathrm{Lip}} |\psi(s,t,x) - x| \lesssim |t-s|^\gamma$$

where the estimate is uniform in $x \in \mathbb{R}^d$; establishing $\psi \in C^\gamma_t C^1_{\mathrm{loc}}$ requires an analogue estimate for $D_x\psi$, where

$$D_x\psi_t(x) - D_x\psi_s(x) = D_x\psi_s(\psi(s,t,x)) D_x\psi(s,t,x) - D_x\psi_s(x).$$

It is easy to see from the above expression that if $\psi \in C^0_t C^2_{\mathrm{loc}}$ (which by time reversal is equivalent to $\Phi \in C^0_t C^2_{\mathrm{loc}}$), then it belongs to $C^\gamma_t C^1_{\mathrm{loc}}$ as well. As shown in the next section, this condition is met if $T^w b$ is regular enough.

4.3.2. Higher regularity. Similarly to the standard ODE case, we can show that the flow Φ inherits the spatial regularity of $T^w b$, i.e., to a more regular averaged functional $T^w b$ corresponds a more regular flow of solutions.

Theorem 4.33. *Let $n \in \mathbb{N}$, $n \geqslant 1$, $\gamma > 1/2$, and assume that one of the following conditions holds*:

- $T^w b \in C^\gamma_t C^{n+1}_x$; *or*
- $T^w b \in C^\gamma_t C^{n+1/2}_x$ *and* $b \in L^\infty_{t,x}$.

Then the YDE associated to θ admits a locally C^n_x-regular flow $(t,x) \mapsto \Phi(t,x)$.

Proof. As before, we give a detailed proof in the case $T^w b \in C^\gamma_t C^{n+1}_x$, and in the end highlight the main differences in the other case. The idea of the proof, similar to that of Theorem 4.30, is to reason by approximation and establish first that, for $b^\varepsilon = \rho^\varepsilon * b$, it holds $\Phi^\varepsilon \in C^\gamma_t C^{n+1}_x$ with an estimate which is uniform in $\varepsilon > 0$; then the conclusion follows from taking the limit as $\varepsilon \to 0$. In order to get uniform estimates, we will show that for any $k \leqslant n$, $D^k_x \Phi^\varepsilon$ satisfies a variational type equation in which the leading term is a linear Young integral. We split the proof in several steps.

Step 1: *k-th order variation equation.* We start by assuming $b \in L^q_t C^\infty_x$ in addition to the assumptions, so that by standard ODE theory the associated flow has C^∞ spatial regularity. We now adopt the following convention: the symbol $\circ$ denotes a suitably chosen matrix product, which can change from line to line. We claim that, for any $1 \leqslant k \leqslant n$, $D_x^k \Phi$ satisfies the variational-type equation

$$D_x^k \Phi(t,x) = \int_0^t D_x^k \Phi(s,x) \circ D_x T^w b(\mathrm{d}r, \Phi(r,x)) + F_k(\{D_x^i \Phi(\cdot,x)\}_{i \leqslant k-1}), \tag{4.33}$$

where the first integral makes sense in the Young sense and the F_k are "polynomial" functions of the form

$$F_k(\{D_x^i \Phi(\cdot,x)\}) = \sum_{\alpha=1}^{k} \sum_{\beta} a_\beta \int_0^t \bigotimes_{i=1}^{k-1} (D^i \Phi(s,x))^{\otimes \beta_i} \circ D^\alpha T^w b(\mathrm{d}r, \Phi(r,x)),$$

where the internal sum is taken over all possible $\beta = (\beta_1, \dots, \beta_{k-1})$ with $\beta_i \in \{1, \dots, k\}$ such that $\sum_i i\beta_i = k$ and a_β are suitable coefficients of combinatorial nature. Observe that, in terms of the variable $D_x^k \Phi(\cdot, x)$, equation (4.33) is a linear YDE of the form $y_t = \int_0^t A_{\mathrm{d}s} y_s + h_t$, as the second term does not have any dependency on $D_x^k \Phi$.

The proof is by induction on k. The case $k = 1$ is immediate. For $k = 2$, differentiating both terms in the variational equation associated to the drift $(t,x) \mapsto b(t, x + w_t)$,

$$D_x \Phi(t,x) = I + \int_0^t D_x \Phi(s,x) \circ D_x b(s, \Phi(s,x) + w_s)\, \mathrm{d}s,$$

we obtain

$$\begin{aligned} &D_x^2 \Phi(t,x) \\ &= \int_0^t D_x^2 \Phi(s,x) \circ D_x b(s, \Phi(s,x) + w_s)\, \mathrm{d}s + \int_0^t D_x \Phi(s,x)^{\otimes 2} \circ D_x^2 b(s, \Phi(s,x) + w_s)\, \mathrm{d}s \\ &= \int_0^t D_x^2 \Phi(s,x) \circ D_x T^w b(\mathrm{d}r, \Phi(r,x)) + \int_0^t D_x \Phi(s,x)^{\otimes 2} \circ D_x^2 T^w b(\mathrm{d}r, \Phi(r,x)), \end{aligned}$$

which is exactly of the form (4.33). Now assume that the statement is true for k, then differentiating (4.33) on both sides we obtain

$$\begin{aligned} D_x^{k+1} \Phi(t,x) &= \int_0^t D_x^{k+1} \Phi(s,x) \circ D_x T^w b(\mathrm{d}r, \Phi(r,x)) \\ &\quad + \int_0^t (D_x^k \Phi(s,x) \otimes D_x \Phi(s,x)) \circ D_x^2 T^w b(\mathrm{d}r, \Phi(r,x)) + \tilde{F}_{k+1}(\{D^i \Phi(\cdot,x)\}_{i \leqslant k}) \end{aligned}$$

where $\tilde{F}_{k+1}(\{D^i \Phi(\cdot,x)\}_{i \leqslant k+1}) = D_x F_k(\{D^i \Phi(\cdot,x)\}_{i \leqslant k})$ and it is easy to check that it is still of "polynomial type".

Step 2: *Inductive estimate on $D_x^k \Phi$.* Fix $R > 0$; we claim that, for any $2 \leqslant k \leqslant n$, there exists a constant $C_k < \infty$ (which depends on R), which is independent of $\|b\|_{L_t^q C_x^\infty}$, such that

$$\sup_{i \leqslant k} \sup_{x \in B_R} \|D^i \Phi(\cdot,x)\|_{C_t^\gamma} < \infty.$$

Again, the proof is inductive. It mainly relies on the fact that $D^i \Phi$ solves the linear YDE (4.33) in combination with the a priori bounds given by Proposition A.2.

We start by proving the claim in the case $k = 2$. In this case we already know by Theorems 4.4 and 4.30 that $\sup_{x\in B_R}(\|\Phi(\cdot,x)\| + \|D\Phi(\cdot,x)\|_{C_t^\gamma}) \leqslant C$; moreover, by properties of the Young integral we have

$$\Big\|\int_0^\cdot D_x\Phi(s,x)^{\otimes 2}\circ D_x^2 T^w b(\mathrm{d}r,\Phi(r,x))\Big\|_{C^\gamma} \lesssim \|D_x\Phi(\cdot,x)\|_{C^\gamma}^2 \Big\|\int_0^\cdot D_x^2 T^w b(\mathrm{d}r,\Phi(r,x))\Big\|_{C^\gamma}$$
$$\lesssim \|D_x\Phi(\cdot,x)\|_{C^\gamma}^2\|\Phi(\cdot,x)\|_{C^\gamma}\|D_x^2T^w b\|_{C_t^\gamma C_x^1} \lesssim \|T^w b\|_{C_t^\gamma C_x^3}$$

as well as the bound $\|\int_0^\cdot D_x T^w b(\mathrm{d}r,\Phi(r,x))\|_{C^\gamma} \lesssim \|T^w b\|_{C_t^\gamma C_x^2}\|\Phi(\cdot,x)\|_{C^\gamma}$. Applying again Proposition A.3 yields the conclusion in this case.

Assume now that the claim holds for k, then by the inductive hypothesis all the term appearing in the sum defining F_{k+1} can be estimated by

$$\Big\|\int_0^\cdot \bigotimes_{i,\beta_i}(D_x^i\Phi(s,x))^{\otimes\beta_i}\circ D_x^\alpha T^w b(\mathrm{d}r,\Phi(r,x))\Big\|_{C^\gamma}$$
$$\lesssim \prod_{i,\beta_i} C_k^{\beta_i}\Big\|\int_0^\cdot D_x^\alpha T^w b(\mathrm{d}r,\Phi(r,x))\Big\|_{C^\gamma} \lesssim \prod_{i,\beta_i} C_k^{\beta_i} C_k^\nu\|D_x^\alpha T^w b\|_{C_t^\gamma C_x^\nu} \lesssim \|T^w b\|_{C_t^\gamma C_x^{\nu+n}}$$

which together with the estimate for $\|\int_0^\cdot DT^w b(\mathrm{d}r,\Phi(r,x))\|_{C^\gamma}$ and the application of Proposition A.2 yields a new constant C_{k+1}.

Step 3: *Approximation procedure.* Let $b^\varepsilon = \rho^\varepsilon * b$ denote by Φ and Φ^ε the flows associated to b and b^ε respectively. Then $\|T^w b^\varepsilon\|_{C_t^\gamma C_x^{n+1}} \leqslant \|T^w b\|_{C_t^\gamma C_x^{n+1}}$ for all $\varepsilon > 0$ and so, by the previous step, we deduce that for any $R > 0$ there exists a suitable constant C such that

$$\sup_{\varepsilon>0}\|\Phi^\varepsilon\|_{L^\infty W^{n,\infty}(B_R)} \leqslant C.$$

But $\Phi^\varepsilon \to \Phi$ uniformly in $[0,T]\times B_R$, which together with the weak-$*$ compactness of balls in $W^{n,\infty}(B_R)$ implies that $\Phi \in L_t^\infty W^{n,\infty}(B_R)$. A slightly more refined argument, analogue to the one from Theorem 4.30, allows to show that, for any fixed $x\in\mathbb{R}^d$, $D^n\Phi^\varepsilon(\cdot,x)$ must converge as $\varepsilon\to 0$ to the unique solution of the variational-type equation (4.33) associated to Φ; with this information at hand it is then possible to show that the limit varies continuously in x and must coincide with $D^n\Phi(\cdot,x)$, thus showing that Φ is not only in $W^{n,\infty}_{\mathrm{loc}}$ but also C^n. We omit the details to avoid unnecessary repetitions.

Step 4: *The case $b\in L^\infty_{t,x}$ with $T^w b \in C_t^\gamma C_x^{n+1/2}$.* In this case Step 1 and Step 3 are identical to the ones above, the only change is in the estimates from Step 2, as we can use the information $\|\Phi(\cdot,x)\|_{\mathrm{Lip}_t} < \infty$ uniformly in $x\in B_R$ to require less regularity for $T^w b$. For instance in the case $k=2$ we have the estimates

$$\Big\|\int_0^\cdot D_x\Phi(s,x)^{\otimes 2}\circ D_x^2 T^w b(\mathrm{d}r,\Phi(r,x))\Big\|_{C^\gamma} \lesssim \|D_x\Phi(\cdot,x)\|_{C^\gamma}^2 \Big\|\int_0^\cdot D_x^2 T^w b(\mathrm{d}r,\Phi(r,x))\Big\|_{C^\gamma}$$
$$\lesssim \|D_x\Phi(\cdot,x)\|_{C^\gamma}^2\|\Phi(\cdot,x)\|_{\mathrm{Lip}}\|D_x^2T^w b\|_{C_t^\gamma C_x^{1/2}} \lesssim \|T^w b\|_{C_t^\gamma C_x^{5/2}}$$

and $\|\int_0^\cdot D_x T^w b(\mathrm{d}r,\Phi(r,x))\|_{C^\gamma} \lesssim \|T^w b\|_{C_t^\gamma C_x^{3/2}}\|\Phi(\cdot,x)\|_{\mathrm{Lip}}$. The general inductive step is similar. ∎

5. Application to transport type PDEs

The aim of this section is to apply the theory of Section 4 in order to solve perturbed first order linear PDEs of the form

$$\partial_t u + b \cdot \nabla u + cu + \dot{w} \cdot \nabla u = 0, \tag{5.1}$$

where $\dot{w}$ denotes the time derivative of w; at this stage, the equation is only formal. However, if we assumed everything smooth, then applying the change of variables $\tilde{u}(t,x) = u(t, x + w_t)$ (similarly for $\tilde{b}$, $\tilde{c}$), (5.1) would be equivalent to

$$\partial_t \tilde{u} + \tilde{b} \cdot \nabla \tilde{u} + \tilde{c}\tilde{u} = 0. \tag{5.2}$$

Equation (5.2) is now meaningful in the classical sense if for instance $\tilde{b}, \tilde{c} \in C^0_{t,x}$, which is equivalent to $b, c \in C^0_{t,x}$; it also makes sense in the weak sense under suitable integrability assumptions on b, c. Moreover, the transformation that defines $\tilde{u}$ in function of u is well defined whenever w is a continuous path.

Based on the above reasoning, we will adopt the convention that u is a solution to (5.1) if and only if $\tilde{u}$ defined as above is a solution to (5.2), and we will study systematically the latter equation. Let us mention that in the case w is a rough path, it is possible to give meaning to (5.1), and the passage from (5.1) to (5.2) can be rigorously justified, see [11].

Although the above discussion holds for general c, we will focus only on two cases of interest, given by transport and continuity equations, namely for $c = 0$ and $c = \operatorname{div} b$ (respectively, $\tilde{c} = 0$ and $\tilde{c} = \operatorname{div} \tilde{b}$).

In Section 4 the proofs were almost identical for $b \in C^0_{t,x}$ with $T^w b \in C^\gamma_t C^{3/2}_x$ and $T^w b \in C^\gamma_t C^2_x$; here instead the difference becomes relevant and the first case is much easier to treat compared to the latter. To our surprise, even if the existence of a Lipschitz flow for the associated ODE is already known, the case $T^w b \in C^\gamma_t C^2_x$ requires the application of refined tools like commutators and the sewing lemma. For this reason, we split the results in two subsections, with the proofs becoming gradually more complex, so that the difficulties arising in the second case become apparent.

5.1. The case of continuous bounded b

Let us mention that in this case the transport equation was treated with similar techniques in [11], while the continuity equation was studied in [41], Chapter 9. More recently ([2]), the transport equation was investigated in the case $b \in L^\infty_{t,x}$ with different techniques.

We start by considering the case $c \equiv 0$. Recall that $\tilde{b}(t,x) = b(t, x + w_t)$ and that in this case the YDE associated to θ corresponds to the ODE associated to $\tilde{b}$, for which existence of a locally $C^1_{t,x}$ flow Φ is known. Let us also recall the notation from Section 4.3, namely $\Phi_t(x) = \Phi(0,t,x)$, $\psi(s,t,\cdot) = \Phi(s,t,\cdot)^{-1}$ and $\psi_t = \Phi_t^{-1}$. With a slight abuse, from now on we will denote $\tilde{u}$ with u instead.

Proposition 5.1. *Let $b \in C^0_{t,x}$ be such that $T^w b \in C^\gamma_t C^{3/2}_x$. Then for any $u_0 \in C^1_x$ there exists a unique solution of*

$$\partial_t u + \tilde{b} \cdot \nabla u = 0 \tag{5.3}$$

with initial condition u_0, which is given by $u_t(x) = u_0(\psi_t(x))$.

Proof. Recall that, by Remark 4.32, $\psi \in C^1_{t,x}$ solves equation

$$\partial_t \psi(t,x) + D_x\psi(t,x)\,\tilde b(t,x) = 0.$$

Therefore $u(t,x) := u_0(\psi(t,x)) \in C^1_{t,x}$ and satisfies

$$\partial_t u(t,x) + \nabla u(t,x)\cdot \tilde b(t,x) = \nabla u_0(\psi(t,x))\cdot[\partial_t\psi(t,x) + D_x\psi(t,x)\tilde b(t,x)] = 0,$$

which shows that it is a solution of (5.3).

Conversely, let u be a solution and for a given $x \in \mathbb{R}^d$ define $z_t = u(t,\Phi_t(x))$. $\dot\Phi_t(x) = \tilde b(t,\Phi_t(x))$, therefore z solves

$$\dot z_t = \partial_t u(t,\Phi_t(x)) + \nabla u(t,\Phi_t(x))\cdot \tilde b(t,\Phi_t(x)) = 0$$

which implies that $u(t,\Phi_t(x)) = u_0(x)$ for all x and thus $u(t,x) = u_0(\psi(t,x))$. ∎

We now turn to the case $c = \operatorname{div} b$, i.e., the continuity equation. Since in general $\operatorname{div} b$ is only defined as a distribution, it makes sense to interpret the equation in a weak sense.

We adopt the following notation: $\mathcal{M}_x(\mathbb{R}^d) = \mathcal{M}_x$ denotes the Banach space of all finite signed Radon measures on $\mathbb{R}^d$, endowed with the total variation norm. We say that $v \in L^\infty_t \mathcal{M}_x$ is weakly continuous if the map $t \mapsto v_t$ is continuous $\mathcal{M}_x$ endowed with the weak-$*$ topology, equivalently if for any $\varphi \in C^\infty_c(\mathbb{R}^d)$, the map $t \mapsto \langle v_t, \varphi\rangle$ is continuous.

Definition 5.2. Let $\tilde b \in C^0_{t,x}$ and $v \in L^\infty_t \mathcal{M}_x$. We say that v is a weak solution of the continuity equation

$$\partial_t v + \nabla\cdot(\tilde b v) = 0 \tag{5.4}$$

if v is weakly continuous and for any $\varphi \in C^\infty_c([0,T]\times\mathbb{R}^d)$ it holds

$$\langle v_t,\varphi_t\rangle - \langle v_0,\varphi_0\rangle = \int_0^t \langle v_s, \partial_t\varphi_s + \tilde b_s\cdot\nabla\varphi_s\rangle\,\mathrm{d}s. \tag{5.5}$$

Proposition 5.3. *Let $b \in C^0_{t,x}$ be such that $T^w b \in C^\gamma_t C^{3/2}_x$. Then for any $v_0 \in \mathcal{M}_x(\mathbb{R}^d)$ there exists a unique weak solution v of (5.4) with initial data v_0, which is given by*

$$v_t(\mathrm{d}x) = \exp\Big(-\int_0^t \operatorname{div} T^w b(\mathrm{d}s,\psi(s,t,x))\Big)v_0(\mathrm{d}x), \tag{5.6}$$

or equivalently, $v_t(\mathrm{d}x)$ is defined by duality as

$$\int_{\mathbb{R}^d}\varphi(x)\,v_t(\mathrm{d}x) = \int_{\mathbb{R}^d}\varphi(\Phi_t(x))\,v_0(\mathrm{d}x), \quad \forall\varphi\in C^\infty_c. \tag{5.7}$$

Remark 5.4. Whenever $\operatorname{div} b \in C^0_{t,x}$, equation (5.6) corresponds to the classical formulation

$$v_t(\mathrm{d}x) = \exp\Big(-\int_0^t \operatorname{div}\tilde b(s,\psi(s,t,x))\,\mathrm{d}s\Big)v_0(\mathrm{d}x).$$

Under the above assumptions, equation (5.6) is still meaningful as a nonlinear Young integral, since $\operatorname{div} T^w b = T^w \operatorname{div} b \in C^\gamma_t C^{1/2}_x$ and the map $s \mapsto \psi(s,t,x)$ is Lipschitz. However, we will only use formula (5.7) in the proof, as it is more practical for explicit computations.

Proof. Let v be defined by (5.7). Then for any $\varphi \in C_c^\infty([0,T]\times\mathbb{R}^d)$ we have

$$\begin{aligned}
\langle v_t,\varphi_t\rangle - \langle v_0,\varphi_0\rangle &= \int_{\mathbb{R}^d} [\varphi_t(\Phi_t(x)) - \varphi_0(x)] v_0(\mathrm{d}x)\\
&= \int_0^t \int_{\mathbb{R}^d} \frac{\mathrm{d}}{\mathrm{d}s}(\varphi_s(\Phi_s(x)))v_0(\,\mathrm{d}x)\,\mathrm{d}s\\
&= \int_0^t \int_{\mathbb{R}^d} [\partial_t\varphi_s(\Phi_s(x)) + \nabla\varphi_s(\Phi_s(x))\cdot b_s(\Phi(s,x))]v_0(\,\mathrm{d}x)\,\mathrm{d}s\\
&= \int_0^t \int_{\mathbb{R}^d} [\partial_t\varphi_s(x) + \nabla\varphi_s(x)\cdot b_s(x)]v_s(\,\mathrm{d}x)\,\mathrm{d}s,
\end{aligned}$$

which shows that v is a weak solution of (5.4).

Since equation (5.5) is linear, it is enough to establish uniqueness in the case $v_0\equiv 0$. Let v be a given weak solution. Then, by standard density arguments, (5.5) extends to all $\varphi\in C_c^1([0,T]\times\mathbb{R}^d)$; take $\varphi_t(x) = u(\psi_t(x))$ with $u\in C_c^\infty(\mathbb{R}^d)$, so that $\varphi\in C_c^1([0,T]\times\mathbb{R}^d)$ and it solves $\partial_t\varphi + \nabla\varphi\cdot b = 0$. Then we obtain

$$\int u(\psi_t(x))v_t(\,\mathrm{d}x) = \langle v_t,\varphi_t\rangle = \langle v_0,\varphi_0\rangle = 0 \quad \forall u\in C_c^\infty.$$

By usual density arguments, the relation then extends to all continuous bounded u; for fixed t, taking $u(x) = \tilde u(\Phi_t(x))$, we deduce that $\langle \tilde u, v_t\rangle = 0$ for all $\tilde u\in C_b^0$, which implies $v_t\equiv 0$ for all t. ∎

5.2. The case of distributional b

We now pass to the case $T^w b\in C_t^\gamma C_x^2$, without assuming any regularity on the distribution b. To the best of our knowledge, this case has never been considered in literature so far; although perturbed linear PDEs have been previously treated in [11, 43], it is always assumed therein at least $b\in L^\infty_{t,x}$ (which can be treated analogously to Section 5.1). However, our approach in the "Young regime", namely for time regularity $\gamma > 1/2$, is undoubtedly similar (and even simpler) to that in the "rough regime" $\gamma\in(1/3,1/2]$ treated in [5]. The use of a commutator lemma also reflects the work [20] and Chapter 9 from [41]. Abstract transport equations in Hölder media have been treated also in [33]; however the results there are, in our opinion, not completely clear, see Remark 5.10 below.

Definition 5.5. Let $T^w b\in C_t^\gamma C_x^1$. We say that $u\in C_t^\gamma C^0_{\mathrm{loc}}$ is a solution of the Young transport equation

$$u(\mathrm{d}t,x) + \nabla u(t,x)\cdot T^w b(\mathrm{d}t,x) = 0 \tag{5.8}$$

if for all $\varphi\in C_c^\infty(\mathbb{R}^d)$ and all $t\in[0,T]$, the following Young integral equation holds:

$$\langle u_t,\varphi\rangle = \langle u_0,\varphi\rangle + \int_0^t \langle u_s, \operatorname{div}\left(T^w_{\mathrm{d}t} b\,\varphi\right)\rangle. \tag{5.9}$$

Remark 5.6. The integral appearing in (5.9) is meaningful as a Young integral, since by assumptions the map $t\mapsto \operatorname{div}(T^w b(t,\cdot)\varphi)$ belongs to $C_t^\gamma C_c^0$ while $t\mapsto u_t\in C_t^\gamma C^0_{\mathrm{loc}}$. An equivalent more practical formulation of (5.9) is the following one: for any $\varphi \in C_c^\infty(\mathbb{R}^d)$, we have the estimate

$$\big|\langle u_{s,t}, \varphi\rangle - \langle u_s, \operatorname{div}(T^w_{s,t} b\varphi)\rangle\big| \lesssim_K |t-s|^{2\gamma}\, \|\varphi\|_{C^1}\, [\![u]\!]_{C^\gamma C^0_K}\, [\![T^w b]\!]_{C^\gamma C^1} \tag{5.10}$$

which is uniform over $(s,t) \in \Delta_T$ but depends on $K = \operatorname{supp}\varphi$; choosing $\varphi = \rho^\varepsilon(x-\cdot)$ with $x \in B_R$ and ρ^ε standard mollifier, since $T^w_{s,t} b \cdot \nabla u_s$ is a well defined distribution, we obtain

$$\sup_{x\in B_R} |\rho^\varepsilon * u_{s,t} - \rho^\varepsilon * (T^w_{s,t} b\cdot \nabla u_s)| \lesssim_{\varepsilon,R} |t-s|^{2\gamma}. \tag{5.11}$$

If in addition $u \in C^0_t C^1_{\mathrm{loc}}$, we can integrate by parts in (5.10) back to obtain

$$\big|\langle u_{s,t}, \varphi\rangle + \langle T^w_{s,t} b\cdot \nabla u_s,\, \varphi\rangle\big| \lesssim_K |t-s|^{2\gamma}\, \|\varphi\|_{C^1}\, [\![u]\!]_{C^\gamma C^0_K}\, [\![T^w b]\!]_{C^\gamma C^1};$$

if $u \in C^\gamma_t C^1_{\mathrm{loc}}$. This necessarily implies the pointwise identity

$$u(t,x) = u_0(x) + \int_0^t T^w b(\mathrm{d}s, x)\cdot \nabla u(s,x) \quad \text{for all } (t,x) \in [0,T]\times \mathbb{R}^d, \tag{5.12}$$

which is meaningful since $T^w b(\cdot,x), \nabla u(\cdot,x) \in C^\gamma_t$. It is therefore clear that for regular b, any classical solution of (5.3) is also a solution in the sense of Definition 5.5.

We start by showing that our candidate solution satisfies Definition 5.5.

Lemma 5.7. *Let $u_0 \in C^1_x$ and define $u(t,x) = u_0(\psi_t(x))$. Then $u \in C^\gamma_t C^0_x \cap C^0_t C^1_{\mathrm{loc}}$ and it is a solution of the Young transport equation (5.8).*

Proof. The regularity of $u \in C^\gamma_t C^0_x \cap C^0_t C^1_{\mathrm{loc}}$ follows from Remark 4.32, since ψ satisfies

$$\sup_x |\psi_t(x) - \psi_s(x)| \lesssim |t-s|^\gamma, \quad \sup_{t,x} |D_x \psi(t,x)| < \infty,$$

combined with the regularity of u_0. Recall that by (4.30), for any $s<t$ it holds

$$\langle u_{s,t}, \varphi\rangle = \int_{\mathbb{R}^d} (u_0(\psi_t(x)) - u_0(\psi_s(x)))\varphi(x)\,\mathrm{d}x = \int_{\mathbb{R}^d} u_0(x) F(s,t,x)\,\mathrm{d}x,$$

where

$$\begin{aligned} F(s,t,x) = \varphi(\Phi_t(x)) \exp\Big(\int_0^t \operatorname{div} T^w b(r, \Phi_r(x))\Big)& \\ - \varphi(\Phi_s(x)) \exp\Big(\int_0^s \operatorname{div} T^w b(r, \Phi_r(x))\Big)&. \end{aligned}$$

By Young's chain rule, we have the estimates

$$\begin{aligned} |\varphi(\Phi_t(x)) - \varphi(\Phi_s(x)) - \nabla\varphi(\Phi_s(x))\cdot T^w_{s,t} b(\Phi_s(x))| &\lesssim |t-s|^{2\alpha}, \\ \Big|\exp\Big(\int_s^t \operatorname{div} T^w b(r,\Phi_r(x))\Big) - 1 - \operatorname{div} T^w_{s,t} b(\Phi_s(x))\Big| &\lesssim |t-s|^{2\alpha}, \\ \Big|\exp\Big(\int_0^s \operatorname{div} T^w b(r,\Phi_r(x))\Big)\Big| &\lesssim 1, \end{aligned}$$

which can all be taken uniform over x belonging to a compact set K. Combining them we deduce that

$$\begin{aligned}
&F(s,t,x)\\
&\quad\approx [\nabla\varphi(\Phi_s(x))\cdot T^w_{s,t}b(\Phi_s(x)) + \operatorname{div} T^w_{s,t}b(\Phi_s(x))]\exp\Big(\int_0^s \operatorname{div} T^w b(r,\Phi_r(x))\Big)\\
&\quad= \operatorname{div}(T^w_{s,t}b\varphi)(\Phi_s(x))\exp\Big(\int_0^s \operatorname{div} T^w b(r,\Phi_r(x))\Big)
\end{aligned}$$

in the sense of the equality holding up to a term of order $|t-s|^{2\alpha}$. Therefore

$$\begin{aligned}
\langle u_{s,t},\varphi\rangle &\approx \int_{\mathbb{R}^d} u_0(x)\operatorname{div}(T^w_{s,t}b\varphi)(\Phi_s(x))\exp\Big(\int_0^s \operatorname{div} T^w b(r,\Phi_r(x))\Big)\,\mathrm{d}x\\
&= \int_{\mathbb{R}^d} u_0(\psi_s(x))\operatorname{div}(T^w_{s,t}b\varphi)(x)\,\mathrm{d}x = \langle u_s,\operatorname{div}(T^w_{s,t}b\varphi)\rangle,
\end{aligned}$$

which implies the conclusion. ■

Remark 5.8. If $T^w b\in C^\gamma_t C^3_x$, then by Theorem 4.33 and Remark 4.32 it holds $\psi\in C^0_t C^2_x\cap C^\gamma_t C^1_x$. It is then possible to check with standard calculations that for $u_0\in C^2_x$, the solution u constructed as above belongs to $C^0_t C^2_x\cap C^\gamma_t C^1_x$ as well. In this case, by Remark 5.6, u also satisfies the stronger pointwise identity (5.12).

By the method of characteristics, we are able to obtain the following preliminary uniqueness result. It is however of limited applicability, see Remark 5.10 below.

Lemma 5.9. *Let $T^w b\in C^\gamma_t C^2_x$, $u\in C^\gamma_t C^2_{\mathrm{loc}}$ be a solution of (5.8). Then $u(t,x)=u_0(\psi_t(x))$.*

Proof. It is enough to show that the function $f_t := u(t,\Phi_t(x))$ is constant; in particular, it suffices to prove that $|f_{s,t}|\lesssim |t-s|^{2\gamma}$ since $\gamma>1/2$. By the regularity assumption on u, it satisfies (5.12) and therefore

$$|u_{s,t}(x)+\nabla u_s(x)\cdot T^w_{s,t}b(x)| \lesssim_R |t-s|^{2\gamma},\quad \forall y\in B_R.$$

Choosing appropriately R we have

$$\begin{aligned}
f_{s,t} &= u_{s,t}(\Phi_t(x)) + u_s(\Phi_t(x)) - u_s(\Phi_s(x))\\
&= u_{s,t}(\Phi_s(x)) + \nabla u_s(\Phi_s(x))\cdot\Phi_{s,t}(x) + O(|t-s|^{2\gamma})\\
&= -\nabla u_s(\Phi_s(x))\cdot T^w_{s,t}b(y) + \nabla u_s(\Phi_s(x))\cdot\Phi_{s,t}(x) + O(|t-s|^{2\gamma})\\
&= O(|t-s|^{2\gamma})
\end{aligned}$$

where in the last passage we used the fact that $\Phi_{s,t}(x)=\int_s^t T^w b(\mathrm{d}r,\Phi_r(x))\mathrm{d}r$. ■

Remark 5.10. The hypothesis $u\in C^\gamma_t C^2_x$ is required to justify the passage

$$u_s(\Phi_t(x)) - u_s(\Phi_s(x)) = \nabla u_s(\Phi_s(x))\cdot\Phi_{s,t}(x) + O(|t-s|^{2\gamma}),$$

which is not true in general under the sole assumption $u \in C_t^\gamma C_x^1$. However, for general $T^w b \in C_t^\gamma C_x^2$, we only know that $\psi \in C_t^\alpha C_{\mathrm{loc}}^0 \cap C_t^0 C_{\mathrm{loc}}^1$ and so the solution constructed by $u_t(x) = u_0(\psi_t(x))$ is not a priori in the class $C_t^\gamma C_x^2$. For this reason, Lemma 5.9 is potentially vacuous, as it might only imply the non existence of $C_t^\gamma C_x^2$-solutions, while leaving open the problem of uniqueness in the class where u constructed as in Lemma 5.7 lives. We believe the same issue arises in Theorems 3.6 and 3.7 from [33], which do not settle the problem of uniqueness.

Observe that the above issue is typical of the Young regime and is completely absent in the case $b \in C_{t,x}^0$, where uniqueness follows immediately from standard arguments.

In order to prove uniqueness of solutions to (5.8) in the class $C_t^\gamma C_{\mathrm{loc}}^0 \cap C_t^0 C_{\mathrm{loc}}^1$, we need to use an appropriate commutator lemma, in the style of [20]. The basic idea is as follows: let $\{\rho_\varepsilon\}_{\varepsilon>0}$ be a family of standard mollifiers (assume $\rho_1 = \rho$ to be supported on B_1 for simplicity), denote $u^\varepsilon = \rho_\varepsilon * u$; by equation (5.11) we deduce that for any $R > 0$, adopting the notation $C_R^0 = C_{B_R}^0$, it holds

$$\|u_{s,t}^\varepsilon + T_{s,t}^w b \cdot \nabla u_s^\varepsilon + R^\varepsilon(u_s, T_{s,t}^w b)\|_{C_R^0} \lesssim_{\varepsilon,R} |t-s|^{2\gamma} \quad \text{uniformly in } 0 \leqslant s \leqslant t \leqslant T,$$

where the estimate is uniform in $\varepsilon > 0$ and the commutator R^ε appearing is the bilinear operator

$$R^\varepsilon(h,g) = (g \cdot \nabla h)^\varepsilon - g \cdot \nabla h^\varepsilon = \rho^\varepsilon * (g \cdot \nabla h) - g \cdot \nabla(\rho^\varepsilon * h). \tag{5.13}$$

Now $u^\varepsilon \in C_t^\gamma C_{\mathrm{loc}}^2$ and so we can apply the same idea of the proof of Lemma 5.9, i.e., study the function $f_t^\varepsilon = u_t^\varepsilon(\Phi_t(x))$, which we expect to be quasi constant; in the estimates, terms of the form $R^\varepsilon(u_s, T_{s,t}^w b)(\Phi_s(x))$ will then start to appear, and so we need to control them as $\varepsilon \to 0$. For this reason we need the following lemma.

Lemma 5.11. *The operator $R^\varepsilon\colon C_{\mathrm{loc}}^0 \times C_{\mathrm{loc}}^1 \to C_{\mathrm{loc}}^0$ defined by (5.13) is such that*

(i) *there exists C independent of ε such that $\|R^\varepsilon(h,g)\|_{C_R^0} \leqslant C \|h\|_{C_{R+1}^0} \|g\|_{C_{R+1}^1}$,*

(ii) *for any fixed $h \in C^0$, $g \in C^1$, it holds $R^\varepsilon(h,g) \to 0$ uniformly on compact sets as $\varepsilon \to 0$.*

Similar statements hold for $R^\varepsilon\colon C_{\mathrm{loc}}^1 \times C_{\mathrm{loc}}^2 \to C_{\mathrm{loc}}^1$.

Proof. The proof is analogue to the one of Lemma II.1 from [20]. It holds

$$R^\varepsilon(h,g)(x) = \int_{B_1} h(x-\varepsilon z) \frac{g(x-\varepsilon z) - g(x)}{\varepsilon} \cdot \nabla\rho(z) \mathrm{d}z - (h \operatorname{div} g)^\varepsilon(x).$$

Thus claim (i) follows from $\|(h \operatorname{div} g)^\varepsilon\|_{C_R^0} \leqslant \|h\|_{C_{R+1}^0} \|g\|_{C_{R+1}^1}$ and

$$\Big| \int_{B_1} h(x-\varepsilon z) \frac{g(x-\varepsilon z) - g(x)}{\varepsilon} \cdot \nabla\rho(z) \mathrm{d}z \Big| \leqslant \|h\|_{C_{R+1}^0} \|g\|_{C_{R+1}^1} \|\nabla\rho\|_{L^1},$$

where the estimate is uniform in $x \in B_R$. Now fix $R > 0$; we can assume that h, g and Dg all have modulus of continuity ω on B_{R+1}. By known properties of convolutions, $(h \operatorname{div} g)^\varepsilon \to h \operatorname{div} g$ uniformly on compact sets; moreover, for all $x \in B_R$ it holds

$$\Big| \frac{g(x-\varepsilon z) - g(x)}{\varepsilon} - Dg(x)z \Big| = \Big| \int_0^1 [Dg(x-\varepsilon\theta z) - Dg(x)] z \mathrm{d}\theta \Big| \leqslant \omega(\varepsilon);$$

combined with a similar estimate for $|h(x-\varepsilon z)-h(x)|$, this implies that, uniformly in $x\in B_R$,

$$\int_{B_1} h(x-\varepsilon z)\frac{g(x-\varepsilon z)-g(x)}{\varepsilon}\cdot\nabla\rho(z)\,\mathrm{d}z \to h(x)\int_{B_1}\nabla\rho(z)\cdot Db(x)z\,\mathrm{d}z,$$

where the last term equals $h(x)\operatorname{div} b(x)$, which implies claim (ii). The statements for $R^\varepsilon\colon C^1_{\mathrm{loc}}\times C^2_{\mathrm{loc}}\to C^1_{\mathrm{loc}}$ follow once we observe that $\partial_i R^\varepsilon(h,g)=R^\varepsilon(\partial_i h,g)+R^\varepsilon(h,\partial_i g)$ and we apply the previous results. ∎

We have now all the ingredients to show uniqueness in the class $C^\gamma_t C^0_x\cap C^0_t C^1_x$.

Theorem 5.12. *Let $T^w b\in C^\gamma_t C^2_x$ and $u\in C^\gamma_t C^0_x\cap C^0_t C^1_x$ be a solution of (5.8). Then*

$$u(t,x)=u_0(\psi_t(x))\quad \forall(t,x)\in[0,T]\times\mathbb{R}^d.$$

Proof. As before, it suffices to show that for any $x\in\mathbb{R}^d$, $f_t:=u_t(\Phi_t(x))$ satisfies $|f_{s,t}|\lesssim |t-s|^{2\gamma}$, as it implies that f is constant. Recall that Φ satisfies the estimate $|x-\Phi_t(x)|\lesssim |t|^\gamma$ uniformly in x, therefore we can fix B_R such that $\Phi_t(x)\in B_{2R}$ for all $t\in[0,T]$ and all $x\in B_R$; from now on all the norms appearing will be localised on B_{2R} without writing it explicitly.

Since u is a solution of (5.8), it satisfies (5.11) and therefore u^ε is such that

$$\|u^\varepsilon_{s,t}+\nabla u^\varepsilon_s\cdot T^w_{s,t}b+R^\varepsilon(u_s,T^w_{s,t}b)\|_{C^0}\lesssim |t-s|^{2\gamma}\quad\text{uniformly in } 0\leqslant s\leqslant t\leqslant T.$$

Define $f^\varepsilon_t=u^\varepsilon_t(\Phi_t(x))$; using the above property and going through similar calculations as in the proof of Lemma 5.9, we deduce that

$$|f^\varepsilon_{s,t}-R^\varepsilon(u_s,T^w_{s,t}b)(\Phi_s(x))|\lesssim_\varepsilon |t-s|^{2\gamma}. \tag{5.14}$$

The estimate above a priori depends on ε, as it involves $\|u^\varepsilon\|_{C^\gamma C^2}$, but we are now going to show that under the assumptions on $T^w b$ and u it is actually uniform in $\varepsilon>0$. This is accomplished with the help of the sewing lemma, see Lemma A.4 from Appendix A.1. Define

$$\Gamma^\varepsilon_{s,t}:=R^\varepsilon(u_s,T^w_{s,t}b)(\Phi_s(x)),$$

so that relation (5.14) can be rephrased as $|f^\varepsilon_{s,t}-\Gamma^\varepsilon_{s,t}|\lesssim|t-s|^{2\gamma}$. We can estimate $\|\delta\Gamma^\varepsilon\|_{2\gamma}$ as follows:

$$\begin{aligned}
|\delta\Gamma^\varepsilon_{s,u,t}| &= |\Gamma^\varepsilon_{s,t}-\Gamma^\varepsilon_{s,u}-\Gamma^\varepsilon_{u,t}|\\
&\leqslant |R^\varepsilon(u_s,T^w_{s,t}b)(\Phi_s(x))-R^\varepsilon(u_s,T^w_{u,t}b)(\Phi_u(x))|+|R^\varepsilon(u_{s,u},T^w_{u,t}b)(\Phi_u(x))|\\
&\leqslant \|R^\varepsilon(u_s,T^w_{u,t}b)\|_{C^1}|\Phi_s(x)-\Phi_u(x)|+\|R^\varepsilon(u_{s,u},T^w_{u,t}b)\|_{C^0}\\
&\lesssim |t-s|^{2\gamma}(\|R^\varepsilon\|\|u\|_{C^0C^1}[\![T^w b]\!]_{C^\gamma C^2}[\![\Phi_\cdot(x)]\!]_\gamma+\|R^\varepsilon\|[\![u]\!]_{C^\gamma C^0}[\![T^w b]\!]_{C^\gamma C^1})\\
&\lesssim |t-s|^{2\gamma}
\end{aligned}$$

where we used the fact that $[\![\Phi_\cdot(x)]\!]_\gamma\lesssim 1$ by Theorem 4.28 and the estimate is uniform in ε, since $\|R^\varepsilon\|_{\mathcal{L}^2(C^i\times C^{i+1};C^i)}\leqslant C_1$ for $i=1,2$ by Lemma 5.11. It follows that

$\|\delta\Gamma^{\varepsilon}\|_{2\gamma} \leqslant C_2$ for some constant independent of ε and therefore by Lemma A.4 (specifically estimate (A.7)) there exists C_3 such that

$$|u_t^{\varepsilon}(\Phi_t(x)) - u_s^{\varepsilon}(\Phi_s(x)) - R^{\varepsilon}(u_s, T_{s,t}^{w}b)(\Phi_s(x))| = |f_{s,t}^{\varepsilon} - \Gamma_{s,t}^{\varepsilon}| \leqslant C_3|t-s|^{2\gamma} \tag{5.15}$$

for all $\varepsilon > 0$ and $s < t$. Since $u_t^{\varepsilon}(\Phi_t(x)) \to u_t(\Phi_t(x))$ and by part (ii) of Lemma 5.11, $R^{\varepsilon}(u_s, T_{s,t}^{w}b)(\Phi_s(x)) \to 0$, taking the limit as $\varepsilon \to 0$ in (5.15) we deduce that $|u_t(\Phi_t(x)) - u_s(\Phi_s(x))| \leqslant C_3|t-s|^{2\gamma}$, which gives the conclusion. ∎

We now pass to study weak solutions of the continuity equation associated to $T^w b$. Given a distribution v, we say that $v \in (C_x^1)^*$ if there exists a constant C such that $|\langle v, \varphi\rangle| \leqslant C\|\varphi\|_{C^1}$ for all smooth φ. We denote by $\|v\|_{(C^1)^*}$ the optimal constant C. Note that, when v is a measure, $\|v_{s,t}\|_{(C^1)^*}$ is the 1-Wasserstein distance between v_t and v_s.

Definition 5.13. Let $T^w b \in C_t^{\gamma} C_x^1$ and let $v \in C_t^{\gamma}(C_x^1)^*$. We say that v is a weak solution of the Young continuity equation

$$v(\mathrm{d}t) + \operatorname{div}(v_t T^w b(\mathrm{d}t)) = 0, \tag{5.16}$$

if there exists a constant C such that for all $\varphi \in C^2(\mathbb{R}^d)$ the following holds:

$$|\langle v_{s,t}, \varphi\rangle - \langle v_s, T_{s,t}^{w}b \cdot \nabla\varphi\rangle| \leqslant C\|\varphi\|_{C^2}\,|t-s|^{2\gamma}. \tag{5.17}$$

Remark 5.14. As before, it can be shown that for smooth b, any classical solution of

$$\partial_t v + \operatorname{div}(v\tilde{b}) = 0,$$

is also a solution in the sense of the definition above. Equations (5.16) and (5.17) can be rephrased as v satisfying the functional Young integral equation

$$v_{s,t} = \operatorname{div}\Big(\int_s^t v_r T^w b(\mathrm{d}r)\Big),$$

where the integral inside the divergence is a well defined element of $(C^1)^*$ since the product between C^1 and $(C^1)^*$ is still an element of $(C^1)^*$. Formulation (5.17) is however more useful for our purposes.

Lemma 5.15. *Let $T^w b \in C_t^{\gamma} C_x^2$, $v_0 \in \mathcal{M}_x$ and define $v \in L_t^{\infty}\mathcal{M}_x$ by*

$$\langle v_t, \varphi\rangle = \int_{\mathbb{R}^d} \varphi(\Phi_t(x)) v_0(\mathrm{d}x) \quad \forall \varphi \in C_c^{\infty}.$$

Then v is a weak solution of (5.16) with initial condition v_0.

Proof. Let us first show that v defined as above belongs to $C_t^{\gamma}(C_x^1)^*$. It holds

$$\begin{aligned}
|\langle v_{s,t}, \varphi\rangle| &= \Big|\int_{\mathbb{R}^d} [\varphi(\Phi_t(x)) - \varphi(\Phi_s(x))] v_0(\mathrm{d}x)\Big| \\
&\leqslant \|\varphi\|_{\mathrm{Lip}} \sup_{x\in\mathbb{R}^d} |\Phi_{s,t}(x)|\|v_0\|_{\mathcal{M}} \lesssim |t-s|^{\gamma}\,\|\varphi\|_{\mathrm{Lip}}\,\|v_0\|_{\mathcal{M}},
\end{aligned}$$

where we used estimate (4.3); it follows that $\|v\|_{C_t^\gamma(C_x^1)^*} \lesssim \|v_0\|_{\mathcal{M}}$. We now check that v is a solution in the sense of Definition 5.13. It holds

$$|\varphi(\Phi_t(x)) - \varphi(\Phi_s(x)) - \nabla\varphi(\Phi_s(x)) \cdot \Phi_{s,t}(x)| \lesssim \|\varphi\|_{C_b^2} |\Phi_{s,t}(x)|^2 \lesssim \|\varphi\|_{C_b^2} |t-s|^{2\gamma},$$

where as before we used (4.3) and the estimate is uniform in x; similarly,

$$|\Phi_{s,t}(x) - T_{s,t}^w b(\Phi_s(x))| \lesssim |t-s|^{2\gamma} \|T^w b\|_{C^\gamma C^1} [\![\Phi_\cdot(x)]\!]_{C^\gamma} \lesssim |t-s|^{2\gamma}.$$

Combining the two estimates we obtain

$$|\varphi(\Phi_t(x)) - \varphi(\Phi_s(x)) - \nabla\varphi(\Phi_s(x)) \cdot T_{s,t}^w b(\Phi_s(x))| \lesssim \|\varphi\|_{C_b^2} |t-s|^{2\gamma},$$

which yields

$$\begin{aligned}
&|\langle v_{s,t}, \varphi\rangle - \langle v_s, T_{s,t}^w b \cdot \nabla\varphi\rangle| \\
&\qquad \leqslant \left| \int_{\mathbb{R}^d} [\varphi(\Phi_t(x)) - \varphi(\Phi_s(x)) - \nabla\varphi(\Phi_s(x)) \cdot T_{s,t}^w b(\Phi_s(x))] v_0(\mathrm{d}x) \right| \\
&\qquad \lesssim \|\varphi\|_{C_b^2} \|v_0\|_{\mathcal{M}} |t-s|^{2\gamma},
\end{aligned}$$

and thus the conclusion. ∎

Theorem 5.16. *For any given $v_0 \in \mathcal{M}_x$, there exists a unique weak solution of (5.16) in the class $v \in L_t^\infty M_x \cap C_t^\gamma(C_x^1)^*$, which is given by the one from Lemma 5.15.*

Proof. As before, by linearity it is enough to show that there exists a unique solution for the initial condition $v_0 \equiv 0$. The basic strategy is the usual one: given any $u_0 \in C_c^\infty$, setting $u_t(x) = u_0(\psi_t(x))$, it is enough to show that the function $f_t := \langle v_t, u_t\rangle$ is constant, as it implies

$$\langle v_t, u_t\rangle = \int u_0(\psi_t(x)) v_t(\mathrm{d}x) = 0,$$

and thus, reasoning as in the proof of Proposition 5.3, that $v_t \equiv 0$. Observe that the function u has compact space-time support, so we do not need to introduce localisations here.

Now we reason following the same lines as in Theorem 5.12, namely we spatially mollify u so that now u^ε solves

$$u_{s,t}^\varepsilon(x) + \nabla u_s^\varepsilon(x) \cdot T_{s,t}^w b(x) = R^\varepsilon(u_s, T_{s,t}^w b)(x) + O_\varepsilon(|t-s|^{2\gamma}) \tag{5.18}$$

and all the terms are in C_x^1 due to the mollification. Define $f_t^\varepsilon = \langle v_t, u_t^\varepsilon\rangle$; then

$$f_{s,t}^\varepsilon = \langle v_{s,t}, u_s^\varepsilon\rangle + \langle v_s, u_{s,t}^\varepsilon\rangle + \langle v_{s,t}, u_{s,t}^\varepsilon\rangle.$$

The last term trivially satisfies $|\langle v_{s,t}, u_{s,t}^\varepsilon\rangle| \lesssim_\varepsilon |t-s|^{2\gamma}$. Combining the estimates

$$\begin{aligned}
|\langle v_{s,t}, u_s^\varepsilon\rangle - \langle v_s, T_{s,t}^w b \cdot \nabla u_s^\varepsilon\rangle| &\lesssim \|u_s^\varepsilon\|_{C^2} |t-s|^{2\gamma} \lesssim_\varepsilon |t-s|^{2\gamma}, \\
|\langle v_s, u_{s,t}^\varepsilon\rangle + \langle v_s, \nabla u_s^\varepsilon \cdot T_{s,t}^w b\rangle - \langle v_s, R^\varepsilon(u_s, T_{s,t}^w b)\rangle| &\lesssim_\varepsilon \|v_s\|_{(C^1)^*} |t-s|^{2\gamma},
\end{aligned}$$

which come respectively from v being a solution of (5.17) and (5.18) above, we overall obtain

$$|f_{s,t}^\varepsilon - \langle v_s, R^\varepsilon(u_s, T_{s,t}^w b)\rangle| \lesssim_\varepsilon |t-s|^{2\gamma}.$$

As before, the estimate a priori depends on ε, but we can apply the sewing lemma for the choice $\Gamma_{s,t} = \langle v_s, R^\varepsilon(u_s, T^w_{s,t} b)\rangle$ for which, by analogue computations to those of Theorem 5.12, it holds

$$\|\delta\Gamma\|_{2\gamma} \leqslant \|R^\varepsilon\|(\llbracket v\rrbracket_{C^\gamma(C^1)^*}\|u\|_{C^0C^1}\llbracket T^w b\rrbracket_{C^\gamma C^2} + \|v\|_{L^\infty M}\llbracket u\rrbracket_{C^\gamma C^0}\llbracket T^w b\rrbracket_{C^\gamma C^1}) \lesssim 1$$

uniformly in $\varepsilon > 0$. Therefore there exists a constant C independent of ε such that

$$|\langle v_t, u^\varepsilon_t\rangle - \langle v_s, u^\varepsilon_s\rangle - \langle v_s, R^\varepsilon(u_s, b_{s,t})\rangle| \leqslant C|t-s|^{2\gamma}.$$

By the properties of R^ε, taking $\varepsilon \to 0$ we deduce $|\langle v_t, u_t\rangle - \langle v_s, u_s\rangle| \lesssim |t-s|^{2\gamma}$ which implies the conclusion. ∎

A. Some tools

This appendix collects some technical estimates and some reminders of various standard results, from certain functional spaces to stochastic integration in the Banach setting.

A.1. Some useful lemmas

The following chaining lemma is a slight variation on the one from [12], Lemma 3.1.

Lemma A.1. *Let E be a Banach space and let $X\colon [0,T] \to E$ be a continuous stochastic process such that, for some $\lambda > 0$,*

$$\mathbb{E}\Big[\exp\Big(\lambda \frac{\|X_t - X_s\|_E^2}{|t-s|^{2\alpha}}\Big)\Big] \leqslant C \quad \forall s \neq t \in [0,T]. \tag{A.1}$$

Then $\mathbb{P}$-a.s. $X \in C^\omega E$ for the modulus $\omega(|t-s|) = |t-s|^\alpha\sqrt{-\log|t-s|}$ and there exists $\beta > 0$ such that

$$\mathbb{E}\big[\exp\big(\beta\llbracket X\rrbracket^2_{C^\omega E}\big)\big] < \infty.$$

In particular, if $X_0 \equiv 0$, then for any $\gamma < \alpha$ there exists $\beta > 0$ such that

$$\mathbb{E}\big[\exp\big(\beta\|X\|^2_{C^\gamma E}\big)\big] < \infty.$$

Proof. Without loss of generality we can assume $T = 1$. Also, we will only show that proof in the case $\alpha = 1/2$, the other cases being entirely analogue. Let us define the random variable

$$R(\lambda) = \sum_{n\in\mathbb{N}}\sum_{k=0}^{2^n-1} 2^{-2n}\exp\Big(\mu\frac{\|X_{(k+1)2^{-n}} - X_{k2^{-n}}\|_E^2}{2^{-n}}\Big).$$

Then it follows from the assumption that $\mathbb{E}[R(\lambda)] \leqslant C$. We can then apply Lemma 3.1 from [12] to deduce that there exist deterministic positive constants K, β such that

$$\exp\Big(\beta\frac{\|X_t - X_s\|_E^2}{|t-s|}\Big) \lesssim |t-s|^{-K} R(\lambda) \quad \forall s \neq t,$$

which implies by taking the logarithm and dividing by $-\log|t-s|$ that

$$\exp\Big(\beta\Big(\sup_{s\neq t}\frac{\|X_t-X_s\|_E}{|t-s|\sqrt{-\log|t-s|}}\Big)^2\Big)=\sup_{s\neq t}\exp\Big(\beta\frac{\|X_t-X_s\|_E^2}{|t-s|(-\log|t-s|)}\Big)\lesssim R(\lambda),$$

which yields the conclusion. Alternatively, it follows from the assumption that

$$\mathbb{E}[B]:=\mathbb{E}\Big[\int_{[0,T]^2}\exp\Big(\lambda\frac{\|X_t-X_s\|_E^2}{|t-s|^{2\alpha}}\Big)\,\mathrm{d}t\;\mathrm{d}s\Big]<\infty,$$

which implies that we can apply the Garsia–Rodemich–Rumsey theorem (see [30]) for the choice $\psi(x)=e^{\lambda x^2}$, $p(x)=x^\alpha$, which gives

$$\|X_t-X_s\|_E\lesssim\int_0^{|t-s|}\sqrt{B-\log u}\,u^{\alpha-1}\,\mathrm{d}u\lesssim\big(\sqrt{B}+\sqrt{-\log|t-s|}\big)\,|t-s|^\alpha,$$

and from which we can again deduce that

$$\sup_{s\neq t}\frac{\|X_t-X_s\|_E}{|t-s|\sqrt{-\log|t-s|}}\lesssim 1+\sqrt{B}$$

and the exponential integrability bound. The final claim follows immediately. ■

We also provide here a simple lemma on a priori bounds on solutions to linear Young differential equations, in the style of Section 6.2 from [39].

Proposition A.2. *Let $A\in C_t^\gamma(0,T;\mathcal{L}(\mathbb{R}^d;\mathbb{R}^d))$, $h\in C^\gamma([0,T];\mathbb{R}^d)$ and $\gamma>1/2$. Then there exists a unique solution to the YDE*

$$x_t=x_0+\int_0^t A_{\mathrm{d}s}x_s+h_t \tag{A.2}$$

and there exist suitable positive constants which only depend on γ such that

$$[\![x]\!]_{C^\gamma}\lesssim[\![A]\!]_{C^\gamma}\|x\|_{C^0}+[\![h]\!]_{C^\gamma}, \tag{A.3}$$

$$\|x\|_{C^0}\lesssim e^{C[\![A]\!]_{C^\gamma}^{1/\gamma}T}(|x_0+h_0|+[\![h]\!]_{C^\gamma}). \tag{A.4}$$

Proof. Since $A\in C_t^\gamma C_x^\infty$, uniqueness of solutions is well known (see for instance [39]), so we are only interested in proving the bounds (A.3) and (A.4). Up to renaming x_0, we can assume $h_0=0$; we can also assume up to rescaling everything that $T=1$.

We adopt the following notation: for $\Delta\leqslant 1$, we consider

$$[\![x]\!]_{\gamma,\Delta}:=\sup_{\substack{0\leqslant s<t\leqslant T\\|s-t|\leqslant\Delta}}\frac{|x_{s,t}|}{|t-s|^\gamma}.$$

Let $\Delta>0$ to be chosen later and let $s<t$ be such that $|t-s|\leqslant\Delta$. By (A.2) it holds

$$\begin{aligned}|x_{s,t}|&\leqslant\Big|\int_s^t A_{\mathrm{d}r}x_r\Big|+|h_{s,t}|\leqslant|A_{s,t}x_s|+C|t-s|^{2\gamma}[\![A]\!]_{C^\gamma}[\![x]\!]_{\gamma,\Delta}+|t-s|^\gamma\|h\|_{C^\gamma}\\&\leqslant|t-s|^\gamma\big([\![A]\!]_{C^\gamma}\|x\|_{C^0}+[\![h]\!]_{C^\gamma}\big)+C\Delta^\gamma[\![A]\!]_{C^\gamma}[\![x]\!]_{\gamma,\Delta},\end{aligned}$$

and so, dividing both sides by $|t-s|^\gamma$, taking the supremum over s, t and choosing Δ such that $C\Delta^\gamma [\![A]\!]_{C^\gamma} \leqslant 1/2$ we obtain

$$[\![x]\!]_{\gamma,\Delta} \leqslant 2\big([\![A]\!]_{C^\gamma} \|x\|_{C^0} + [\![h]\!]_{C^\gamma}\big). \tag{A.5}$$

We now distinguish two cases. If $[\![A]\!]_{C^\gamma}$ is such that $(2+C)[\![A]\!]_{C^\gamma} \leqslant 1/2$, then it follows from (A.5) with the choice $\Delta = 1$ and the trivial estimate $\|x\|_{C^0} \leqslant |x_0| + [\![x]\!]_\gamma$ that

$$[\![x]\!]_\gamma \lesssim [\![A]\!]_{C^\gamma}|x_0| + [\![h]\!]_{C^\gamma} \lesssim |x_0| + [\![h]\!]_{C^\gamma},$$

which immediately implies the conclusion. Suppose instead the opposite and choose Δ such that $1/4 \leqslant (2+C)\Delta^\gamma [\![A]\!]_{C^\gamma} \leqslant 1/2$; define $I_n = [(n-1)\Delta, n\Delta]$, $J_n = \sup_{t\in I_n}|x_t|$; then estimates similar to the one done above show that

$$J_{n+1} \leqslant |x_{n\Delta}| + \Delta^\gamma\, [\![x]\!]_{C^\gamma(I_n)} \leqslant |x_{n\Delta}|(1 + 2\Delta^\gamma [\![A]\!]_{C^\gamma}) + 2[\![h]\!]_{C^\gamma} \lesssim J_n + [\![h]\!]_{C^\gamma},$$

which implies recursively that for a suitable constant C it holds $J_n \leqslant C^n(|x_0| + [\![h]\!]_{C^\gamma})$. Since $n \sim \Delta^{-1} \sim [\![A]\!]_{C^\gamma}^{1/\gamma}$ we deduce that

$$\|x\|_{C^0} = \sup_n J_n \lesssim C^{[\![A]\!]_{C^\gamma}^{1/\gamma}}(|x_0| + [\![h]\!]_{C^\gamma}),$$

which gives (A.4); this, combined with $\Delta^{-\gamma} \sim [\![A]\!]_{C^\gamma}$, estimate (A.5) and the basic inequality

$$[\![x]\!]_{C^\gamma} \lesssim \Delta^{-\gamma}\|x\|_{C^0} + [\![x]\!]_{\gamma,\Delta},$$

yields estimate (A.3). ∎

Similarly to the above lemma, we also have the following result.

Lemma A.3. *Let $A \in C^\gamma_t \mathrm{Lip}_x$ be such that $A(t,0) = 0$ for all $t \geqslant 0$, let $h \in C^\gamma_t$, and let x be a solution of the nonlinear YDE*

$$x_t = x_0 + \int_0^t A(\mathrm{d}s, x_s) + h_t.$$

Then there exist suitable positive constants, which only depend on γ, such that

$$\|x\|_{C^\gamma} \lesssim e^{C[\![A]\!]_{C^\gamma}^{1/\gamma} T}(1 + [\![A]\!]_{C^\gamma \mathrm{Lip}})\,(|x_0 + h_0| + [\![h]\!]_{C^\gamma}). \tag{A.6}$$

Proof. Since x is a solution to the nonlinear YDE, for any $s < t$ it holds

$$\begin{aligned}|x_{s,t}| &\leqslant |A_{s,t}(x_s)| + C|t-s|^{2\gamma}\, [\![A]\!]_{C^\gamma \mathrm{Lip}}\, [\![x]\!]_{C^\gamma([s,t])} + |h_{s,t}| \\ &\leqslant |t-s|^\gamma [\![A]\!]_{C^\gamma \mathrm{Lip}}\, |x_s| + C|t-s|^{2\gamma}\, [\![A]\!]_{C^\gamma \mathrm{Lip}}\, [\![x]\!]_{C^\gamma([s,t])} + |t-s|^\gamma\, [\![h]\!]_{C^\gamma},\end{aligned}$$

where in the second line we used the fact that $A(s,0) = 0$ by hypothesis. The rest of the proof from here on is identical to the one of Lemma A.2 and we omit it. The inequality (A.6) is a combination of inequalities (A.3) and (A.4). ∎

We conclude this section by recalling the sewing lemma, which is a fundamental tool in the theory of rough paths. Consider an interval $[0, T]$ and a Banach space E; let Δ_n denote the n-simplex on $[0, T]$, so that $\Delta_n = \{(t_1, \dots, t_n) : 0 \leqslant t_1 \leqslant \cdots \leqslant t_n \leqslant T\}$. Given a map $\Gamma\colon \Delta_2 \to E$, we define $\delta\Gamma\colon \Delta_3 \to E$ by

$$\delta\Gamma_{s,u,t} := \Gamma_{s,t} - \Gamma_{s,u} - \Gamma_{u,t}.$$

We say that $\Gamma \in C_2^{\alpha,\beta}([0, T]; E)$ if $\Gamma_{t,t} = 0$ for all $t \in [0, T]$ and $\|\Gamma\|_{\alpha,\beta} < \infty$, where

$$\|\Gamma\|_\alpha := \sup_{s<t} \frac{\|\Gamma_{s,t}\|_E}{|t-s|^\alpha}, \quad \|\delta\,\Gamma\|_\beta := \sup_{s<u<t} \frac{\|\delta\,\Gamma_{s,u,t}\|_E}{|t-s|^\beta}, \quad \|\Gamma\|_{\alpha,\beta} := \|\Gamma\|_\alpha + \|\delta\,\Gamma\|_\beta\,.$$

Let us remark that for a map $f\colon [0, T] \to E$, we still denote by $f_{s,t}$ the increment $f_t - f_s$.

Lemma A.4 (Sewing lemma). *Let α and β be such that $0 < \alpha \leqslant 1 < \beta$. For any $\Gamma \in C_2^{\alpha,\beta}([0, T]; E)$ there exists a unique map $\mathcal{J}\Gamma \in C^\alpha([0, T]; E)$ such that $(\mathcal{J}\Gamma)_0 = 0$ and*

$$\|(\mathcal{J}\Gamma)_{s,t} - \Gamma_{s,t}\|_E \leqslant C\, \|\delta\Gamma\|_\beta\, |t-s|^\beta, \tag{A.7}$$

where the constant C only depends on β. In particular, the map $\mathcal{J}\colon C_2^{\alpha,\beta} \to C^\alpha$ is linear and bounded and there exists a constant C', which only depends on β and T, such that

$$\|\mathcal{J}\Gamma\|_{C^\alpha} \leqslant C'\|\Gamma\|_{\alpha,\beta}. \tag{A.8}$$

For given Γ, the map $\mathcal{J}\Gamma$ is characterised as the unique limit of Riemann–Stieltjes sums: for any $t > 0$,

$$(\mathcal{J}\Gamma)_t = \lim_{|\Pi|\to 0} \sum_i \Gamma_{t_i,t_{i+1}}.$$

The notation above means that for any sequence of partitions $\Pi_n = \{0 = t_0 < t_1 < \cdots < t_{k_n} = t\}$ with mesh $|\Pi_n| = \sup_{i=1,\dots,k_n} |t_i - t_{i-1}| \to 0$ as $n \to \infty$, it holds

$$(\mathcal{J}\Gamma)_t = \lim_{n\to\infty} \sum_{i=0}^{k_n-1} \Gamma_{t_i,t_{i+1}}.$$

For a proof, see Lemma 4.2 from [28]. Let us point out that estimate (A.7) is extremely useful even in cases even when $\mathcal{J}\Gamma$ is already known, as it asserts that in order to control $\|(\mathcal{J}\Gamma)_{s,t} - \Gamma_{s,t}\|$ it is enough to have an estimate for $\|\delta\Gamma\|_\beta$.

A.2. Function spaces

We recall here the definition and basic properties of the function spaces we consider, which are Bessel potential spaces $L^{s,p}(\mathbb{R}^d)$ and Besov spaces $B^s_{p,q}(\mathbb{R}^d)$. In particular, in view of application to regularity estimates from Section 3.3, we need interpolation estimates and heat kernel estimates for such spaces. Bessel potential spaces are a subclass of Triebel–Lizorkin spaces, which will be also introduced. Most of the material is classical and covered in the monographs [4] and [55].

Definition A.5. For $s \geqslant 0$, we call Bessel potential and we denote it by G^s the linear operator with Fourier symbol given by $(1+|\xi|^2)^{-s/2}$, with the convention $G^0 = I$. For any $p \in (1,\infty)$, G^s is a continuous embedding of L^p into itself and it satisfies the semigroup property $G^t G^s = G^{t+s}$. For $p \in (1,\infty)$ and $s \geqslant 0$, we define the Bessel potential space $L^{s,p}$ as $G^s(L^p)$ (with the convention $L^{0,p} = L^p$), endowed with the norm

$$\|f\|_{L^{s,p}} := \|g\|_{L^p} \quad \text{if } f = G^s g.$$

It follows immediately from the definition and the semigroup property that G^t provides an isomorphism of $L^{s,p}$ and $L^{s+t,p}$ in the sense that $\|G^t f\|_{L^{s+t,p}} = \|f\|_{L^{s,p}}$. This allows also to define $L^{s,p}$ for negative values of s as the set of distributions f such that $G^s f \in L^p$. Whenever $s = m$ integer, the space $L^{s,p}$ coincides with the classical Sobolev space $W^{m,p}$, with equivalent norm. Similarly to Sobolev spaces, Bessel embeddings are available; in particular if $sp > d$, we have the continuous embedding $L^{s,p} \hookrightarrow C^\gamma$ with $\gamma = s - d/p$, whenever γ is not an integer.

Definition A.6. Let $\mathcal{A}$ be the annulus $\bar{B}_{8/3} \setminus B_{3/4}$. A dyadic pair is a couple of functions (χ,φ) such that $\chi \in C_c^\infty(B_{4/3})$, $\varphi \in C_c^\infty(\mathcal{A})$ and such that

$$\chi(\xi) + \sum_{j=0}^{\infty} \varphi(2^{-j}\xi) = 1 \quad \forall \xi \in \mathbb{R}^d$$

as well as

$$|j-j'| \geqslant 2 \;\Rightarrow\; \operatorname{supp}\varphi(2^{-j}\cdot) \cap \operatorname{supp}\varphi(2^{-j'}\cdot) = \emptyset.$$

Given such a dyadic pair, we define the operator Δ_{-1} by $\Delta_{-1} f = \mathcal{F}^{-1}(\chi \mathcal{F} f)$, and similarly Δ_j for $j \geqslant 0$ by $\Delta_j f = F^{-1}(\varphi(2^{-j}\cdot)\mathcal{F} f)$.

Definition A.7. For $s \in \mathbb{R}$, $(p,q) \in [1,\infty]^2$, we define the Besov space $B^s_{p,q}$ as the set of all tempered distributions f such that

$$\|f\|^q_{B^s_{p,q}} := \sum_{j=0}^{\infty} 2^{sjq}\, \|\Delta_j f\|^q_{L^p} < \infty.$$

The spaces $B^s_{2,2}$ coincide with the (fractional) Sobolev spaces H^s, which also coincide with $L^{s,2}$; however, for $p \neq 2$ Bessel and Besov spaces do not coincide. The space $B^s_{\infty,\infty}$ coincides with C^s whenever s is not an integer. Also in the case of Besov spaces, embedding theorems are available; in particular, $B^s_{p,q} \hookrightarrow B^{s-d/p}_{\infty,\infty}$, which coincides with C^γ whenever $\gamma = s - d/p$ is not an integer. Let us also point out that the exponent q is most of the time not particularly relevant, as for any $\tilde{q} < q$ and any $\varepsilon > 0$ we have the embeddings $B^s_{p,\tilde{q}} \hookrightarrow B^s_{p,q} \hookrightarrow B^{s-\varepsilon}_{p,\tilde{q}}$.

Definition A.8. For $s \in \mathbb{R}$, $(p,q) \in [1,\infty]^2$, we define the Triebel–Lizorkin space $F^s_{p,q}$ as the set of all tempered distributions f such that

$$\|f\|_{F^s_{p,q}} := \Big\| \Big(\sum_{j=0}^{\infty} 2^{sjq} |\Delta_j f(\cdot)|^q \Big)^{1/q} \Big\|_{L^p(\mathbb{R}^d)} < \infty.$$

Both definitions of Besov and Triebel–Lizorkin spaces are independent of the dyadic pair (χ, φ), in the sense that different pairs yields the same space of distributions with equivalent norms. Bessel spaces $L^{s,p}$ correspond to $F^s_{p,2}$; the spaces $F^s_{p,q}$ and $B^s_{p,q}$ coincide if and only if $p = q$, in which case $F^s_{p,p} = B^s_{p,p} = W^{s,p}$ are sometimes referred to as fractional Sobolev spaces, see [19]. In the case $p \neq q$, suitable embeddings between $F^s_{p,q}$ and $B^s_{p,q}$ follow immediately from Minkowski's inequality, since their norms can be regarded respectively as $L^p(\mathbb{R}^d, \lambda; \ell^q(\mathbb{N}, \mu))$- and $\ell^q(\mathbb{N}, \mu; L^p(\mathbb{R}^d, \lambda))$-norms, where λ is the Lebesgue measure in $\mathbb{R}^d$ and μ is the counting measure on $\mathbb{N}$. In particular, for $p > q$ it holds $B^s_{p,q} \hookrightarrow F^s_{p,q}$, while for $p < q$ we have the reversed embedding.

We now state a simple interpolation-like inequality for Bessel and Besov spaces. Since we do not have a direct reference for this result, we also provide a quick proof.

Lemma A.9. *Let $s > 0$ and $p \in [2, \infty)$. Then for any $\varepsilon > 0$ there exists a constant c_ε such that*

$$\| f \|_{L^{s,p}} \leqslant c_\varepsilon \| f \|_{L^p}^{1-\theta} \| f \|_{L^{s+\varepsilon,p}}^{\theta}, \quad \text{where } \theta = \frac{s}{s+\varepsilon} \cdot$$

The same statement holds with the $L^{s,p}$ norm replaced by the $B^s_{p,q}$ norm.

Proof. We use here the equivalent norm for $L^{s,p}$ given by $\| \cdot \|_{F^s_{p,2}}$ as defined above. For any $N \in \mathbb{N}$ it holds

$$\begin{aligned}
\| f \|_{F^s_{p,2}} &= \Big\| \Big(\sum_j 2^{2sj} |\Delta_j f| \Big)^{1/2} \Big\|_{L^p} \\
&\leqslant \Big\| \Big(\sum_{j \leqslant N} 2^{2sj} |\Delta_j f| \Big)^{1/2} \Big\|_{L^p} + \Big\| \Big(\sum_{j > N} 2^{2sj} |\Delta_j f| \Big)^{1/2} \Big\|_{L^p} \\
&\leqslant 2^{sN} \Big\| \Big(\sum_{j \leqslant N} |\Delta_j f| \Big)^{1/2} \Big\|_{L^p} + 2^{-\varepsilon N} \Big\| \Big(\sum_{j > N} 2^{2(s+\varepsilon)j} |\Delta_j f| \Big)^{1/2} \Big\|_{L^p} \\
&\leqslant 2^{sN} \| f \|_{L^p} + 2^{-\varepsilon N} \| f \|_{F^{s+\varepsilon}_{p,2}} .
\end{aligned}$$

Choosing N such that $2^{sN} \| f \|_{L^p} \sim 2^{-\varepsilon N} \| f \|_{F^{s+\varepsilon}_{p,2}}$, we obtain the conclusion for $L^{s,p}$.

A similar proof can be carried out for $B^s_{p,q}$; alternatively in this case one can use Hölder's inequality as follows:

$$\begin{aligned}
\| f \|^q_{B^s_{p,q}} &= \sum_j 2^{sqj} \| \Delta_j f \|^q_{L^p} \\
&\leqslant \Big(\sum_j 2^{sqj/\theta} \| \Delta_j f \|^q_{L^p} \Big)^{\theta} \Big(\sum_j \| \Delta_j f \|^q_{L^p} \Big)^{1-\theta} = \| f \|^{\theta q}_{B^{s/\theta}_{p,q}} \| f \|^{(1-\theta)q}_{B^0_{p,q}},
\end{aligned}$$

which gives the conclusion for the choice $\theta = s/(s+\varepsilon)$. ∎

We also need to recall the action of the heat flow P_t on such spaces; with a slight abuse of notation, we will denote by P_t both the convolution operator and the Gaussian density itself.

Lemma A.10. *For any $s \in \mathbb{R}$, $\rho > 0$, $p \in (1, \infty)$ and for any $f \in L^{s,p}$, $t > 0$, it holds*

$$\| P_t f \|_{L^{s+\rho,p}} \lesssim t^{-\rho/2} \, \| f \|_{L^{s,p}}.$$

Similarly, for any $s \in \mathbb{R}$, $\rho > 0$, $p, q \in [1, \infty]$ and for any $f \in B^s_{p,q}$, $t > 0$, it holds

$$\| P_t f \|_{B^{s+\rho}_{p,q}} \lesssim t^{-\rho/2} \, \| f \|_{B^s_{p,q}}.$$

Both statements are classical. The first one follows immediately from the fact that, due to the scaling $P_t = t^{-d/2} P_1(t^{-1/2}\cdot)$, it holds $\| P_t \|_{L^{\rho,1}} = t^{-\rho/2} \| P_1 \|_{L^{\rho,1}}$. See Proposition 5 at page 2414 of [42] for a proof in a more general context of the second statement.

A.3. A primer on stochastic integration in UMD Banach spaces

In this section we recall several results on abstract stochastic integration which are needed in order to complete the proof of Theorem 3.11; we believe they are also of independent interest and therefore provide a general presentation. In view of application to Section 3.3, we only need results for martingale type 2 spaces, which however yield the restriction to work on L^p-based spaces with $p \geqslant 2$; weakening this condition to the case $p = 1$ would highly enhance the results, as discussed in Remark 3.19, which is why we also discuss here UMD Banach spaces. Although with this more general theory we are currently not able to overcome the obstacle, we believe it might be of help for future developments and improvements.

All the material presented here is taken from [56], [57]. Also, we restrict for simplicity to the case W is a real valued Brownian motion (the extension to the vector valued case $W \in \mathbb{R}^d$ being straightforward), but the theory is far more general as it considers the case of H-cylindrical Brownian motion, H being an abstract Hilbert space. This gives rise to γ-Radonifying norms $\gamma(H, E)$; in our simple setting, $H = \mathbb{R}$, for any Banach space E it holds $\| \cdot \|_{\gamma(H,E)} = \| \cdot \|_E$.

Definition A.11. Let $p \in [1, 2]$. A Banach space E has martingale type p if there exists a constant $C \geqslant 0$ such that for all finite E-valued martingale difference sequences $(d_n)_{n=1}^N$ it holds

$$\mathbb{E}\Big[\Big\| \sum_{n=1}^N d_n \Big\|_E^p\Big] \leqslant C^p \sum_{n=1}^N \mathbb{E}[\| d_n \|_E^p].$$

The least admissible constant is denoted by $C_{p,E}$.

Examples of martingale type spaces are the following:

- Every Banach space has martingale type 1.
- Every Hilbert space has martingale type 2.
- A closed subspace of a Banach space of martingale type p has still martingale type p.
- If E has martingale type p and $(S, \mathcal{A}, \mu)$ is a measure space, then $L^r(S; E)$, with $r \in [1, \infty)$, has martingale type $p \wedge r$; in particular Lebesgue spaces $L^p(\mathbb{R}^d)$ have martingale type $p \wedge 2$.

- Let (E_0, E_1) be an interpolation couple such that E_i has martingale type $p_i \in [1,2]$, let $\theta \in (0,1)$ and consider $p \in [1,2]$ such that $1/p = (1-\theta)/p_0 + \theta/p_1$. Then both the complex and real interpolation spaces E_θ and $\tilde{E}_\theta$ have martingale type p.

For the last two examples, see Propositions 7.1.3 and 7.1.4 from [36]. It follows from the previous list of examples that Sobolev spaces $W^{k,p}(\mathbb{R}^d)$ with $p \in [2,\infty)$ have martingale type as they can be identified with closed subspaces of $L^p(\mathbb{R}^d)^{\otimes n}$ for suitable n; Bessel potential spaces $L^{s,p}(\mathbb{R}^d)$ with general s are isomorphic to $L^p(\mathbb{R}^d)$, with isomorphism given by $G^s = (1-\Delta)^{s/2}$, therefore for $p \in [2,\infty)$ they have martingale type 2. In the case of Besov spaces $B^s_{p,q}$ with $p,q \in [2,\infty)$, again it can be shown that they have martingale type 2, either by constructing them as interpolation spaces (see for instance Section 17.3 from [40]) or reasoning as follows: by definition, any $\varphi \in B^s_{p,q}$ can be identified with a sequence $\{\Delta_j \varphi\}_j \subset L^p(\mathbb{R}^d)$ with suitable summability, namely such that it belongs to $\ell^q(\mathbb{N},\mu; L^p(\mathbb{R}^d))$, where $\mu(\{j\}) = 2^{-sqj}$; in the case $p,q \in [2,\infty)$, by the previous examples it has martingale type 2.

Now let W be a real valued $\mathcal{F}_t$-Brownian motion on a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geqslant 0}, \mathbb{P})$, $\{\mathcal{F}_t\}_{t\geqslant 0}$ being a filtration satisfying the usual conditions. For martingale type 2 spaces it is possible to define stochastic integrals analogously to the standard case: for an adapted elementary process $\phi: \mathbb{R}_+ \times \Omega \to E$, namely of the form

$$\phi(t,\omega) = \sum_{i=1}^{n-1} x_i \, \mathbb{1}_{(t_i,t_{i+1}]\times F_i}(t,\omega),$$

where $0 \leqslant t_1 < t_2 < \cdots < t_n$, $x_i \in E$, $F_i \in \mathcal{F}_{t_i}$, we set

$$\int_0^\cdot \phi \,\mathrm{d}W := \sum_{i=1}^{n-1} x_i \, \mathbb{1}_{F_i}(W_{\cdot\wedge t_{i+1}} - W_{\cdot\wedge t_i}).$$

Using the martingale type 2 property it is then possible to show that the L^2 norm of the process defined in this way is controlled by $\|\phi\|_{L^2(\mathbb{R}_+\times\Omega,E)}$, see Theorem 4.6 from [57]. By standard approximation procedures, together with Doob's maximal inequality, the following analogue of standard Itô integration can then be proven.

Theorem A.12. *Let $\phi: \mathbb{R}_+ \times \Omega \to E$ be a progressively measurable process satisfying*

$$\|\phi\|^2_{L^2(\mathbb{R}_+\times\Omega,E)} = \mathbb{E}\Big[\int_0^{+\infty} \|\phi_t\|_E^2 \,\mathrm{d}t\Big] < \infty.$$

Then $\int \phi \,\mathrm{d}W$ is well defined as an E-valued martingale with paths in $C_b(\mathbb{R}_+; E)$ and

$$\mathbb{E}\Big[\sup_{t\geqslant 0}\Big\|\int_0^t \phi_s \,\mathrm{d}W_s\Big\|_E^2\Big] \leqslant 4C_{2,E}^2\, \mathbb{E}\Big[\int_0^{+\infty} \|\phi_t\|_E^2 \,\mathrm{d}t\Big]. \tag{A.9}$$

Let us also remark that it follows immediately from the definition for simple processes and the usual approximation procedure that, for any ϕ as above and any deterministic $\varphi^* \in E^*$, the following identity holds:

$$\Big\langle \varphi^*, \int_0^\cdot \phi_s \,\mathrm{d}W_s\Big\rangle = \int_0^\cdot \langle \varphi^*, \phi_s\rangle \,\mathrm{d}W_s \tag{A.10}$$

where the integral on the right-hand side is a standard real valued stochastic integral.

We are now ready to complete the proof of Theorem 3.11.

Proof of Theorem 3.11. Let us first show the following general fact: given a separable Banach space E and two E-valued random variables X and Y such that for any φ^* in a linearly dense subspace of E^* it holds

$$\langle \varphi^*, X\rangle = \langle \varphi^*, Y\rangle \quad \mathbb{P}\text{-a.s.,} \tag{A.11}$$

then necessarily $X = Y$ $\mathbb{P}$-a.s. Indeed, it follows from the linear density assumption that relation (A.11) holds for any $\varphi^* \in E^*$; by separability of E and the Hahn–Banach theorem, it is possible to find a countable collection $\{\varphi_n^*\}_n \subset E^*$ such that $\|\varphi_n^*\|_{E^*} = 1$ for all n and

$$\|x\| = \sup_n |\langle \varphi_n^*, x\rangle| \quad \forall x \in E.$$

By (A.11) and the fact that the supremum is over a countable set, we can find a set Γ of full probability such that

$$\|X - Y\|_E = \sup_n |\langle \varphi_n^*, X - Y\rangle| = 0 \quad \forall \omega \in \Gamma,$$

which proves the claim. Now let $b \in C_c^\infty([0,T] \times \mathbb{R}^d; \mathbb{R})$, so that it can be identified with an element of $L^2(0,T; H^\alpha(\mathbb{R}^d))$ for any $\alpha > 0$; choose α big enough so that H^α embeds into continuous functions vanishing at infinity. Then thanks to relation (A.10), equation (3.9) can be written as: for a given $x \in \mathbb{R}^d$, $\mathbb{P}$-a.s. it holds

$$\begin{aligned}\langle \delta_x, \int_s^t b\left(r, \cdot + W_r^H\right) \mathrm{d}r\rangle &= \langle \delta_x, \int_s^t P_{\tilde{c}_H |r-s|^{2H}}\, b\left(r, \cdot + W_{s,r}^{2,H}\right) \mathrm{d}r\rangle \\ &\quad + \langle \delta_x, \int_s^t \int_u^t P_{\tilde{c}_H |r-u|^{2H}} \nabla b\left(r, \cdot + W_{u,r}^{2,H}\right) c_H |r-u|^{H-1/2}\, \mathrm{d}r \cdot \mathrm{d}B_u\rangle,\end{aligned}$$

where the first two integrals are interpreted as (random) Bochner integrals while the last one as a stochastic integral in H^α (with the inner integral being a random Bochner integral); integrability and predictability are straightforward due to the regularity of b and the properties of $W_{u,r}^{2,H}$. As the collection $\{\delta_x\}_{x\in\mathbb{R}^d}$ is linearly dense in $H^{-\alpha}$ and H^α is separable, we can apply the general fact above to deduce that, for $s < t$ fixed, the random variables above coincide on a set of full probability, without the need of testing against δ_x. This is exactly formula (3.10). ■

In the setting of martingale type 2 spaces, a one-sided Burkholder's inequality is available; we state it with the optimal asymptotic behaviour of the constants, which is needed in the estimates in Section 3.3. It was first shown by Seidler in [51].

Theorem A.13 (Theorem 4.7 from [57]). *Let E be a martingale type 2 space. Then for any progressively measurable process $\phi\colon \mathbb{R}_+ \times \Omega \to E$ and $p \in (0,\infty)$, there exists a constant $\tilde{C}_{p,E}$ such that*

$$\mathbb{E}\Big[\sup_{t\geqslant 0} \Big\| \int_0^t \phi_s\, \mathrm{d}W_s \Big\|_E^2\Big] \leqslant \tilde{C}_{p,E}^p\, \mathbb{E}\Big[\Big(\int_0^\infty \|\phi_s\|_E^2\, \mathrm{d}s\Big)^{p/2}\Big]. \tag{A.12}$$

In particular, it is possible to choose $\tilde{C}_{p,E}$ such that $\tilde{C}_{p,E} \leqslant C_E \sqrt{p}$ for any $p \geqslant 2$, where C_E is a universal constant that only depends on the space E.

This concludes the exposition of results needed in the proofs of this work. In the rest of this appendix, we present a brief account on stochastic integration in UMD Banach spaces.

Some of the major drawbacks of martingale type 2 spaces are the fact that they do not include L^p spaces with $p < 2$, Burkholder's inequality is in general only one-sided and it is not sharp, which is troublesome in applications to maximal regularity of mild solutions of SPDEs. This motivates the introduction of a larger class of spaces. As before, we only consider the case of a real valued W, but the theory extends to W being a cylindrical H-Brownian motion for an Hilbert space H.

Definition A.14. A Banach space E is called a UMD space (i.e., it has unconditional martingale differences) for some $p \in (1, \infty)$ if there exists a constant $\beta \geqslant 0$ such that for all E-valued L^p-martingale differences $(d_n)_{n\geq 1}$ and signs $(\varepsilon_n)_{n\geq 1}$ one has

$$\mathbb{E}\Big[\Big\| \sum_{n=1}^{d} \varepsilon_n d_n \Big\|_E^p\Big] \leqslant \beta^p\, \mathbb{E}\Big[\Big\| \sum_{n=1}^{N} d_n \Big\|_E^p\Big] \quad \forall N \geqslant 1.$$

The least admissible constant is denoted by $\beta_{p,E}$.

It can be shown that if E is UMD for some $p \in (1, \infty)$, then it is actually UMD for all $p \in (1, \infty)$. Examples are the following (here $p' \in (1, \infty)$ denotes the conjugate of p):

- Every Hilbert space H is UMD with $\beta_{p,H} = \max\{p, p'\}$.
- If E is a UMD Banach space and $(S, \mathcal{A}, \mu)$ is a measure space, then $L^p(\mu; E)$ is a UMD space with $\beta_{p,L^p(\mu;E)} = \beta_{p,E}$.
- E is UMD if and only if E^* is UMD and it holds $\beta_{p,E} = \beta_{p',E^*}$.

In the case of UMD spaces, it is possible again to construct stochastic integrals in a suitable class of predictable processes and to obtain two-sided Burkholder inequalities.

Theorem A.15 (Theorem 5.5 from [57]). *Let E be a UMD Banach space and let $p \in (1, \infty)$. For all progressively measurable processes $\phi: \mathbb{R}_+ \times \Omega \to E$ we have*

$$\frac{1}{\beta_{p,E}} \|\phi\|_{L^p(\Omega;\gamma^p(L^2(\mathbb{R}_+),X))} \leqslant \mathbb{E}\Big[\Big\| \int_0^\infty \phi \,\mathrm{d}W \Big\|^p\Big]^{1/p} \leqslant \beta_{p,E}\, \|\phi\|_{L^p(\Omega;\gamma^p(L^2(\mathbb{R}_+),X))}.$$

In the above statement, $\gamma^p(L^2(\mathbb{R}_+), X)$ stands for the p-th γ-Radonifying norm; we omit the precise definition, which can be found in [56], [57]. There are special cases in which the γ-Radonifying norm is equivalent to other norms with a simpler expression, in particular when $E = L^q(\mu)$, in which case there is an isomorphism of Banach spaces

$$\gamma^p(L^2(\mathbb{R}_+), L^p(\mu)) = L^p(\mu; L^2(\mathbb{R}_+))$$

and so the previous inequality can be reformulated as

$$\begin{aligned}\mathbb{E}\Big[\Big\| \int_0^\infty \phi_s \,\mathrm{d}W_s \Big\|_{L^q(\mu)}^p\Big] &\sim_{p,q} \mathbb{E}\Big[\Big\|\Big(\int_0^\infty \phi_s^2(\cdot) \,\mathrm{d}s\Big)^{1/2} \Big\|_{L^q(\mu)}^p\Big] \\ &= \mathbb{E}\Big[\Big(\int_S \Big(\int_0^\infty \phi^2(s,x)\,\mathrm{d}s\Big)^{q/2} \mathrm{d}\mu(x)\Big)^{p/q}\Big].\end{aligned}$$

In the case $q \geqslant 2$, an application of Minkowski's inequality then yields

$$\mathbb{E}\Big[\Big\| \int_0^\infty \phi_s \,\mathrm{d}W_s \Big\|_{L^q(\mu)}^p\Big] \lesssim_{p,q} \mathbb{E}\Big[\Big(\int_0^\infty \|\phi(s,\cdot)\|_{L^q(\mu)}^2 \,\mathrm{d}s\Big)^{p/2}\Big],$$

which is consistent with the aforementioned results for martingale type 2 spaces. In the general case instead, assuming we want to estimate the $L^q(\mathbb{R}^d)$ norm of an averaged operator by means of the Itô–Tanaka formula (3.10), we would then need to estimate a term of the form (we omit the constants for simplicity)

$$\mathbb{E}\Big[\Big(\int_{\mathbb{R}^d}\Big(\int_s^t\Big(\int_u^t P_{|r-u|^{2H}}\nabla b(r,x+W_{u,r}^{2,H})|r-u|^{H-1/2}\,\mathrm{d}r\Big)^2\mathrm{d}u\Big)^{q/2}\mathrm{d}x\Big)^{p/q}\Big],$$

which we are currently not able to do. The techniques employed in Section 3.3 rely quite crucially on the simplifications given by a formula of the form (A.12).

Acknowledgements. We thank Mark Veraar and Simone Floreani for a very useful discussion on integration in UMD Banach spaces.

Funding. The authors were supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through the Hausdorff Center for Mathematics under Germany's Excellence Strategy -EXC-2047/1- 390685813 and through CRC 1060, projekt number 211504053.

References

[1] Amine, O., Baños, D. and Proske, F.: C^∞-regularization by noise of singular ODE's. Preprint 2018, arXiv: 1710.05760.

[2] Amine, O., Mansouri, A. R. and Proske, F.: Well-posedness of the deterministic transport equation with singular velocity field perturbed along fractional Brownian paths. Preprint 2020, arXiv: 2003.06200.

[3] Athreya, S., Butkovsky, O. and Mytnik, L.: Strong existence and uniqueness for stable stochastic differential equations with distributional drift. *Ann. Probab.* **48** (2020), no. 1, 178–210.

[4] Bahouri, H., Chemin, J. Y. and Danchin, R.: *Fourier analysis and nonlinear partial differential equations*. Grundlehren der Mathematischen Wissenschaften 343. Springer, Heidelberg, 2011.

[5] Bailleul, I. and Gubinelli, M.: Unbounded rough drivers. *Ann. Fac. Sci. Toulouse Math. (6)* **26** (2017), no. 4, 795–830.

[6] Baños, D, Nilssen, T. and Proske, F.: Strong existence and higher order Fréchet differentiability of stochastic flows of fractional Brownian motion driven SDEs with singular drift. *J. Dynam. Differential Equations* **32** (2020), no. 4, 1819–1866.

[7] Bayart, F. and Heurteaux, Y.: On the Hausdorff dimension of graphs of prevalent continuous functions on compact sets. In *Further developments in fractals and related fields*, 25–34. Trends Math., Birkhäuser/Springer, New York, 2013.

[8] Beck, L., Flandoli, F., Gubinelli, M. and Maurelli, M.: Stochastic ODEs and stochastic linear PDEs with critical drift: regularity, duality and uniqueness. *Electron. J. Probab.* **24** (2019), Paper no. 136, 72 pp.

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

[10] Butkovsky, O., Dareiotis, K. and Gerencsér, M.: Approximation of SDEs – a stochastic sewing approach. *Probab. Theory Related Fields* **181** (2021), no. 4, 975–1034.

[11] Catellier, R.: Rough linear transport equation with an irregular drift. *Stoch. Partial Differ. Equ. Anal. Comput.* **4** (2016), no. 3, 477–534.

[12] Catellier, R. and Gubinelli, M.: Averaging along irregular curves and regularisation of ODEs. *Stochastic Process. Appl.* **126** (2016), no. 8, 2323–2366.

[13] Chaudru de Raynal, P. E., Menozzi, S. and Priola, E.: Weak well-posedness of multidimensional stable driven SDEs in the critical case. *Stoch. Dyn.* **20** (2020), no. 6, 2040004, 20 pp.

[14] Christensen, J. P. R.: On sets of Haar measure zero in Abelian Polish groups. *Israel J. Math.* **13** (1972), no. 3-4, 255–260.

[15] Coutin, L., Duboscq, R. and Reveillac, A.: The Itô–Tanaka trick: a non-semimartingale approach. Preprint 2019, arXiv: 1907.03629.

[16] Da Prato, G. and Flandoli, F.: Pathwise uniqueness for a class of SDE in Hilbert spaces and applications. *J. Funct. Anal.* **259** (2010), no. 1, 243–267.

[17] Da Prato, G., Flandoli, F., Röckner, M. and Veretennikov, A. Y.: Strong uniqueness for SDEs in Hilbert spaces with nonregular drift. *Ann. Probab.* **44** (2016), no. 3, 1985–2023.

[18] Davie, A. M.: Uniqueness of solutions of stochastic differential equations. *Int. Math. Res. Not. IMRN* (2007), no. 24, Art. ID rnm124, 26 pp.

[19] Di Nezza, E., Palatucci, G. and Valdinoci, E.: Hitchhiker's guide to the fractional Sobolev spaces. *Bull. Sci. Math.* **136** (2012), no. 5, 521–573.

[20] DiPerna, R. J. and Lions, P.-L.: Ordinary differential equations, transport theory and Sobolev spaces. *Invent. Math.* **98** (1989), no. 3, 511–547.

[21] Flandoli, F.: *Random perturbation of PDEs and fluid dynamic models*. Lectures from the 40th Probability Summer School held in Saint-Flour, 2010. Lecture Notes in Mathematics 2015, Springer, Heidelberg, 2011.

[22] Flandoli, F.: Regularizing properties of Brownian paths and a result of Davie. *Stoch. Dyn.* **11** (2011), no. 2-3, 323–331.

[23] Flandoli, F., Gubinelli, M. and Priola, E.: Well-posedness of the transport equation by stochastic perturbation. *Invent. Math.* **180** (2010), no. 1, 1–53.

[24] Flandoli, F., Issoglio, E. and Russo, F.: Multidimensional stochastic differential equations with distributional drift. *Trans. Amer. Math. Soc.* **369** (2017), no. 3, 1665–1688.

[25] Fraser, J. M. and Hyde, J. T.: The Hausdorff dimension of graphs of prevalent continuous functions. *Real Anal. Exchange* **37** (2011/12), no. 2, 333–352.

[26] Fraysse, A.: Regularity criteria for almost every function in Sobolev spaces. *J. Funct. Anal.* **258** (2010), no. 6, 1806–1821.

[27] Fraysse, A. and Jaffard, S.: How smooth is almost every function in a Sobolev space? *Rev. Mat. Iberoam.* **22** (2006), no. 2, 663–682.

[28] Friz, P. K. and Hairer, M.: *A Course on rough paths. With an introduction to regularity structures*. Universitext, Springer, Cham, 2014.

[29] Galeati, L. and Gubinelli, M.: Prevalence of ρ-irregularity and related properties. Preprint 2020, arXiv: 2004.00872.

[30] Garsia, A. M., Rodemich, E. and Rumsey H.: A real variable lemma and the continuity of paths of some Gaussian processes *Indiana Univ. Math. J.* **20** (1970/71), no. 6, 565–578.

[31] Gerencsér, M.: Regularisation by regular noise. Preprint 2020, arXiv: 2009.08418.

[32] Harang, F. A. and Perkowski, N.: C^∞ regularization of ODEs perturbed by noise. *Stoch. Dyn.* **21** (2021), no. 8, Paper no. 2140010, 29 pp.

[33] Hu, Y. and Lê, K.: Nonlinear Young integrals and differential systems in Hölder media. *Trans. Amer. Math. Soc.* **369** (2017), no. 3, 1935–2002.

[34] Hunt, B. R.: The prevalence of continuous nowhere differentiable functions. *Proc. Amer. Math. Soc.* **122** (1994), no. 3, 711–717.

[35] Hunt, B. R., Sauer, T. and Yorke, J. A.: Prevalence: a translation-invariant "almost every" on infinite-dimensional spaces. *Bull. Amer. Math. Soc. (N.S.)* **27** (1992), no. 2, 217–238.

[36] Hytönen, T., van Neerven, J., Veraar, M. C. and Weis, L.: *Analysis in Banach spaces. Vol. II: Probabilistic methods and operator theory*. A Series of Modern Surveys in Mathematics 67, Springer, Cham, 2017.

[37] Krylov, N. V. and Röckner, M.: Strong solutions of stochastic equations with singular time dependent drift. *Probab. Theory Related Fields* **131** (2005), no. 2, 154–196.

[38] Lê, K.: A stochastic sewing lemma and applications. *Electron. J. Probab.* **25** (2020), Paper no. 38, 55 pp.

[39] Lejay, A.: Controlled differential equations as Young integrals: a simple approach. *J. Differential Equations* **249** (2010), no. 8, 1777–1798.

[40] Leoni, G.: *A first course in Sobolev spaces*. Second edition. Graduate Studies in Mathematics 181, American Mathematical Society, Providence, RI, 2017.

[41] Maurelli, M.: *Regularization by noise in finite dimension*. PhD thesis, Scuola Normale Superiore di Pisa, Pisa, 2016.

[42] Mourrat, J. C. and Weber, H.: Global well-posedness of the dynamic Φ^4 model in the plane. *Ann. Probab.* **45** (2017), no. 4, 2398–2476.

[43] Nilssen, T.: Rough linear PDEs with discontinuous coefficients – existence of solutions via regularization by fractional Brownian motion. *Electron. J. Probab.* **25** (2020), Paper no. 34, 33 pp.

[44] Nualart, D.: *The Malliavin calculus and related topics*. Second edition. Probability and its Applications (New York), Springer-Verlag, Berlin, 2006.

[45] Nualart, D. and Ouknine, Y.: Regularization of differential equations by fractional noise. *Stochastic Process. Appl.* **102** (2002), no. 1, 103–116.

[46] Ott, W. and Yorke, J.: Prevalence. *Bull. Amer. Math. Soc. (N.S.)* **42** (2005), no. 3, 263–290.

[47] Peres, Y. and Sousi, P.: Dimension of fractional Brownian motion with variable drift. *Probab. Theory Related Fields* **165** (2016), no. 3-4, 771–794.

[48] Picard, J.: Representation formulae for the fractional Brownian motion. In *Séminaire de probabilités XLIII*, 3–70. Lecture Notes in Math. 2006, Springer, Berlin, 2011.

[49] Priola, E.: Pathwise uniqueness for singular SDEs driven by stable processes. *Osaka J. Math.* **49** (2012), no. 2, 421–447.

[50] Samko, G. S., Kilbas, A. A. and Marichev, O. I.: *Fractional integrals and derivatives. Theory and applications.* Gordon and Breach Science Publishers, Yverdon, 1993.

[51] Seidler, J.: Exponential estimates for stochastic convolutions in 2-smooth Banach spaces. *Electron. J. Probab.* **15** (2010), 1556–1573.

[52] Shaposhnikov, A. V.: Some remarks on Davie's uniqueness theorem. *Proc. Edinb. Math. Soc. (2)* **59** (2016), no. 4, 1019–1035.

[53] Shaposhnikov, A. V.: Correction to the paper "Some remarks on Davie's uniqueness theorem". Preprint 2017, arXiv: 1703.06598.

[54] Shaposhnikov, A. V. and Wresch, L.: Pathwise vs. path-by-path uniqueness. Preprint 2020, arXiv: 2001.02869.

[55] Triebel, H.: *Theory of function spaces II*. Monographs in Mathematics 84, Birkhäuser Verlag, Basel, 1992.

[56] van Neerven, J., Veraar, M. C. and Weis, L.: Stochastic integration in UMD Banach spaces. *Ann. Probab.* **35** (2007), no. 4, 1438–1478.

[57] van Neerven, J., Veraar, M. C. and Weis, L.: Stochastic integration in Banach spaces – a survey. In *Stochastic analysis: a series of lectures*, 297–332. Progr. Probab. 68, Birkhäuser/Springer, Basel, 2015.

[58] Veretennikov, A. Y.: On strong solutions and explicit formulas for solutions of stochastic integral equations. *Sbornik: Mathematics* **39** (1981), no. 3, 387–403.

[59] Zvonkin, A. K.: A transformation of the phase space of a diffusion process that removes the drift. *Mat. Sb. (N.S.)* **93 (135)** (1974), no. 1, 129–149.

Received April 2, 2020; revised March 24, 2021. Published online July 6, 2021.

Lucio Galeati
Institute for Applied Mathematics & Hausdorff Center for Mathematics, Endenicher Allee 60, 53111 Bonn, Germany;
lucio.galeati@iam.uni-bonn.de

Massimiliano Gubinelli
Institute for Applied Mathematics & Hausdorff Center for Mathematics, Endenicher Allee 60, 53111 Bonn, Germany;
gubinelli@iam.uni-bonn.de