

DISTRIBUTION DEPENDENT SDES DRIVEN BY ADDITIVE FRACTIONAL BROWNIAN MOTION

LUCIO GALEATI, FABIAN A. HARANG, AND AVI MAYORCAS

ABSTRACT. We study distribution dependent stochastic differential equations with irregular, possibly distributional drift, driven by an additive fractional Brownian motion of Hurst parameter $H \in (0, 1)$. We establish strong well-posedness under a variety of assumptions on the drift; these include the choice

$$B(\cdot, \mu) = (f * \mu)(\cdot) + g(\cdot), \quad f, g \in B_{\infty, \infty}^{\alpha}, \quad \alpha > 1 - \frac{1}{2H},$$

thus extending the results by Catellier and Gubinelli [9] to the distribution dependent case. The proofs rely on some novel stability estimates for singular SDEs driven by fractional Brownian motion and the use of Wasserstein distances.

CONTENTS

1.	Introduction	1
2.	Main results	8
3.	SDEs driven by fBm	14
4.	Proofs of the main results	27
5.	Refined results in the convolutional case	32
	Appendix A. Some useful lemmas	40
	References	45

1. INTRODUCTION

In this work we consider a distribution dependent SDE (henceforth DDSDE) of the form

$$X_t = \xi + \int_0^t B_s(X_s, \mathcal{L}(X_s)) ds + W_t \tag{1.1}$$

where $B : \mathbb{R}_+ \times \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \rightarrow \mathbb{R}^d$, ξ is an \mathbb{R}^d -valued random variable and W is a \mathbb{R}^d -valued stochastic process independent of ξ . The drift B and the law of (ξ, W) are prescribed, while the process X is the unknown and $\mathcal{L}(X_t)$ denotes the law of its marginal at time t .

Date: June 10, 2022.

Key words and phrases. Distribution dependent SDEs, Singular drifts, Regularization by noise, Fractional Brownian motion.

AMS 2020 Mathematics Subject Classification: Primary: 60H50, 60H10; Secondary: 60G22, 60L90.

Acknowledgments. FH gratefully acknowledges financial support from the STORM project 274410, funded by the Research Council of Norway. LG is funded by the DFG under Germany's Excellence Strategy - GZ 2047/1, project-id 390685813.

Usually in the literature W is sampled as a standard Brownian motion; in this case the DDSDE is also called a McKean–Vlasov SDE, after the pioneering work [34] where it was first introduced.

The importance of McKean–Vlasov equations is due to their connection to systems of N particles subject to a mean field interaction of the form

$$X_t^{i,N} = \xi^i + \int_0^t B_s(X_s^{i,N}, L^N(X_s^{(N)})) ds + W_t^i, \quad L^N(X_t^{(N)}) := \frac{1}{N} \sum_{i=1}^N \delta_{X_t^{i,N}} \quad (1.2)$$

where (ξ^i, W^i) are typically taken to be i.i.d. copies of (ξ, W) and $L^N(X_t^{(N)})$ stands for the empirical measure of the system at time t . One expects the DDSDE (1.1) to be the mean field limit of (1.2) in the sense that, as N goes to infinity, $L^N(X_t^{(N)})$ converges weakly to $\mathcal{L}(X_t)$ with probability 1.

Another feature of DDSDEs in the Brownian noise case is their connection to nonlinear Fokker–Planck PDEs (also called McKean–Vlasov equations) of the form

$$\partial_t \rho + \nabla \cdot ((B_t(\cdot, \rho) \rho)) = \frac{1}{2} \Delta \rho, \quad \rho_0 = \mathcal{L}(\xi), \quad (1.3)$$

which describe the evolution of the marginal $\rho_t = \mathcal{L}(X_t)$; in particular, both (1.1) and (1.3) provide a macroscopic, compact description of the system (1.2), allowing one to reduce its complexity. For this reason, DDSDEs have found applications in numerous fields, see the review [27] and the references therein; let us also mention their connection to mean-field games [31].

Classical results concerning the well-posedness of the DDSDE (1.1) and the mean-field limit property go back to Sznitman [44] and Gärtner [20]; in the last years the field has witnessed substantial contributions both from the analytic and probabilistic communities. On the one hand, new methods based on entropy inequalities [14, 28, 6] and modulated energy methods [41, 42] have allowed for the rigorous derivation of mean field limits for fairly singular B ; while on the other, DDSDEs with irregular drifts are related to the flourishing field of *regularization by noise* phenomena. The latter topic was initiated by Zvonkin [49] and Veretennikov [47] in the case of standard SDEs, see [13] for a general overview; recently many authors have applied similar techniques in the DDSDE case, see for instance [5, 35, 40, 10, 25].

Contrary to the previously mentioned works, here we will study DDSDEs in which W is sampled as a *fractional Brownian motion* (fBm for short) of Hurst parameter $H \in (0, 1)$. Our main reasons for doing so are the following:

1. It was shown in [11], revisiting the ideas of Tanaka [45], that for Lipschitz B the mean-field limit of (1.2) to (1.1) holds for *any* choice of the process W , regardless of it being Markov or a semimartingale. In particular the DDSDE has a physical meaning and still provides a compact description of a much more complex system of interacting particles.
2. Several regularization by noise results for standard SDE are available for W sampled as an fBm (or similar fractional processes), see [37, 9, 32, 1, 3] for a short selection.

In light of Point 2. above, it is natural to expect similar results to hold for DDSDEs with singular (possibly even distributional in space) drifts and W sampled as an fBm; by Point 1., they are relevant in the study of particle systems with singular interactions (for instance with a discontinuity at the origin, as typical of Coulomb and Riesz-type potentials).

Let us mention that there is a certain degree of arbitrariness in choosing W to be sampled as an fBm, as one could consider other non-Markovian, non-martingale processes. We believe

our choice to be simple enough while at the same time representing what one might expect for a larger class of processes (e.g. Gaussian processes satisfying a local non-determinism condition). In this sense, this work also serves as a comparison to the results from [19], where we explored in detail the DDSDE (1.1) in the opposite regime where no assumption whatsoever is imposed on W , thus no regularization can be observed.

Despite the above motivations, singular DDSDEs driven by fBm (or similar fractional processes) so far have not received the same attention as their Brownian counterparts; to the best of our knowledge, the only previous work treating these kind of equations is [4]. After the first version of this manuscript came out, a different approach based on relative entropy methods has been proposed in [21].

One possible reason for this is the substantial new difficulties presented by such equations: fBm with parameter $H \neq 1/2$ is neither a Markov process, nor a semimartingale, so techniques based on Itô calculus are not applicable. This includes in particular the connection to parabolic semigroups, the martingale problem formulation and the use of Zvonkin transform (or Itô–Tanaka trick), all techniques used extensively in the aforementioned works in the Brownian case. It also prevents the use of standard arguments, which typically rely on establishing uniqueness of the law $\rho_t = \mathcal{L}(X_t)$ through PDE analysis of (1.3) and then fixing the law in the DDSDE and treating it as a standard SDE.

Treating DDSDEs driven by fBm thus requires a novel set of tools and ideas; our strategy in this paper builds on the work of Catellier and Gubinelli [9], which represented a major breakthrough in the study of standard SDEs driven by fBm of the form

$$X_t = \xi + \int_0^t b_s(X_s) ds + W_t. \quad (1.4)$$

Therein the authors develop a *pathwise approach* to the equation, based on *nonlinear Young integrals* and *Girsanov transform*, that allows to give meaning to (1.4) and establish its *path-by-path uniqueness*, for drifts b of poor regularity, possibly even distributional. Their results and techniques have been revisited in subsequent works [17, 23, 18, 22]; in general it suffices to require

$$b \in \begin{cases} L_T^q B_{\infty, \infty}^\alpha & \text{with } \alpha > 1 - \frac{1}{2H} + \frac{1}{Hq} & \text{if } H \leq 1/2 \\ C_T^{\alpha H} C_x^0 \cap C_T^0 C_x^\alpha & \text{with } \alpha > 1 - \frac{1}{2H} & \text{if } H > 1/2 \end{cases} \quad (1.5)$$

see for instance Theorem 15 and Corollary 2 from [17]. Here $B_{\infty, \infty}^\alpha$ denote Besov–Hölder spaces; see Section 1.1 below for the relevant definitions and notations in use throughout the article.

For the sake of exposition, let us ignore for the moment the additional time regularity required in (1.5) in the case $H > 1/2$, since it is mostly of a technical nature; then condition (1.5) roughly amounts to the drift b enjoying a spatial regularity $B_{\infty, \infty}^\alpha$ with $\alpha > 1 - 1/(2H)$. Observe that for all $H \in (0, 1)$ this includes values $\alpha < 1/2$, while for $H < 1/2$ we are even allowed to take $\alpha < 0$, namely distributional b . To the best of our knowledge, no work after [9] has improved on the allowed range of α .

With the above theory at hand, we can interpret the DDSDE (1.1) by rewriting it as

$$X_t = \xi + \int_0^t \bar{b}_s(X_s) ds + W_t, \quad \bar{b}_t(\cdot) := B_t(\cdot, \mathcal{L}(X_t));$$

namely, X solves the SDE with drift \bar{b} , in the Catellier-Gubinelli sense, where \bar{b} depends in a nontrivial way on the law of X itself. This interpretation comes with a natural fixed point formulation: given a process X , we can associate to it a “flow of measures” $\mu_t = \mathcal{L}(X_t)$ and a drift $b_t^\mu := B_t(\cdot, \mu_t)$, then solve the associated SDE, which gives a new process $Y = \mathcal{I}(X)$; thus X is a solution to (1.1) if and only if it is a fixed point for \mathcal{I} .

Alternatively, one could start with the flow of measures $\mu = \{\mu_t\}_{t \in [0, T]}$ and set up the fixed point procedure for this object, by defining $\mathcal{J}(\mu)_t = \mathcal{L}(X_t)$ for X solution to b^μ . These two interpretations are in fact equivalent: once μ is completely determined, the DDSDE reduces to a standard SDE with fixed drift b^μ , to which the previous results can be applied; see Lemma 4.4 for more details. Throughout the article we will exploit both interpretations whenever useful.

Given the above interpretation, we need two main ingredients to develop a solution theory:

1. Firstly, B must have the properties that b^μ satisfies (1.5) for any μ . of interest and that the solution-to-drift map $X(\mapsto \mu) \mapsto b^\mu$ is Lipschitz in suitable topology.
2. Secondly, we must develop stability estimates for the drift-to-solution map $b \mapsto Y$, in an appropriate topology that complements the stability of $\mu \mapsto b^\mu$.

Once these points are established, the contractivity of the overall map $X \mapsto b^\mu \mapsto \mathcal{I}(X)$ follows.

There are however major problems with the program outlined above; to describe them without too many technicalities, let us consider here the most relevant case $B(\mu) = f * \mu + g$ for time homogeneous $f, g \in B_{\infty, \infty}^\alpha$, $\alpha > 1 - 1/(2H)$. In this case, the map $\mu \mapsto b^\mu$ is naturally Lipschitz in the total variation topology, in the sense that

$$\|B(\mu^1) - B(\mu^2)\|_{B_{\infty, \infty}^\alpha} \lesssim \|\mu^1 - \mu^2\|_{TV};$$

however due to the lack of an underlying parabolic PDE (1.3) (and the associated maximum principle) in the fBm setting, it is not obvious how to control the drift-to-solution map $b \mapsto Y$ in this topology, i.e. how to bound $\|\mathcal{L}(Y_t^1) - \mathcal{L}(Y_t^2)\|_{TV}$ as a function of $\|b^1 - b^2\|_{B_{\infty, \infty}^\alpha}$.

One of the main intuitions of the current work, which allows us to overcome this difficulty, is the understanding that although the regularity $B_{\infty, \infty}^\alpha$ is needed in order to solve the SDE (1.4), one may establish stability estimates in the weaker norm $B_{\infty, \infty}^{\alpha-1}$. Roughly speaking, given two solutions X^1, X^2 to (1.4) associated to different initial data and drifts (ξ^i, b^i) , for any $p \in [1, \infty)$ we have

$$\mathbb{E} \left[\sup_{t \in [0, T]} |X_t^1 - X_t^2|^p \right]^{1/p} \lesssim \mathbb{E} [|\xi^1 - \xi^2|^p]^{1/p} + \|b^1 - b^2\|_{B_{\infty, \infty}^{\alpha-1}} \quad (1.6)$$

see Theorem 3.13 and Corollary 3.17 for the rigorous statements. This property is naturally analogous to standard ODE theory, where solvability requires b Lipschitz, but stability estimates are in the supremum norm.

In our setting, it implies that B only needs to enjoy some *multiscale regularity* of the form

$$\|B(\mu)\|_{B_{\infty, \infty}^\alpha} \lesssim 1, \quad \|B(\mu^1) - B(\mu^2)\|_{B_{\infty, \infty}^{\alpha-1}} \lesssim d(\mu^1, \mu^2)$$

for another notion of distance $d(\mu^1, \mu^2)$, possibly different from the total variation one. The right choice for d turns out to be the family of p -Wasserstein distances $d_p(\mu^1, \mu^2)$, which complements the bound (1.6) thanks to the basic property $d_p(\mathcal{L}(X_t^1), \mathcal{L}(X_t^2)) \leq \mathbb{E}[|X_t^1 - X_t^2|^p]^{1/p}$.

Overall, the newly found stability estimate (1.6) and the use of Wasserstein distance allow us to fulfill Points 1.-2. outlined above and to solve the DDSDE (2.1) for a large class of drifts B , see Theorems 2.4 and 2.5 for the precise statements; this includes the case $B(\mu) = f * \mu + g$ mentioned above.

For the sake of this preliminary discussion we have ignored the time regularity requirement in (1.5), but it does indeed play a relevant role, making the proofs a bit more technical and requiring us to treat the cases $H > 1/2$ and $H \leq 1/2$ slightly differently; see Section 4 for more details.

Let us stress that, since we are not allowed to use the same tools as in the Brownian setting, our results are not optimal for the choice $H = 1/2$, sharper ones being available for instance in [40, 25]. Nevertheless, they still provide some new insights, with the stability estimate (1.6) being new in this setting as well. This also partially answers the ongoing debate from [40, 25, 26] on whether the drift B should be taken Lipschitz continuous in the measure argument μ w.r.t. the total variation distance, the Wasserstein one or a weighted mix of the two: the use of Wasserstein distance allows the drift to be Lipschitz continuous in the different regularity scale $B_{\infty, \infty}^{\alpha-1}$, which is strictly negative in the regime $\alpha \in (0, 1)$, which is admissible in (1.5) for $H = 1/2$.

A major open problem coming from this work is the mean-field convergence (and associated propagation of chaos property) of the particle system (1.2) to (1.1), for the class of singular drifts for which we establish well-posedness of the DDSDE in Theorem 2.4. Our techniques are currently not enough to give a full answer; recently, several authors have investigated the Brownian setting using alternative tools based on Girsanov theorem and Large Deviations, see [30, 29, 46, 24]. Contrary to Itô calculus, these tools are available for fBm as well, thus we hope they may be of help in future investigations.

Another interesting question posed by the current work is whether our results can be further improved, in the sense of allowing values of $\alpha < 1 - 1/(2H)$, at least in some special cases. Theorems 2.6 and 2.7 suggest an affirmative answer for convolutional drifts $B(\mu) = b * \mu$, see also the discussion at the beginning of Section 5; this is in analogy with the Brownian case, where standard SDE theory requires roughly $b \in L_x^\infty$, but the nonlinear PDE (1.3) can be solved for roughly $b \in W_x^{-1, \infty}$.

We conclude this introduction with the structure of the paper. In Section 1.1 we introduce all relevant notations adopted in the paper and recall some well-known facts. Section 2 contains all our main results and Section 2.1 relevant examples of drifts B satisfying them. We present in detail the Catellier–Gubinelli theory of SDEs driven by fBm in Section 3, where we prove our main stability results (Theorem 3.13 and Corollary 3.17 from Section 3.3) as well as some new auxiliary results on the regularity of the law of solutions (Section 3.4). Sections 4 and 5 contain the proofs of our main results, respectively Theorems 2.4, 2.5, 2.6 and 2.7. Finally, we have included in Appendix A a collection of useful analytic lemmas used throughout the paper.

1.1. Notations, conventions and well-known facts. Throughout the article we will always work on a finite time interval $[0, T]$, although arbitrarily large; we will never deal with estimates

on the infinite interval $[0, +\infty)$. We write $a \lesssim b$ whenever there exists a constant $C > 0$ such that $a \leq Cb$. To stress the dependence $C = C(\lambda)$ on a particular parameter λ , we will write $a \lesssim_\lambda b$. For $p \in [1, \infty]$ and where it will not cause confusion, we write p' to denote the dual exponent to p , that is $1/p + 1/p' = 1$, with the interpretation $p = 1 \iff p' = \infty$.

Throughout the article, whenever not mentioned explicitly, we will consider an underlying probability space $(\Omega, \mathcal{F}, \mathbb{P})$; any σ -algebra appearing is assumed to be \mathbb{P} -complete. If Ω has a topological structure, then $\mathcal{B}(\Omega)$ denotes its Borel σ -algebra (again up to \mathbb{P} -completion).

We denote by $\mathbb{E}_{\mathbb{P}}$, or simply \mathbb{E} , expectation w.r.t. \mathbb{P} ; Given a Banach space E and $p \in [1, \infty]$, we will frequently consider E -valued random variables X in the space $L^p_{\Omega} E := L^p(\Omega, \mathcal{F}, \mathbb{P}; E)$, with norm $\|X\|_{L^p_{\Omega} E} = \mathbb{E}[\|X\|_E^p]^{1/p}$ (essential supremum if $p = \infty$).

We denote by $\mathcal{L}_{\mathbb{P}}(X)$, or simply $\mathcal{L}(X)$, the law of X on E , namely the pushforward measure $\mathbb{P} \circ X^{-1} = X\#\mathbb{P}$; more generally, we adopt the notation $F\#\mu$ for the pushforward of a measure μ under a measurable map F . Given a measure $\mu \in \mathcal{P}(C_T)$, we mention in particular the pushforward $\mu_t := e_t\#\mu$ where $e_t(h) = h_t$ denotes the evaluation map, $e_t : C_T \rightarrow \mathbb{R}^d$.

1.1.1. *Function spaces on $[0, T]$.* Given a metric space (M, d_M) , we denote by $C_T M = C([0, T]; M)$ the space of all continuous functions $f : [0, T] \rightarrow M$; for $\gamma \in (0, 1)$, we set $C_T^\gamma M = C^\gamma([0, T]; M)$ to be the subset of γ -Hölder continuous functions, namely

$$\llbracket f \rrbracket_{\gamma, M} := \sup_{s \neq t \in [0, T]} \frac{d_M(f_t, f_s)}{|t - s|^\gamma} < \infty.$$

If $(E, \|\cdot\|_E)$ is a Banach space, then $C_T E$ and $C_T^\gamma E$ are Banach spaces with norms

$$\|f\|_{C_T E} = \sup_{t \in [0, T]} \|f_t\|_E, \quad \|f\|_{C_T^\gamma E} = \|f\|_{C_T E} + \llbracket f \rrbracket_{\gamma, E}.$$

In the case $E = \mathbb{R}^n$ for some $n \in \mathbb{N}$, whenever it doesn't create confusion we will simply use C_T , C_T^γ and $\|f\|_\gamma$ in place of $C_T \mathbb{R}^d$, $C_T^\gamma \mathbb{R}^d$, $\|f\|_{C_T^\gamma}$; moreover for any $[s, t] \subset [0, T]$ we set

$$\llbracket f \rrbracket_{\gamma, [a, b]} := \sup_{s \neq t \in [a, b]} \frac{|f_t - f_s|}{|t - s|^\gamma}.$$

Given a Banach space E and $q \in [1, \infty]$, we denote by $L_T^q E = L^q(0, T; E)$ the Bochner–Lebesgue space of strongly measurable $f : [0, T] \rightarrow E$ such that

$$\|f\|_{L_T^q E} = \left(\int_0^T \|f_t\|_E^q dt \right)^{\frac{1}{q}} < \infty$$

with usual modification for $q = \infty$; as before we write L_T^q for $L_T^q \mathbb{R}^n$.

1.1.2. *Function spaces on \mathbb{R}^d .* Given $d, m \in \mathbb{N}$, we denote by $C(\mathbb{R}^d; \mathbb{R}^m)$ the space of continuous, bounded functions $f : \mathbb{R}^d \rightarrow \mathbb{R}^m$, endowed with the supremum norm $\|f\|_{C_x^0}$; whenever it doesn't create confusion we will simply write C_x^0 . $C_c^\infty = C_c^\infty(\mathbb{R}^d; \mathbb{R}^m)$, $C_x^n = C^n(\mathbb{R}^d; \mathbb{R}^m)$ denote respectively compactly supported smooth functions and n -times differentiable functions with continuous, bounded derivatives up to order n ; $\mathcal{S} = \mathcal{S}(\mathbb{R}^d; \mathbb{R}^m)$ denote Schwartz functions, \mathcal{S}' their dual. Given f , we denote by Df its Jacobian, i.e. the collection of first order derivatives

$(\partial_j f_i)_{i,j}$, possibly interpreted in the distributional sense. For $\alpha \in (0, 1)$, $C_x^\alpha = C^\alpha(\mathbb{R}^d; \mathbb{R}^m)$ stand for the Banach space of Hölder continuous functions, with norm

$$\|f\|_{C_x^\alpha} := \|f\|_{C_x^0} + \llbracket f \rrbracket_{C_x^\alpha}, \quad \llbracket f \rrbracket_{C_x^\alpha} := \sup_{x \neq y \in \mathbb{R}^d} \frac{|f(x) - f(y)|}{|x - y|^\alpha}.$$

The definition of C_x^α extends canonically to $\alpha \in (1, +\infty)$ by imposing that $f \in C_x^\alpha$ if $f \in C_x^{\lfloor \alpha \rfloor}$ and its derivatives of order $\lfloor \alpha \rfloor$ belong to $C_x^{\alpha - \lfloor \alpha \rfloor}$, where $\lfloor \alpha \rfloor$ denotes the integer part of α . We denote by $C_{loc}^\alpha = C_{loc}^\alpha(\mathbb{R}^d; \mathbb{R}^m)$ the vector space of all continuous $f : \mathbb{R}^d \rightarrow \mathbb{R}^m$ such that $\varphi f \in C_x^\alpha$ for all $\varphi \in C_c^\infty$; we say that $f^n \rightarrow f$ in C_{loc}^α if $\varphi f^n \rightarrow \varphi f$ in C_x^α for all $\varphi \in C_c^\infty$.

Given $\alpha \in \mathbb{R}$ and $p \in [1, \infty]$, we denote by $B_{p,p}^\alpha = B_{p,p}^\alpha(\mathbb{R}^d; \mathbb{R}^m)$ the associated (inhomogeneous) Besov space, given by distributions $f \in \mathcal{S}'$ such that

$$\|f\|_{B_{p,p}^\alpha} := \begin{cases} \left(\sum_{n=-1}^{+\infty} 2^{\alpha n p} \|\Delta_n f\|_{L^p}^p \right)^{\frac{1}{p}} < \infty & \text{for } p \in [1, \infty) \\ \sup_{n \geq -1} 2^{\alpha n} \|\Delta_n f\|_{L^\infty} < \infty & \text{for } p = \infty \end{cases}$$

where Δ_n denote the Littlewood–Paley blocks associated to a partition of the unity. We refer to the monograph [2] for details on Besov spaces; throughout the paper we will frequently employ their properties, like Besov embeddings, Bernstein estimates for $\Delta_n f$ or the regularity of $f * g$ for f, g in different Besov spaces. Let us also mention that, although the Littlewood–Paley definition will be the most relevant for our purposes, Besov spaces admit alternative equivalent characterizations based on either interpolation or Gagliardo–Nirenberg type integral seminorms, see for instance [33]. For $\alpha \in \mathbb{R}_+ \setminus \mathbb{N}$, the spaces C_x^α and $B_{\infty,\infty}^\alpha$ coincide; however for clarity we will continue to write C_x^α for $\alpha \geq 0$ and $B_{\infty,\infty}^\alpha$ otherwise.

The notations from this section and the previous one can be combined to define $C_T^\gamma C_x^\alpha, L_T^q B_{p,p}^\alpha$, etc.; similarly, we define $C_T^\gamma C_{loc}^\alpha$ to be the vector space of all $f : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^m$ such that $\varphi f \in C_T^\gamma C_x^\alpha$ for all $\varphi \in C_c^\infty$, with convergence $f^n \rightarrow f$ in $C_T^\gamma C_{loc}^\alpha$ if $\varphi f^n \rightarrow \varphi f$ in $C_T^\gamma C_x^\alpha$. Given a function f of time and space, Df always denotes its Jacobian in the space variable only.

1.1.3. Probability measures and Wasserstein distance. Given a separable Banach space E , we denote by $\mathcal{P}(E)$ the set of probability measures over E ; we write $\mu^n \rightharpoonup \mu$ for weak convergence of measures, in the sense of testing against continuous bounded functions.

Given $\mu, \nu \in \mathcal{P}(E)$, $\Pi(\mu, \nu)$ stands for the set of all possible couplings of (μ, ν) , i.e. the subset of $\mathcal{P}(E \times E)$ with first and second marginals given respectively by μ and ν . For any $p \in [1, \infty)$, we define

$$d_p(\mu, \nu) := \inf_{m \in \Pi(\mu, \nu)} \left(\int_{E \times E} \|x - y\|_E^p m(dx, dy) \right)^{1/p}$$

which is a well defined quantity (possibly taking value $+\infty$). By [48, Theorem 4.1], an optimal coupling $\bar{m} \in \Pi(\mu, \nu)$ realizing the above infimum always exists.

Similarly we define $\mathcal{P}_p(E)$ to be set of p -integrable probability measures; that is, $\mu \in \mathcal{P}_p(E)$ if $\mu \in \mathcal{P}(E)$ and

$$\|\mu\|_p := \left(\int_E \|x\|_E^p \mu(dx) \right)^{1/p} < \infty.$$

It is well known that $d_p(\mu, \nu) < \infty$ for $\mu, \nu \in \mathcal{P}_p(E)$ and that $(\mathcal{P}_p(E), d_p)$ is a complete metric space, usually referred to as the p -Wasserstein space on E ; let us stress however that our definition of $d_p(\mu, \nu)$ holds for all $\mu, \nu \in \mathcal{P}(E)$. We recall that, given a sequence $\{\mu^n\}_n \subset \mathcal{P}_p(E)$, $d_p(\mu^n, \mu) \rightarrow 0$ is equivalent to $\mu^n \rightharpoonup \mu$ weakly and $\|\mu^n\|_p \rightarrow \|\mu\|_p$, see [48, Theorem 6.9].

Given $\mu \in \mathcal{P}(\mathbb{R}^d)$, with a slight abuse of notation we will write $\mu \in L^q(\mathbb{R}^d)$ (or simply L_x^q) for $q \in [1, \infty]$ to indicate that μ admits a density $\mu(dx) = \rho(x) dx$ with respect to the d -dim. Lebesgue measure, such that $\rho \in L_x^q$.

1.1.4. *Fractional Brownian motion.* A real valued continuous process $\{W_t, t \in [0, T]\}$ is a fractional Brownian motion (fBm) with Hurst parameter $H \in (0, 1)$ if it is a centered Gaussian process with covariance function

$$\mathbb{E}[W_t W_s] = \frac{1}{2}(|t|^{2H} + |s|^{2H} - |t - s|^{2H});$$

an \mathbb{R}^d -valued process W is a d -dimensional fBm if its components are independent 1-dimensional fBms. All the results we are going to recall here are classical and can be found in [36, 39].

For $H = 1/2$, fBm corresponds to classical Brownian motion (Bm), but for $H \neq 1/2$ it is not a semimartingale nor a Markov process; its trajectories are \mathbb{P} -a.s. in $C_T^{H-\varepsilon}$ for any $\varepsilon > 0$.

Given an fBm W of parameter H on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, it's always possible to construct a standard Bm B on it such that the following *canonical representation* holds:

$$W_t = \int_0^t K_H(t, s) dB_s$$

where K_H is a Volterra-type kernel and B and W generate the same filtration. Given a filtration $\{\mathcal{F}_t\}_{t \in [0, T]}$, we say that W is an \mathcal{F}_t -fBm if the associated B is an \mathcal{F}_t -Bm in the classical sense.

Closely related to the canonical representation are a version of Girsanov theorem for fBm (see e.g. [37, Theorem 2]) and the *strong local non-determinism* (LND) of fBm: for any $H \in (0, 1)$ there exists $c_H > 0$ such that

$$\text{Var}[W_t | \mathcal{F}_s] \geq c_H |t - s|^{2H} I_d \quad \forall t > s.$$

The LND property plays a key role in establishing the regularising features of W , cf. [23, 16].

2. MAIN RESULTS

Let us recall that the focus here is an abstract DDSDE of the form

$$X_t = \xi + \int_0^t B_s(X_s, \mu_s) ds + W_t, \quad \mu_t = \mathcal{L}(X_t) \quad \forall t \in [0, T] \quad (2.1)$$

where $\mathcal{L}(\xi) = \mu_0$, ξ independent of W and W is sampled as a fBm of parameter $H \in (0, 1)$.

We want to identify general conditions for measurable drifts $B : [0, T] \times \mathcal{P}_p(\mathbb{R}^d) \rightarrow B_{\infty, \infty}^\alpha$, $\alpha \in \mathbb{R}$, such that we can develop a solution theory for (2.1). As explained in the introduction, our strategy consists in setting up a fixed point for $\mu \mapsto b_t^\mu := B_t(\mu_t) \mapsto X \mapsto \tilde{\mu}_t := \mathcal{L}(X_t)$.

To this end, the assumptions on B should enforce two facts: for any flow of measures $\mu \in C_T \mathcal{P}_p$, the associated drift $b_t^\mu := B_t(\mu)$ is regular enough to solve (1.4), namely b^μ must satisfy condition (1.5); the map $\mu \mapsto b^\mu$ should be stable in suitable topologies. Last but not least, the eligible B should include cases of particular interest (most notably $B(\mu) = b * \mu$), see Section 2.1 below.

Corresponding to the above requirements, for $H > 1/2$ we define the following space:

Definition 2.1. For $\alpha, \beta \in (0, 1)$ and $p \in [1, \infty)$, let $\mathcal{H}_p^{\beta, \alpha}$ denote the class of continuous functions $B : [0, T] \times \mathbb{R}^d \times \mathcal{P}_p(\mathbb{R}^d) \rightarrow \mathbb{R}^d$ satisfying the following condition: there exists $C > 0$ such that

- i. For all $(t, x, \mu) \in [0, T] \times \mathbb{R}^d \times \mathcal{P}_p(\mathbb{R}^d)$, $|B_t(x, \mu)| \leq C$.
- ii. For all $(s, t) \in [0, T]^2$, $(x, y) \in (\mathbb{R}^d)^2$ and $(\mu, \nu) \in \mathcal{P}_p(\mathbb{R}^d) \times \mathcal{P}_p(\mathbb{R}^d)$, we have

$$|B_t(x, \mu) - B_s(y, \nu)| \leq C(|t - s|^{\alpha\beta} + |x - y|^\alpha + d_p(\mu, \nu)^\alpha).$$

- iii. For all $t \in [0, T]$ and $\mu, \nu \in \mathcal{P}_p(\mathbb{R}^d)$

$$\|B_t(\cdot, \mu) - B_t(\cdot, \nu)\|_{B_{\infty, \infty}^{\alpha-1}} \leq C d_p(\mu, \nu).$$

Whenever it does not create confusion, we will simply denote by $\|B\|$ the optimal constant C .

Corresponding to the above requirements, for $H \leq 1/2$ we define the following space:

Definition 2.2. For $\alpha \in \mathbb{R}$, $p \in [1, \infty)$ and $q \in [1, \infty]$, let $\mathcal{G}_p^{q, \alpha}$ denote the class of measurable functions $B : [0, T] \times \mathcal{P}_p(\mathbb{R}^d) \rightarrow B_{\infty, \infty}^\alpha$ satisfying the following condition: there exists $h \in L_T^q$ such that

- i. For all $(t, \mu) \in [0, T] \times \mathcal{P}_p(\mathbb{R}^d)$, we have $\|B_t(\mu)\|_{B_{\infty, \infty}^\alpha} \leq h_t$.
- ii. For all $(t, \mu, \nu) \in [0, T] \times \mathcal{P}_p(\mathbb{R}^d) \times \mathcal{P}_p(\mathbb{R}^d)$, we have $\|B_t(\mu) - B_t(\nu)\|_{B_{\infty, \infty}^{\alpha-1}} \leq h_t d_p(\mu, \nu)$.

Whenever it does not create confusion, we will simply denote by $\|B\|$ the optimal constant $\|h\|_{L_T^q}$.

Remark 2.3. It is readily checked that for $\alpha \leq \tilde{\alpha}$, $p \geq \tilde{p}$ and $q \leq \tilde{q}$ we have $\mathcal{G}_{\tilde{p}}^{\tilde{q}, \tilde{\alpha}} \subset \mathcal{G}_p^{q, \alpha}$. Similarly, for $\alpha \leq \tilde{\alpha}$, $\beta \leq \tilde{\beta}$ and $p \geq \tilde{p}$ it holds $\mathcal{H}_{\tilde{p}}^{\tilde{\beta}, \tilde{\alpha}} \subset \mathcal{H}_p^{\beta, \alpha}$.

Roughly speaking, we say that X is a solution to the DDSDE (2.1) if, setting $b_t^\mu := B_t(\mathcal{L}(X_t))$, then X is a solution to the standard SDE (1.4) associated to b^μ , being interpreted in the Catellier–Gubinelli sense whenever b^μ is singular; the pathwise theory for singular SDEs will be recalled in detail in Section 3. All the concepts of strong existence, pathwise uniqueness and uniqueness in law for DDSDEs then follow from the standard ones, see Definition 4.2 from Section 4.2.

Our first main result is the well-posedness of DDSDE (2.1) under suitable conditions on B ; it can be seen as an extension of [17, Theorem 15] to the distribution dependent case. The proof of the following theorem is given in Sections 4.1 and 4.2.

Theorem 2.4. Let $H > 1/2$ and let $B \in \mathcal{H}_p^{H, \alpha}$ for parameters

$$\alpha > 1 - \frac{1}{2H} > 0, \quad p \in [1, \infty). \quad (2.2)$$

Then for any $\mu_0 \in \mathcal{P}_p(\mathbb{R}^d)$, strong existence, pathwise uniqueness and uniqueness in law hold for the DDSDE (2.1).

Similarly, let $H \leq 1/2$ and let $B \in \mathcal{G}_p^{q, \alpha}$ for parameters

$$\alpha > 1 + \frac{1}{Hq} - \frac{1}{2H}, \quad \alpha \in \mathbb{R}, \quad q \in (2, \infty], \quad p \in [1, \infty). \quad (2.3)$$

Then for any $\mu_0 \in \mathcal{P}_p(\mathbb{R}^d)$, strong existence, pathwise uniqueness and uniqueness in law hold for the DDSDE (2.1).

Given a DDSDE (2.1), we will consider either (ξ, B) or (μ_0, B) to be the data of the problem, where we recall that $\mathcal{L}(\xi) = \mu_0$. As already mentioned in the introduction, the solution X is entirely determined by the associated flow of measures $\mu \in C_T\mathcal{P}_p$ given by $\mu_t = \mathcal{L}(X_t)$: once this is known, the drift $b_t^\mu = B_t(\mu_t)$ is determined as well and so we can reconstruct the strong solution X (or construct another copy of it on any probability space of interest). For this reason, it is quite useful to regard $\mu \in C_T\mathcal{P}_p$ to be itself a solution to the DDSDE; the exact equivalence between μ and X will be discussed rigorously in Lemma 4.4 from Section 4.2.

The next theorem provides stability estimates for the data-to-solution map $(\mu_0, B) \mapsto \mu$ (respectively $(\xi, B) \mapsto X$), showing that it is locally Lipschitz. The next theorem is proved in Section 4.3.

Theorem 2.5. *Let $\mu_0, \nu_0 \in \mathcal{P}_p$ for some $p \in [1, \infty)$. Then the following holds:*

- i. *For $H > 1/2$, let B^1, B^2 , be drifts in $\mathcal{H}_p^{H,\alpha}$ with parameters satisfying (2.2) and let $M > 0$ be a constant such that $\|B^i\| \leq M$. Then there exists a constant $C = C(\alpha, H, T, M, p)$ such that, for any $\mu_0^i \in \mathcal{P}_p(\mathbb{R}^d)$, the associated solutions $\mu^i \in C_T\mathcal{P}_p$ satisfy*

$$\sup_{t \in [0, T]} d_p(\mu_t^1, \mu_t^2) \leq C(d_p(\mu_0^1, \mu_0^2) + \|B^1 - B^2\|_\infty), \quad (2.4)$$

where

$$\|B^1 - B^2\|_\infty := \sup_{(t, \mu) \in [0, T] \times \mathcal{P}_p} \|B^1(t, \mu) - B^2(t, \mu)\|_{B_{\infty, \infty}^{\alpha-1}}.$$

If X^1, X^2 are two associated solutions, in the sense of stochastic processes, defined on the same probability space, then there exists $\gamma > 1/2$ such that

$$\mathbb{E} \left[\|X^1 - X^2\|_{\gamma; [0, T]}^p \right]^{1/p} \leq C(\|\xi^1 - \xi^2\|_{L_\Omega^p} + \|B^1 - B^2\|_\infty). \quad (2.5)$$

- ii. *For $H \leq 1/2$, let B^1, B^2 , be drifts in $\mathcal{G}_p^{q,\alpha}$ with parameters satisfying (2.3) and let $M > 0$ be a constant such that $\|B^i\| \leq M$. Then there exists a constant $C = C(\alpha, H, T, M, p, q)$ such that, for any $\mu_0^i \in \mathcal{P}_p(\mathbb{R}^d)$, the associated solutions $\mu^i \in C_T\mathcal{P}_p$ satisfy*

$$\sup_{t \in [0, T]} d_p(\mu_t^1, \mu_t^2) \leq C(d_p(\mu_0^1, \mu_0^2) + \|B^1 - B^2\|_{q, \infty}). \quad (2.6)$$

where

$$\|B^1 - B^2\|_{q, \infty} := \left(\int_0^T \sup_{\mu \in \mathcal{P}_p} \|B^1(t, \mu) - B^2(t, \mu)\|_{B_{\infty, \infty}^{\alpha-1}}^q dt \right)^{1/q}.$$

If X^1, X^2 are two associated solutions, in the sense of stochastic processes, defined on the same probability space, then there exists $\gamma > 1/2$ such that

$$\mathbb{E} \left[\|X^1 - X^2\|_{\gamma; [0, T]}^p \right]^{1/p} \leq C(\|\xi^1 - \xi^2\|_{L_\Omega^p} + \|B^1 - B^2\|_{q, \infty}). \quad (2.7)$$

As the settings of Theorems 2.4 and 2.5 are very general, they do not allow one to exploit any specific structure of the DDSDE in consideration to obtain sharper results. A prototypical example of such structure, which arises in many practical applications, is given by convolutional drifts $B_t(x, \mu) := (b_t * \mu)(x)$. The associated DDSDE takes the form

$$X_t = \xi + \int_0^t (b_s * \mathcal{L}(X_s))(X_s) ds + W_t \quad \forall t \in [0, T]. \quad (2.8)$$

As before we allow the drift b to be distributional, at least of the form $b \in L_T^1 B_{p,p}^\alpha$ for some $\alpha \in \mathbb{R}$, $p \in [1, \infty]$; at this stage pointwise evaluation of $b_s * \mathcal{L}(X_s)$ is not meaningful, instead we again interpret the equation in the Catellier–Gubinelli sense.

The heuristic idea behind the next results is that we can use the convolutional structure in a recursive way: assuming we are given a solution X with sufficiently regular law $\mathcal{L}(X_\cdot)$, this in turn leads to an improved regularity for the effective drift $b * \mathcal{L}(X_\cdot)$, compared to the original b . The argument can be made rigorous by establishing a priori estimates and working with smooth approximations; as a result, we are able to establish well-posedness for (2.8) in situations where the general Theorem 2.4 does not apply.

In both results we are going to present, we will need some additional regularity for the initial data μ_0 , in the form of an integrability assumption. This is because, as explained in the introduction, the lack of an underlying parabolic PDE prevents us from proving a smoothing effect at strictly positive times analogous to that of parabolic equations; rather, in order to develop a priori estimates, we will show that such integrability is propagated by the dynamics.

The next result shows existence and uniqueness of solutions to (2.8) in a suitable class, under an additional condition on $\operatorname{div} b$, which is by now quite standard since the pioneering work [12]. The proof of the next theorem is given in Section 5.1.

Theorem 2.6. *Let $H \in (0, 1)$, $q \in (2, \infty]$, $p \in [1, \infty]$, p' its conjugate exponent. Assume that $\operatorname{div} b \in L_T^1 L_x^\infty$ and either:*

- i. if $H > 1/2$, then $b \in C_T^{\alpha H} L_x^p \cap C_T^0 B_{p,p}^\alpha$ for some $\alpha > 1 - \frac{1}{2H}$;*
- ii. if $H \leq 1/2$, then $b \in L_T^q B_{p,p}^\alpha$ with $1 > \alpha > 1 - \frac{1}{2H} + \frac{1}{Hq}$.*

Then for any $\mu_0 \in L_x^{p'}$ there exists a strong solution to (2.8), which satisfies

$$\sup_{t \in [0, T]} \|\mathcal{L}(X_t)\|_{L_x^{p'}} < \infty; \quad (2.9)$$

moreover uniqueness holds, both pathwise and in law, in the class of solutions satisfying (2.9).

Our second result in the convolutional case is established under $L_T^q L_x^p$ -type assumptions on b ; here instead of relying on a bound for $\operatorname{div} b$, we exploit Girsanov-based arguments to establish integrability of $\mathcal{L}(X_t)$. This technique however only works in the regime $H \leq 1/2$. Section 5.2 contains the proof of the following result.

Theorem 2.7. *Let $d \geq 2$, $H \leq 1/2$, $(r, p, q) \in [1, \infty)^2 \times (2, \infty]$ be such that*

$$r > \frac{d}{d-1}, \quad \frac{1}{q} + \frac{Hd}{p} < \frac{1}{2}. \quad (2.10)$$

Then for any $b \in L_T^q L_x^p$ and $\mu_0 \in L_x^r$, there exists a strong solution to (2.1), which satisfies

$$\sup_{t \in [0, T]} \|\mathcal{L}(X_t)\|_{L_x^{\tilde{r}}} < \infty \quad \forall \tilde{r} < r; \quad (2.11)$$

moreover uniqueness holds, both pathwise and in law, in the class of solutions satisfying (2.11).

Remark 2.8. Condition (2.10) can be generalized in a way that allows values $r \leq d/(d-1)$ and that applies for $d = 1$, see Theorem 5.9 in Section 5.2 for more details. We warn the reader not to interpret Theorems 2.6 and 2.7 as full pathwise uniqueness (resp. uniqueness in law) statements: in general they do not exclude the existence of irregular solutions X which do not satisfy condition (2.9) (resp. (2.11)). However, as the proofs show, any solution constructed as

the limit of smooth drifts $b^n \rightarrow b$ does satisfy (2.9) (resp. (2.11)), thus it is the only physical solution to the DDSDE (2.1).

2.1. Examples. To illustrate the variety of situations to which Theorems 2.4 and 2.5 apply, we provide here several examples of functions contained in $\mathcal{G}_p^{q,\alpha}$ and $\mathcal{H}_p^{\beta,\alpha}$.

Example 2.9. Let $\alpha \in \mathbb{R}$, and for any $y \in \mathbb{R}^d$, $b : [0, T] \times \mathbb{R}^d \rightarrow B_{\infty, \infty}^\alpha$ be a measurable map, $b_t(\cdot, y) := b_t(y)(\cdot)$, and suppose there exists $h \in L_T^q$ for some $q \in [1, \infty]$ such that

$$\|b_t(\cdot, y)\|_{B_{\infty, \infty}^\alpha} \leq h_t, \quad \|b_t(\cdot, y) - b_t(\cdot, y')\|_{B_{\infty, \infty}^{\alpha-1}} \leq h_t |y - y'| \quad \forall t \in [0, T], (y, y') \in \mathbb{R}^{2d}.$$

Define now another measurable map $B : [0, T] \times \mathcal{P}(\mathbb{R}^d) \rightarrow B_{\infty, \infty}^\alpha$ by

$$B_t(\cdot, \mu) := \int_{\mathbb{R}^d} b_t(\cdot, y) \mu(dy), \quad \forall (t, \mu) \in [0, T] \times \mathcal{P}_p(\mathbb{R}^d)$$

where the integral is meaningful in the Bochner sense; then $B \in \mathcal{G}_p^{q,\alpha}$ for any $p \in [1, \infty)$.

Indeed, by the hypothesis on b , it is readily checked that

$$\|B_t(\cdot, \mu)\|_\alpha \leq \int_{\mathbb{R}^d} \|b_t(\cdot, y)\|_\alpha \mu(dy) \leq h_t \quad \forall t \in [0, T], \mu \in \mathcal{P}(\mathbb{R}^d);$$

given $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$, let $m \in \mathcal{P}(\mathbb{R}^{2d})$ be an optimal coupling for $d_1(\mu, \nu)$, then

$$\|B_t(\cdot, \mu) - B_t(\cdot, \nu)\|_{B_{\infty, \infty}^{\alpha-1}} \leq \int_{\mathbb{R}^{2d}} \|b_t(\cdot, y) - b_t(\cdot, y')\|_{B_{\infty, \infty}^{\alpha-1}} m(dy, dy') \leq h_t \int_{\mathbb{R}^{2d}} |y - y'| m(dy, dy')$$

which implies that

$$\|B_t(\cdot, \mu) - B_t(\cdot, \nu)\|_{B_{\infty, \infty}^{\alpha-1}} \leq h_t d_1(\mu, \nu) \leq h_t d_p(\mu, \nu) \quad \forall p \in [1, \infty).$$

Example 2.10. Given $\alpha, \beta \in (0, 1)$, assume that $b : [0, T] \times \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ satisfies

$$|b_t(x, y)| \leq C, \quad |b_t(x, y) - b_s(x', y')| \leq C(|t - s|^{\alpha\beta} + |x - x'|^\alpha + |y - y'|^\alpha)$$

for some $C > 0$, uniformly over s, t, x, x', y, y' ; we can identify b with the map $b : [0, T] \times \mathbb{R}^d \rightarrow C_x^\alpha = B_{\infty, \infty}^\alpha$ given by $(t, y) \mapsto b_t(\cdot, y)$. Assume additionally that for the same constant C it holds

$$\|b_t(\cdot, y) - b_t(\cdot, y')\|_{B_{\infty, \infty}^{\alpha-1}} \leq C |y - y'|$$

and define $B : [0, T] \times \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \rightarrow \mathbb{R}^d$ by

$$B_t(x, \mu) := \int_{\mathbb{R}^d} b_t(x, y) \mu(dy).$$

Then $B \in \mathcal{H}_p^{\beta,\alpha}$ for any $p \in [1, \infty)$. The verification of Conditions i. and iii. of Definition 2.1 is identical to that of Example 2.9, so we only need to focus on Condition ii. for $p = 1$.

Given $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$, let m be an optimal coupling for $d_1(\mu, \nu)$, then

$$\begin{aligned} |B_t(x, \mu) - B_s(x', \nu)| &= \left| \int_{\mathbb{R}^d} b_t(x, y) \mu(dy) - \int_{\mathbb{R}^d} b_s(x', y') \nu(dy') \right| \\ &\leq \int_{\mathbb{R}^{2d}} |b_t(x, y) - b_s(x', y')| m(dy, dy') \\ &\leq C \left(|t - s|^{\alpha\beta} + |x - x'|^\alpha + \int_{\mathbb{R}^{2d}} |y - y'|^\alpha m(dy, dy') \right) \\ &\leq C (|t - s|^{\alpha\beta} + |x - x'|^\alpha + d_1(\mu, \nu)^\alpha) \end{aligned}$$

where in the last step we used Jensen's inequality and the optimality of m .

Example 2.11. Consider now $B_t(\cdot, \mu) := b_t * \mu$, where $b \in L_T^q B_{\infty, \infty}^\alpha$ for some $\alpha \in \mathbb{R}$; then $B \in \mathcal{G}_p^{q, \alpha}$ for any $p \in [1, \infty)$.

Indeed, the verification of Condition *i.* from Definition 2.2 is the same as in Example 2.9, where now we can take $h = \|b\|_{B_{\infty, \infty}^\alpha} \in L_T^q$. Moreover by Lemma A.7 in Appendix A, for any $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$ and any $p \in [1, \infty)$ it holds

$$\|B_t(\cdot, \mu) - B_t(\cdot, \nu)\|_{B_{\infty, \infty}^{\alpha-1}} = \|b_t * (\mu - \nu)\|_{B_{\infty, \infty}^{\alpha-1}} \lesssim \|b_t\|_{B_{\infty, \infty}^\alpha} d_p(\mu, \nu).$$

Similarly, given $\alpha, \beta \in (0, 1)$, let $b \in C_T^{\alpha\beta} C_x^0 \cap C_T C_x^\alpha$ and set $B_t(\cdot, \mu) = b_t * \mu$; then $B \in \mathcal{H}_p^{\beta, \alpha}$ for any $p \in [1, \infty)$.

The verification of Conditions *i.* and *ii.* from Definition 2.1 follows from Example 2.10, as we can simply set $\tilde{b}_t(x, y) := b_t(x - y)$ and apply the calculations therein to \tilde{b} . Condition *i.* instead follows as above from an application of Lemma A.7.

Finally, let us point out that all the computations carry over to the case $B_t(\cdot, \mu) = b_t^1 * \mu + b_t^2$ for $b^i \in L_T^q B_{\infty, \infty}^\alpha$ (resp. $b^i \in C_T^{\alpha\beta} C_x^0 \cap C_T C_x^\alpha$).

Example 2.12. Let $b : [0, T] \times \mathbb{R}^d \rightarrow B_{\infty, \infty}^\alpha$ be as in Example 2.9 and $\phi : \mathbb{R}^d \rightarrow \mathbb{R}^d$ be a globally Lipschitz with constant $[\phi]_{Lip}$; define $B : [0, T] \times \mathcal{P}_1(\mathbb{R}^d) \rightarrow B_{\infty, \infty}^\alpha$ by

$$B_t(\cdot, \mu) = b_t(\cdot, \langle \phi, \mu \rangle), \quad \text{where } \langle \phi, \mu \rangle := \int_{\mathbb{R}^d} \phi(x) \mu(dx).$$

Then $B \in \mathcal{G}_p^{q, \alpha}$ for any $p \in [1, \infty)$. Similarly, given b as in Example 2.10, with B defined as above, it is easy to verify that $B \in \mathcal{H}_p^{\alpha, \beta}$ for any $p \in [1, \infty)$.

As a prototypical example, one may consider $b \in B_{\infty, \infty}^\alpha$ and define

$$B_t(\cdot, \mu) = B(\cdot, \mu) := b(\cdot - \langle x, \mu \rangle) \quad \text{where } \langle x, \mu \rangle := \int_{\mathbb{R}^d} x \mu(dx)$$

in which case, similarly to before, it holds $B \in \mathcal{G}_p^{q, \alpha}$ for any $q \in [1, \infty]$ (resp. $B \in \mathcal{H}_p^{\beta, \alpha}$ for any $\beta \in (0, 1)$) and $p \in [1, \infty)$.

We highlight that this class of examples are quite important since B is only defined on $\mathcal{P}_1(\mathbb{R}^d)$ and not on the whole $\mathcal{P}(\mathbb{R}^d)$, thus making the use of other notions of distance between measures (e.g. total variation norm) more difficult to handle. It can be further generalized to the case $\phi : \mathbb{R}^d \rightarrow \mathbb{R}^m$ for another $m \in \mathbb{N}$ (namely, B is determined by m statistics associated to μ) or to dependence on p -moments like $B_t(\cdot, \mu) = b_t(\cdot, \|\mu\|_p)$ for $\mu \in \mathcal{P}_p(\mathbb{R}^d)$; for $p > 1$ we can also allow ϕ to grow more than linearly at infinity.

3. SDES DRIVEN BY FBM

In this section we revisit the theory of singular SDEs driven by fBm, in order to derive useful estimates to apply later to the DDSDE setting. Sections 3.1 and 3.2 serve as a recap of key facts, respectively the pathwise meaning of singular SDEs and the regularising properties of fractional Brownian motion. Sections 3.3 and 3.4 instead provide novel results, Theorem 3.13 being the most important for our purposes.

Although the material of Sections 3.1-3.2 is strongly based on the works previous [9, 17, 15, 23], we felt obliged to provide the proofs of several key results for technical but rather important reasons. On the one hand, the aforementioned works are focused entirely on a pathwise setting, never establishing clear probabilistic concepts of solutions (cf. Definitions 3.4-3.5 below); on the other hand, previously singular drifts $b \in L_T^q B_{\infty, \infty}^\alpha$ were treated in [9] only in the autonomous case, while in [17] when they are compactly supported in space. As neither option fits our setting nicely (consider drifts of the form $b = \tilde{b} * \mu_t$) we extend the results therein to suit our analysis of DDSDEs.

3.1. Pathwise SDEs as nonlinear Young equations. Consider a standard SDE of the form

$$X_t = X_0 + \int_0^t b(s, X_s) ds + W_t, \quad \forall t \in [0, T], \quad (3.1)$$

where $b \in L_T^1 B_{\infty, \infty}^\alpha$ with $\alpha \in \mathbb{R}$ and W is an \mathbb{R}^d -valued fractional Brownian motion.

When $\alpha > 0$, the SDE has a classical meaning; it can be solved pathwise by standard ODE theory if b is regular enough, e.g. $\alpha > 1$. We will say that b is a *distributional drift* (sometimes *distributional field*) if instead $\alpha < 0$, in which case pointwise evaluation is not allowed, and we cannot give meaning to the integral appearing in (3.2) in the classical Lebesgue sense.

To deal with distributional drifts, we will employ the *nonlinear Young integral* framework, first developed in [9]; to present it, we first need the concept of *averaged field*.

Let us give an heuristic motivation before going into technical details. In the regular regime $\alpha > 0$, if X is a solution to (3.1), by the change of variables $\theta_t := X_t - W_t$ we find that θ solves

$$\theta_t = \theta_0 + \int_0^t b(s, \theta_s + W_s) ds. \quad (3.2)$$

Closely related to the above integral is the averaging of the field b along the curve W , namely the space-time function

$$T^W b(t, x) := \int_0^t b(s, x + W_s) ds \quad (3.3)$$

which we call an *averaged field*; we will write $T_{s,t}^W b(x) := T^W b(t, x) - T^W b(s, x)$.

As long as b is at least measurable and bounded, both integrals appearing in (3.2) and (3.3) are well defined. However, for distributional b , while equation (3.2) breaks down, the averaged field $T^W b$ is still meaningful in the distributional sense, see [17, Section 3.1]; moreover, depending on the properties of W , $T^W b$ might even be continuous or (higher order) differentiable in the spatial variable.

The fundamental intuition of [9] is that the regularity of $T^W b$ can be used to give meaning to (3.2), thus also to (3.1), by reformulating the SDE as a *nonlinear Young equation*.

As the next statement shows, given any space-time function $A : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ and path $\theta : [0, T] \rightarrow \mathbb{R}^d$ of suitable regularity, it's possible to give meaning to $\int_0^t \partial_t A(s, \theta_s) ds$ also when $\partial_t A$ is not well defined anymore.

Proposition 3.1. *Let $\gamma > 1/2$ and consider a function $A \in C_T^\gamma C_{loc}^1$ and a path $\theta \in C_T^\gamma$. Then for any interval $[s, t] \subset [0, T]$ and any sequence of partitions \mathcal{D}_n of $[s, t]$ with mesh converging to zero, the following limit exists and is independent of the chosen sequence:*

$$\int_s^t A(ds, \theta_s) := \lim_{n \rightarrow \infty} \sum_{[u, v] \in \mathcal{D}_n} A_{u, v}(\theta_u).$$

We will refer to it as a nonlinear Young integral. Furthermore:

- i. The integral is additive: $\int_s^t A(du, \theta_u) = \int_s^r A(du, \theta_u) + \int_r^t A(du, \theta_u)$ for any $r \in [s, t]$.
- ii. If $\partial_t A$ exists and is continuous, then $\int_0^t A(du, \theta_u) = \int_0^t \partial_u A(u, \theta_u) du$.
- iii. The map from $C_T^\gamma C_{loc}^1 \times C_T^\gamma$ to C_T^γ given by $(A, \theta) \mapsto \int_0^\cdot A(du, \theta_u)$ is linear in A and continuous in both variables. Namely, if $A^n \rightarrow A$ in $C_T^\gamma C_{loc}^1$ and $\theta^n \rightarrow \theta$ in C_T^γ , then $\int_0^\cdot A^n(ds, \theta_s^n) \rightarrow \int_0^\cdot A(ds, \theta_s)$ in C_T^γ .

Proof. The statement is a particular subcase of [15, Theorem 2.7]. \square

We provided the statement only for $A \in C_T^\gamma C_{loc}^1$ as this setting is sufficient for our purposes, but let us mention that the theory is more general and allows to consider $A \in C_T^\gamma C_{loc}^\nu$, $\theta \in C_T^\rho$ for $\gamma + \nu\rho > 1$. With the above result at hand, we can now define nonlinear Young equations.

Definition 3.2. Let $A \in C_T^\gamma C_{loc}^1$ with $\gamma > 1/2$, $\theta_0 \in \mathbb{R}^d$; we say that θ is a solution to the nonlinear Young equation associated to (θ_0, A) if $\theta \in C_T^\gamma$ and

$$\theta_t = \theta_0 + \int_0^t A(ds, \theta_s) \quad \forall t \in [0, T]. \quad (3.4)$$

For later use, we provide the following technical lemma; loosely speaking it shows that solutions to nonlinear Young equations have a closure property.

Lemma 3.3. *Let $\gamma > 1/2$, $A \in C_T^\gamma C_{loc}^1$ and $\{A^n\}_{n \in \mathbb{N}}$ be a sequence converging to A in $C_T^\gamma C_{loc}^1$; suppose that for each n there exists a solution θ^n associated to (θ_0, A^n) and that $\theta^n \rightarrow \theta$ in C_T^γ . Then θ solves the nonlinear Young equation associated to (θ_0, A) .*

Proof. This is a direct consequence of Point iii. of Proposition 3.1. By assumption

$$\theta_t^n = \theta_0 + \int_0^t A^n(ds, \theta_s^n) \quad \forall t \in [0, T], n \in \mathbb{N}$$

and we can pass to the limit on both sides thanks to the continuity of $(\theta, A) \mapsto \int_0^\cdot A(ds, \theta_s)$. \square

We are now ready to explain what it means for X to be a solution to (3.1) when b is distributional but $T^W b$ is regular enough: roughly speaking, we impose the condition $X = \theta + W$, where θ solves the nonlinear YDE associated to $A = T^W b$, which is the natural extension of (3.2). Although so far we have always dealt with a stochastic process W , this is a *pathwise notion* of solution, in the sense that for any fixed realization of $W(\omega)$ such that $T^{W(\omega)} b \in C_T^\gamma C_{loc}^1$ we have an analytically well-defined equation of the form (3.4). This is encoded in the next definition, inspired by [18, Section 4.3], which contains a more in-depth discussion of various related concepts.

Definition 3.4. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, (ξ, W) an $\mathbb{R}^d \times C_T$ -valued random variable defined on it and let b be a distributional field. We say that another C_T -valued random variable X on $(\Omega, \mathcal{F}, \mathbb{P})$ is a *pathwise solution* to the SDE (3.1) associated to (b, ξ, W) if there exists $\Omega' \subset \Omega$ with $\mathbb{P}(\Omega') = 1$ and a deterministic $\gamma > 1/2$ such that for all $\omega \in \Omega'$ the following hold:

- i. $T^{W(\omega)}b \in C_T^\gamma C_{loc}^1$;
- ii. $\theta(\omega) := X(\omega) - W(\omega) \in C_T^\gamma$;
- iii. $\theta(\omega)$ satisfies the nonlinear Young equation

$$\theta_t(\omega) = \xi(\omega) - W_0(\omega) + \int_0^t T^{W(\omega)}b(ds, \theta_s(\omega)) \quad \forall t \in [0, T].$$

The following definition relates standard probabilistic notions of weak and strong solutions and of uniqueness to the notion of pathwise existence given in Definition 3.4.

Definition 3.5. Let b be a distributional field, $\nu \in \mathcal{P}(\mathbb{R}^d \times C_T)$. A tuple $(\Omega, \mathcal{F}, \mathbb{P}; X, \xi, W)$ given by a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a $C_T \times \mathbb{R}^d \times C_T$ -valued random variable is a *weak solution* to the SDE (3.1) associated to (b, ν) if $\mathcal{L}_{\mathbb{P}}(\xi, W) = \nu$ and X is a pathwise solution associated to (b, ξ, W) in the sense of Definition 3.4. We say that X is a *strong solution* if it is adapted to the filtration $\mathcal{F}_t = \sigma\{\xi, W_s \mid s \leq t\}$. *Weak uniqueness* holds for the SDE associated to (b, ν) if any given weak solutions $(\Omega^i, \mathcal{F}^i, \mathbb{P}^i; X^i, \xi^i, W^i)$, $i = 1, 2$, associated to the same data (b, ν) , satisfy $\mathcal{L}_{\mathbb{P}^1}(X^1) = \mathcal{L}_{\mathbb{P}^2}(X^2)$. Similarly, *pathwise uniqueness* holds if any two given solutions (X^i, ξ, W) defined on the same probability space, w.r.t. the same (b, ξ, W) , satisfy $X^1 = X^2$ \mathbb{P} -a.s.

In line with the above definition, we will use the standard terminology that weak (resp. strong) existence holds for the SDE associated to (b, ν) to mean that we can construct a weak (resp. strong) solution $(\Omega, \mathcal{F}, \mathbb{P}; X, \xi, W)$. In particular, if strong existence holds, then $(\Omega, \mathcal{F}, \mathbb{P})$ can be chosen to be the canonical space, namely with $\Omega = \mathbb{R}^d \times C_T$, $\mathbb{P} = \nu$ and \mathcal{F} the completion of $\mathcal{B}(\mathbb{R}^d \times C_T)$ under ν .

Remark 3.6. If $b \in C_T C_{loc}^1$, then any classical solution to (3.1) is of the form $X = W + \theta$ for $\theta \in C_T^1$; moreover in this case $T^W b \in C_T^1 C_{loc}^1$ and $\partial_t T^W b(t, x) = b(t, x + W_t)$. It then follows from Point ii. of Proposition 3.1 that in this setting the concept of pathwise solution from Definition 3.4 is equivalent to the standard one. Moreover for $b \in C_T C_{loc}^1$ standard ODE theory guarantees pathwise uniqueness, uniqueness in law and strong existence of solutions for the SDE associated to (b, ν) for any choice of $\nu \in \mathcal{P}(\mathbb{R}^d \times C_T)$.

The next lemma provides a simple condition to establish uniqueness of solutions to (3.1).

Lemma 3.7. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, (X, ξ, W) be a triple defined on it such that X solves the SDE associated to (b, ξ, W) in the sense of Definition 3.4. If $T^{X(\omega)}b \in C_T^\gamma C_{loc}^1$ for \mathbb{P} -a.e. ω , then any other solution \tilde{X} defined on the same probability space and associated to (b, ξ, W) must coincide with it, in the sense that $X = \tilde{X}$ \mathbb{P} -a.s.*

Proof. The statement is a useful rewriting of [17, Remark 15]. □

Let us stress that, even when the assumptions of Lemma 3.7 are met, pathwise uniqueness doesn't immediately follow, unless one can additionally show that X is a strong solution.

3.2. Regularity of averaged fields and Girsanov transform for fBm. In Section 3.1 we have treated the SDE (3.1) in full generality, but in the remainder of Section 3 we will deal with a slightly more specific setting. We will always take W to be an \mathbb{R}^d -valued fBm of parameter $H \in (0, 1)$ and ξ to be random initial data independent of it; in particular $W_0 \equiv 0$ and $\nu = \mathcal{L}(\xi, W) = \mathcal{L}(\xi) \otimes \mathcal{L}(W) = \mu_0 \otimes \mu^H$ for some $\mu_0 \in \mathcal{P}(\mathbb{R}^d)$, where $\mu^H \in \mathcal{P}(C_T)$ denotes the law of fBm of parameter $H \in (0, 1)$. Therefore for fixed H we can regard the data of the problem to be the pair (μ_0, b) ; if the initial data $\xi = x_0 \in \mathbb{R}^d$ is deterministic, with a slight abuse we will write (x_0, b) in place of (δ_{x_0}, b) .

We begin by showing the \mathbb{P} -a.s. regularity of averaged fields $T^W b$ for W sampled as an fBm. We continue to make use of the intuitive notation

$$\int_0^t b(r, x + W_r) dr = T^W b,$$

despite the fact that in general these objects will not be defined as Lebesgue integrals; rather they are random variables defined on $(\Omega, \mathcal{F}, \mathbb{P})$ constructed as the unique limits of $\int_0^t b^n(r, x + W_r) dr$ for any sequence $b^n \rightarrow b$ in appropriate topologies. More precisely, for $b \in L_T^q B_{\infty, \infty}^\alpha$ satisfying the assumptions of Proposition 3.8 below, one can consider any sequence b^n of smooth bounded fields such that $b^n \rightarrow b$ in $L_T^q B_{\infty, \infty}^{\alpha-\varepsilon}$ for all $\varepsilon > 0$.

Proposition 3.8. *Let $b \in L_T^q B_{\infty, \infty}^\alpha$ with $\alpha < 0$, $q \in (2, \infty]$, W be a fBm of parameter $H \in (0, 1)$; suppose (α, q) satisfy*

$$\gamma := 1 - \frac{1}{q} + \alpha H > \frac{1}{2}. \quad (3.5)$$

Then for any $\tilde{\gamma} < \gamma$ there exists an increasing function K (depending on d, T and the above parameters) such that

$$\mathbb{E} \left[\exp \left(\frac{\eta}{\|b\|_{L_T^q B_{\infty, \infty}^\alpha}^2} \left[\int_0^t b(r, x + W_r) dr \right]_{\tilde{\gamma}}^2 \right) \right] \leq K(\eta) \quad \forall \eta > 0 \quad (3.6)$$

uniformly over $x \in \mathbb{R}^d$ and $b \in L_T^q B_{\infty, \infty}^\alpha$, $b \neq 0$.

Proof. As the proof follows quite closely the ones given in [17, Section 3.3], we only provide a sketch. Let b be smooth and compactly supported, otherwise one can argue by density; up to reasoning componentwise, scaling and shifting, we can assume $x = 0$, $b \in C_c^\infty(\mathbb{R}^d)$ and $\|b\|_{L_T^q B_{\infty, \infty}^\alpha} = 1$, so it will never appear in the computations in the sequel.

Set $W_{s,t}^{(2)} = \mathbb{E}[W_t | \mathcal{F}_s]$ for $\mathcal{F}_s = \sigma\{W_r : r \leq s\}$, then by [17, Lemma 5] there exist $c_H, \tilde{c}_H > 0$ such that

$$\begin{aligned} \int_s^t b(r, W_r) dr &= \int_s^t P_{\tilde{c}_H |r-s|^{2H}} b(r, W_{s,r}^{(2)}) dr \\ &\quad + c_H \int_s^t \int_u^t P_{\tilde{c}_H |r-u|^{2H}} \nabla b(r, W_{u,r}^{(2)}) |r-u|^{H-1/2} dr \cdot dB_u \end{aligned} \quad (3.7)$$

where P_t denotes the Gaussian heat kernel, and B_t is a standard Brownian motion in \mathbb{R}^d . In the following we will drop the constants c_H, \tilde{c}_H , as they don't play any significant role. Thus for any fixed $s < t$, it holds

$$\int_s^t b(r, W_r) dr = I_{s,t}^1 + I_{s,t}^2 = I_{s,t}^1 + \int_s^t J_{u,t} \cdot dB_u$$

where

$$I_{s,t}^1 := \int_s^t P_{|r-s|^{2H}} b(r, W_{s,r}^{(2)}) dr, \quad J_{u,t} := \int_u^t P_{|r-u|^{2H}} \nabla b(r, W_{u,r}^{(2)}) |r-u|^{H-1/2} dr.$$

Let us show how to obtain exponential estimates for I^2 , the ones for I^1 being similar. Going through analogous computations to [17, Theorem 4], invoking heat kernel type estimates, it holds

$$\begin{aligned} |J_{u,t}| &\lesssim \int_u^t \|P_{|r-u|^{2H}} \nabla b_r\|_{L_x^\infty} |r-u|^{H-1/2} dr \\ &\lesssim \int_u^t \|b_r\|_{B_{\infty,\infty}^\alpha} |r-u|^{H\alpha-1/2} dr \\ &\lesssim \|b\|_{L_T^q B_{\infty,\infty}^\alpha} |t-u|^{1/2-1/q+H\alpha} = |t-u|^{\gamma-1/2}. \end{aligned}$$

Applying Burkholder-Davis-Gundy inequality with optimal asymptotic behaviour for large p , we deduce that

$$\mathbb{E}[|I_{s,t}^2|^p]^{1/p} \lesssim \sqrt{p} \mathbb{E}\left[\left(\int_s^t |J_{u,t}|^2 du\right)^{p/2}\right]^{1/p} \lesssim \sqrt{p} |t-s|^\gamma$$

Putting everything together, there exists a constant $C > 0$ such that for any $\eta > 0$

$$\mathbb{E}\left[\exp\left(\eta \left|\frac{I_{s,t}^2}{|t-s|^\gamma}\right|^2\right)\right] = \sum_{n \in \mathbb{N}} \frac{\eta^n}{n!} \mathbb{E}\left[\left|\frac{I_{s,t}^2}{|t-s|^\gamma}\right|^{2n}\right] \leq \sum_{n \in \mathbb{N}} (C\eta)^n \frac{n^n}{n!}$$

and by Stirling's approximation the last series is convergent for any $\eta < (Ce)^{-1}$. Together with similar estimates for I^1 , we conclude that there exists $\tilde{\eta} > 0$ sufficiently small and $C > 0$ such that

$$\mathbb{E}\left[\exp\left(\eta \left|\frac{\int_s^t b(r, W_r) dr}{|t-s|^\gamma}\right|^2\right)\right] \leq C \quad \forall \eta \leq \tilde{\eta}, s < t.$$

The above estimate together with [17, Lemma 18] implies that for any $\tilde{\gamma} < \gamma$ there exist $\bar{\eta} > 0$ and $\kappa > 0$ such that

$$\mathbb{E}\left[\exp\left(\eta \left[\int_0^\cdot b(r, W_r) dr\right]_{\tilde{\gamma}}^2\right)\right] \leq \kappa \quad \forall \eta \leq \bar{\eta}. \quad (3.8)$$

It remains to show that we can improve the above inequality by allowing any value $\eta > 0$, so that we reach (3.6). To do so, we will resort to an interpolation trick, similar in style to techniques already applied in [17, Theorem 15], [9, Corollary 4.6].

First, observe that if α, q, H satisfy (3.5) and we fix $\tilde{\gamma} < \gamma$, then we can find ε sufficiently small so that $\gamma^\varepsilon = 1 - 1/q - (\alpha - \varepsilon)H > 1/2$ and $\tilde{\gamma} < \gamma^\varepsilon$; then by estimate (3.8) (for $\alpha - \varepsilon$ in place of α) and linearity, there exist $\bar{\eta} > 0$ and $\kappa > 0$ such that

$$\mathbb{E}\left[\exp\left(\frac{\bar{\eta}}{\|\tilde{b}\|_{L_T^q B_{\infty,\infty}^{\alpha-\varepsilon}}^2} \left[\int_0^\cdot \tilde{b}(r, W_r) dr\right]_{\tilde{\gamma}}^2\right)\right] \leq \kappa \quad \forall \tilde{b} \in L_T^q B_{\infty,\infty}^{\alpha-\varepsilon}, \tilde{b} \neq 0. \quad (3.9)$$

As before we can assume $\|b\|_{L_T^q B_{\infty,\infty}^\alpha} = 1$ and we fix $\varepsilon > 0$ as above. Then for any $N \in \mathbb{N}$ we can decompose b as

$$b_t = b_t^{1,N} + b_t^{2,N}, \quad b_t^{1,N} = \sum_{j \leq N} \Delta_j b_t, \quad b_t^{2,N} = \sum_{j > N} \Delta_j b_t$$

where Δ_j denote Littlewood-Paley blocks. There exists $C > 0$ such that

$$\|b^{1,N}\|_{L_T^q C_x^0} \leq C 2^{-N\alpha}, \quad \|b^{2,N}\|_{L_T^q B_{\infty,\infty}^{\alpha-\varepsilon}} \leq C 2^{-N\varepsilon}.$$

Now for a given $\eta > 0$, choose $N = N(\eta) \in \mathbb{N}$ such that $\eta \leq C^{-2} 2^{2N\varepsilon-1} \bar{\eta}$ and decompose b as above; w.l.o.g. we may assume that $b^{2,N} \neq 0$, otherwise the stated estimate is trivial. Clearly under (3.5) it holds that $\tilde{\gamma} \leq 1 - 1/q$, therefore setting $\beta = 1 - 1/q - \tilde{\gamma}$ we have

$$\left[\int_0^\cdot b^{1,N}(r, W_r) dr \right]_{\tilde{\gamma}} \leq (1+T)^\beta \|b^{1,N}\|_{L_T^q C_x^0} \leq C(1+T)^\beta 2^{-N\alpha} =: C_{N(\eta)},$$

where the estimate is deterministic; combining it with (3.9) applied to $\tilde{b} = b^{2,N}$, we get

$$\begin{aligned} \mathbb{E} \left[\exp \left(\eta \left[\int_0^\cdot b(r, W_r) dr \right]_{\tilde{\gamma}}^2 \right) \right] &\leq \mathbb{E} \left[\exp \left(2\eta \left[\int_0^\cdot b^{1,N}(r, W_r) dr \right]_{\tilde{\gamma}}^2 + 2\eta \left[\int_0^\cdot b^{2,N}(r, W_r) dr \right]_{\tilde{\gamma}}^2 \right) \right] \\ &\leq \exp(2\eta C_{N(\eta)}^2) \mathbb{E} \left[\exp \left(\frac{\bar{\eta}}{\|b^{2,N}\|_{L_T^q B_{\infty,\infty}^{\alpha-\varepsilon}}^2} \left[\int_0^\cdot b^{2,N}(r, W_r) dr \right]_{\tilde{\gamma}}^2 \right) \right] \\ &\leq \exp(2\eta C_{N(\eta)}^2) \kappa \end{aligned}$$

where the estimate now holds for all $\eta \geq 0$. \square

Corollary 3.9. *Let $b \in L_T^q B_{\infty,\infty}^\alpha$ with $\alpha < 0$, $q \in (2, \infty]$, W be a fBm of parameter $H \in (0, 1)$ and let $\rho \in (0, 1]$; suppose (α, ρ, q) satisfy*

$$\gamma := 1 - \frac{1}{q} + (\alpha - \rho)H > \frac{1}{2}. \quad (3.10)$$

Then for any $\tilde{\gamma} < \gamma$ there exists an increasing function K (depending on d, T and the previous parameters) such that

$$\mathbb{E} \left[\exp \left(\eta \left| \frac{\left[\int_0^\cdot b(r, x + W_r) dr - \int_0^\cdot b(r, y + W_r) dr \right]_{\tilde{\gamma}}^2}{\|b\|_{L_T^q B_{\infty,\infty}^\alpha} |x - y|^\rho} \right| \right) \right] \leq K(\eta) \quad \forall \eta \geq 0 \quad (3.11)$$

uniformly over $x \neq y \in \mathbb{R}^d$ and $b \in L_T^q B_{\infty,\infty}^\alpha$, $b \neq 0$; as a consequence, for any $\varepsilon > 0$, \mathbb{P} -a.s. $T^W b \in C_T^{\gamma-\varepsilon} C_{loc}^{\rho-\varepsilon}$. Suppose now $b \in L_T^q B_{\infty,\infty}^\alpha$ with $\alpha < 1$, $q \in (2, \infty]$ satisfying

$$\alpha - \frac{1}{Hq} > 1 - \frac{1}{2H}, \quad (3.12)$$

then the following hold:

- i. There exists $\tilde{\gamma} > 1/2$ such that \mathbb{P} -a.s. $T^W b \in C_T^{\tilde{\gamma}} C_{loc}^1$.*
- ii. There exists $\tilde{\gamma} > 1/2$ such that for any $b^1, b^2 \in L_T^q B_{\infty,\infty}^\alpha$ and any $n \in \mathbb{N}$*

$$\mathbb{E} \left[\left[\int_0^\cdot b^1(r, W_r) dr - \int_0^\cdot b^2(r, W_r) dr \right]_{\tilde{\gamma}; [0,\tau]}^n \right]^{1/n} \lesssim_n \left(\int_0^\tau \|b_r^1 - b_r^2\|_{B_{\infty,\infty}^{\alpha-1}}^q dr \right)^{1/q} \quad \forall \tau \in [0, T].$$

- iii. If $H < 1/2$, $\alpha < 0$, there exists $\tilde{\gamma} > H + 1/2$ and an increasing function K such that*

$$\mathbb{E} \left[\exp \left(\frac{\eta}{\|b\|_{L_T^q B_{\infty,\infty}^\alpha}^2} \left[\int_0^\cdot b(r, W_r) dr \right]_{\tilde{\gamma}}^2 \right) \right] \leq K(\eta) \quad \forall \eta \geq 0, b \in L_T^q B_{\infty,\infty}^\alpha, b \neq 0.$$

Proof. Given b as above, $x \neq y$ fixed, define $\tilde{b}(t, \cdot) = |x - y|^{-\rho} [b(t, x + \cdot) - b(t, y + \cdot)]$; by properties of Besov spaces

$$\|\tilde{b}\|_{L_T^q B_{\infty, \infty}^{\alpha - \rho}} \lesssim \|b\|_{L_T^q B_{\infty, \infty}^{\alpha}}.$$

Inequality (3.11) follows from (3.6) applied to \tilde{b} , since by assumption (3.10) $\tilde{\alpha} = \alpha - \rho$ satisfies (3.5); $T^W b$ belonging to $C_T^{\gamma - \varepsilon} C_{loc}^{\rho - \varepsilon}$ is a consequence of Garsia-Rodemich-Rumsay lemma.

We now assume (3.12) holds and prove points i.-iii.

If $b \in L_T^q B_{\infty, \infty}^{\alpha}$, then $D_x b \in L_T^q B_{\infty, \infty}^{\alpha - 1}$ with $\tilde{\alpha} = \alpha - 1$ satisfying (3.5), so we can find $\rho > 0$ small enough such that $(\tilde{\alpha}, \rho)$ satisfy (3.10) as well. It follows that $T^W D_x b = D_x T^W b \in C_T^{\gamma - \varepsilon} C_{loc}^0$, namely $T^W b \in C_T^{\gamma - \varepsilon} C_{loc}^1$, for any $\varepsilon > 0$, showing i..

For $\tau = T$, the statement in part ii. is again a consequence of (3.6) (for $x = 0$ and $\tilde{\alpha} = \alpha - 1$) and the linearity of $b \mapsto T^W b$. For general $\tau \in [0, T]$, define $\tilde{b}_t^i = b_t^i \mathbb{1}_{[0, \tau]}(t)$ and observe that

$$\left[\int_0^\cdot (b^1 - b^2)(r, W_r) dr \right]_{\gamma; [0, \tau]} = \left[\int_0^\cdot (\tilde{b}^1 - \tilde{b}^2)(r, W_r) dr \right]_{\gamma; [0, T]}$$

the estimate for general τ thus follows applying the one for $\tau = T$ to \tilde{b}^i .

Finally, in order to prove iii. it is enough to show that

$$\gamma = 1 - \frac{1}{q} + \alpha H > H + 1/2,$$

as in that case we can find $\tilde{\gamma} \in (H + 1/2, \gamma)$ such that (3.6) holds. But the above condition on γ is exactly (3.12). \square

In order to apply Lemma 3.7, we need some information on the pathwise properties of weak solutions X . From this perspective, techniques based on Girsanov theorem are very natural, as they suggest that $T^X b$ may have the same regularity as $T^W b$. As already mentioned, Girsanov transform holds for fBm, see [37]; sufficient conditions in order to apply it in our context (in particular to check that Novikov condition is satisfied) can be found in [17, Section 4.2.2], to which we also refer for more details on the explicit formula for $d\mathbb{P}/d\mathbb{Q}$.

Proposition 3.10. *Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, \mathbb{P})$ be a filtered probability space, W be an \mathcal{F}_t -fBm of parameter $H \in (0, 1)$ and h be an \mathcal{F}_t -adapted process with trajectories in C_T^γ , $\gamma > H + 1/2$, such that $h_0 = 0$ and*

$$\mathbb{E}_{\mathbb{P}}[\exp(\eta[h]_\gamma^2)] \leq K(\eta) < \infty \quad \forall \eta \in \mathbb{R}.$$

Then there exists another probability measure \mathbb{Q} , given by Girsanov theorem, such that $h + W$ is distributed as an \mathcal{F}_t -fBm under \mathbb{Q} . Moreover \mathbb{P} and \mathbb{Q} are equivalent and it holds

$$\mathbb{E}_{\mathbb{Q}} \left[\left(\frac{d\mathbb{P}}{d\mathbb{Q}} \right)^n + \left(\frac{d\mathbb{Q}}{d\mathbb{P}} \right)^n \right] < \infty \quad \forall n \in \mathbb{N} \quad (3.13)$$

where the above estimate only depends on the function K .

Proof. Follows almost exactly as the proof of [17, Theorem 14]. \square

Remark 3.11. For $H \leq 1/2$ and $b \in L_T^q B_{\infty, \infty}^{\alpha}$ with (α, q, H) satisfying (3.12), it follows from Corollary 3.9 and Proposition 3.10 that we can construct a weak solution $(\Omega, \mathcal{F}, \mathbb{P}; X, W)$ to the SDE associated to (x_0, b) , with the property that there exists a measure \mathbb{Q} equivalent to \mathbb{P} such that $\mathcal{L}_{\mathbb{Q}}(X) = \mathcal{L}_{\mathbb{P}}(x_0 + W)$; moreover all the moments of $d\mathbb{P}/d\mathbb{Q}$ and $d\mathbb{Q}/d\mathbb{P}$ can be controlled

in a way that depends on $\|b\|_{L_T^q B_{\infty,\infty}^\alpha}$ but not on the specific (x_0, b) . In particular, the estimates can be performed uniformly over $x_0 \in \mathbb{R}^d$ and $\|b\|_{L_T^q B_{\infty,\infty}^\alpha} \leq M$ for a fixed parameter $M > 0$.

Similarly, in the case $H > 1/2$, given $b \in E = C_T^{\alpha H} C_x^0 \cap C_T^0 C_x^\alpha$ for some $\alpha > 1 - 1/(2H)$, using the regularity of fBm trajectories it's easy to check that the map $t \mapsto b(t, x_0 + W_t)$ belongs \mathbb{P} -a.s. to $C_T^{\alpha H - \varepsilon}$ for any $\varepsilon > 0$. Furthermore, reasoning as in the proof of [17, Theorem 15], it can be shown that there exists $\gamma > H + 1/2$ and an increasing function K such that

$$\mathbb{E} \left[\exp \left(\eta \left\| \int_0^\cdot b(r, x_0 + W_r) dr \right\|_\gamma^2 \right) \right] \leq K(\eta) < \infty \quad \forall \eta \geq 0.$$

Therefore also in this case we can apply Proposition 3.10 to construct weak solutions to the SDE. Moreover the function K only depends on $\|b\|_E$, therefore as before all estimates are uniform over $x_0 \in \mathbb{R}^d$ and $b \in E$ with $\|b\|_E \leq M$, M fixed parameter.

If both cases, if in addition b is smooth, then the weak solution constructed in this way necessarily coincides with the unique strong one; thus the above reasoning also provide uniform estimates for the solutions associated to smooth drifts.

3.3. Stability estimates for SDEs. In light of the above results, in the remainder of Section 3 we will always impose the following assumption on the drift b .

Assumption 3.12. *Given $H \in (0, 1)$, b satisfies one of the following:*

- *If $H > 1/2$, then $b \in C_T^{\alpha H} C_x^0 \cap b \in C_T^0 C_x^\alpha$ for some*

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

equivalently, there exists a constant $C > 0$ s.t., for all $(s, t, x, y) \in [0, T]^2 \times \mathbb{R}^{2d}$, such that

$$|b(t, x)| \leq C, \quad |b(t, x) - b(s, y)| \leq C(|t - s|^{\alpha H} + |x - y|^\alpha).$$

- *If $H \leq 1/2$, then $b \in L_T^q B_{\infty,\infty}^\alpha$ for some (α, q) satisfying (3.12).*

In both cases we will use the notation $\|b\|_E$ for $E = C_T^{\alpha H} C_x^0 \cap C_T^0 C_x^\alpha$ when $H > 1/2$, respectively $E = L_T^q B_{\infty,\infty}^\alpha$ when $H \leq 1/2$.

We are now ready to present the main result of this section. Its novelty, compared to previous results like those in [9, 17], lies in the comparison of two solutions driven by different drifts b^i , which gives rise to the term $\|b^1 - b^2\|_{L_T^q B_{\infty,\infty}^{\alpha-1}}$ (and not $\|b^1 - b^2\|_{L_T^q B_{\infty,\infty}^\alpha}$!) appearing in (3.15). On a more technical level, we drop slightly restrictive assumptions from previous works, like working with autonomous drifts as in [9], or compactly supported in space ones like in [17]; let us also point out that we fill a gap in the statements of Theorem 4.23 and Corollary 4.24 from [17], which do not cover the case of $b \in L_T^q B_{\infty,\infty}^\alpha$ with $q < \infty$ and $\alpha > 0$.

Theorem 3.13. *Let W be an fBm of parameter $H \in (0, 1)$ and let b satisfy Assumption 3.12. Then for any $x_0 \in \mathbb{R}^d$ strong existence, pathwise uniqueness and uniqueness in law hold for the SDE*

$$X_t = x_0 + \int_0^t b(r, X_r) dr + W_t \tag{3.14}$$

in the sense of Definition 3.5. Given $x_0^i \in \mathbb{R}^d$ and b^i satisfying Assumption 3.12, $i = 1, 2$, denote by X^i the solutions associated to (x_0^i, b^i) and let $M > 0$ be a constant such that $\|b^i\|_E \leq M$ for $i = 1, 2$. Let (α, \tilde{q}) be another pair satisfying (3.12) with the same α as in Assumption 3.12 and

$\tilde{q} \leq q$. Then there exists $\gamma > 1/2$ with the following property: for any $p \in [1, \infty)$ there exists a constant $C > 0$ (depending on $\gamma, p, M, T, d, \tilde{q}$ and the parameters appearing in Assumption 3.12) such that

$$\mathbb{E} \left[\|X^1 - X^2\|_{\gamma; [0, \tau]}^p \right]^{1/p} \leq C \left(|x_0^1 - x_0^2| + \|b^1 - b^2\|_{L^{\tilde{q}}(0, \tau; B_{\infty, \infty}^{\alpha-1})} \right) \quad \forall \tau \in [0, T]. \quad (3.15)$$

Proof. We will only treat the case $H \leq 1/2$, the other one being almost identical.

Let us first assume b^i to be smooth functions and show that (3.15) holds; in this case by Remark 3.6 strong existence and uniqueness hold automatically. Moreover by Remark 3.11, there exist probability measures \mathbb{Q}^i equivalent to \mathbb{P} such that $\mathcal{L}_{\mathbb{Q}^i}(X^i) = \mathcal{L}_{\mathbb{P}}(x_0^i + W)$, with moment estimates depending on M but not on (x_0^i, b^i) ; the solutions decompose as $X^i = x_0^i + h^i + W^i$ with $h_0^i = 0$, $h^i \in C_T^{\tilde{\gamma}}$ with $\tilde{\gamma} > H + 1/2$ and such that

$$\mathbb{E} \left[\exp(\eta \|h^i\|_{\gamma}^2) \right] \leq K(\eta) < \infty \quad \forall \eta \geq 0$$

where again K depends on M but not on the specific (x_0^i, b^i) .

For any $\lambda \in [0, 1]$, let us define $x_0^\lambda := x_0^2 + \lambda(x_0^1 - x_0^2)$, $h^\lambda := h^2 + \lambda(h^1 - h^2)$, so that $X^2 + \lambda(X^1 - X^2) = x_0^\lambda + h^\lambda + W$. By Taylor expansion and elementary addition and subtraction, the difference $Y = X^1 - X^2$ satisfies

$$Y_t = x_0^1 - x_0^2 + \int_0^t \left(\int_0^1 Db^1(r, x^\lambda + h_r^\lambda + W_r) d\lambda \right) \cdot Y_r dr + \int_0^t (b^1 - b^2)(r, X_r^2) dr;$$

let us define

$$A_t := \int_0^t \int_0^1 Db^1(r, x_0^\lambda + h_r^\lambda + W_r) d\lambda dr, \quad \psi_t := \int_0^t (b^1 - b^2)(r, X_r^2) dr.$$

In order to get estimates for Y , it turns out to be useful to reinterpret the above equation as a linear Young differential equation of the form

$$Y_t = Y_0 + \int_0^t A_{ds} Y_s + \psi_t. \quad (3.16)$$

Indeed for any $\gamma > 1/2$, we can apply [15, estimate (3.16), Theorem 3.9] to obtain the existence of a constant $C = C(\gamma)$ such that for any $\tau \leq T$ it holds

$$\begin{aligned} \|Y\|_{\gamma; [0, \tau]} &\leq C \exp(C\tau(1 + \|A\|_{\gamma; [0, \tau]}^2)) (|Y_0| + (1 + \tau^\gamma) \|\psi\|_{\gamma; [0, \tau]}) \\ &\lesssim_T \exp(CT \|A\|_{\gamma}^2) (|x_0^1 - x_0^2| + \|\psi\|_{\gamma; [0, \tau]}) \end{aligned}$$

and so our task reduces to finding estimates for quantities of the form

$$\mathbb{E}_{\mathbb{P}} \left[\exp(\eta \|A\|_{\gamma}^2) \right], \quad \mathbb{E}_{\mathbb{P}} \left[\|\psi\|_{\gamma; [0, \tau]}^p \right].$$

We start by estimating ψ , which is the simplest term. Recalling that $\mathcal{L}_{\mathbb{Q}^2}(X^2) = \mathcal{L}_{\mathbb{P}}(x_0^2 + W)$, by Point ii. of Corollary 3.9 and Cauchy inequality we can find $\gamma > 1/2$ such that, for any $p \geq 1$,

$$\begin{aligned} \mathbb{E}_{\mathbb{P}}[\|\psi\|_{\gamma;[0,\tau]}^p] &= \mathbb{E}_{\mathbb{P}}\left[\left[\int_0^\cdot (b^1 - b^2)(r, X_r^2) dr\right]_{\gamma;[0,\tau]}^p\right] = \mathbb{E}_{\mathbb{Q}^2}\left[\left[\int_0^\cdot (b^1 - b^2)(r, X_r^2) dr\right]_{\gamma;[0,\tau]}^p \frac{d\mathbb{P}}{d\mathbb{Q}^2}\right] \\ &\leq \mathbb{E}_{\mathbb{Q}^2}\left[\left(\frac{d\mathbb{P}}{d\mathbb{Q}^2}\right)^2\right]^{1/2} \mathbb{E}_{\mathbb{Q}^2}\left[\left[\int_0^\cdot (b^1 - b^2)(r, X_r^2) dr\right]_{\gamma;[0,\tau]}^{2p}\right]^{1/2} \\ &\lesssim_M \mathbb{E}_{\mathbb{P}}\left[\left[\int_0^\cdot (b^1 - b^2)(r, x_0^2 + W_r) dr\right]_{\gamma;[0,\tau]}^{2p}\right]^{1/2} \lesssim_p \|b^1 - b^2\|_{L^{\tilde{q}}(0,\tau;B_{\infty,\infty}^{\alpha-1})}. \end{aligned}$$

In order to get estimates for A , observe first of all that by convexity of $z \mapsto \exp(\eta z^2)$, it holds

$$\mathbb{E}[\exp(\eta \|h^\lambda\|_\gamma^2)] \leq \lambda \mathbb{E}[\exp(\eta \|h^1\|_\gamma^2)] + (1 - \lambda) \mathbb{E}[\exp(\eta \|h^2\|_\gamma^2)] \leq K(\eta) < \infty$$

where the estimate is uniform in λ and η ; therefore by Proposition 3.10, for any λ there exists a probability \mathbb{Q}^λ equivalent to \mathbb{P} such that $\mathcal{L}_{\mathbb{Q}^\lambda}(h^\lambda + W) = \mathcal{L}_{\mathbb{P}}(W)$; moreover estimates of the form (3.13) only depend on K and thus on M , but not (x_0^i, b^i) . Therefore by Jensen's inequality and Proposition 3.8, we can find $\gamma > 1/2$ such that, for any $\eta \geq 0$, it holds that

$$\begin{aligned} \mathbb{E}_{\mathbb{P}}[\exp(\eta \|A\|_\gamma^2)] &= \mathbb{E}_{\mathbb{P}}\left[\exp\left(\eta \left[\int_0^1 \int_0^\cdot Db^1(r, x_0^\lambda + h_r^\lambda + W_r) dr\right]_\gamma^2\right)\right] \\ &\leq \int_0^1 \mathbb{E}_{\mathbb{P}}\left[\exp\left(\eta \left[\int_0^\cdot Db^1(r, x_0^\lambda + h_r^\lambda + W_r)\right]_\gamma^2\right)\right] \\ &= \int_0^1 \mathbb{E}_{\mathbb{Q}^\lambda}\left[\exp\left(\eta \left[\int_0^\cdot Db^1(r, x_0^\lambda + h_r^\lambda + W_r)\right]_\gamma\right) \frac{d\mathbb{P}}{d\mathbb{Q}^\lambda}\right] \\ &\leq \int_0^1 \mathbb{E}_{\mathbb{P}}\left[\exp\left(2\eta \left[\int_0^\cdot Db^1(r, x_0^\lambda + W_r) dr\right]_\gamma^2\right)\right]^{1/2} \mathbb{E}_{\mathbb{Q}^\lambda}\left[\left(\frac{d\mathbb{P}}{d\mathbb{Q}^\lambda}\right)^2\right]^{1/2} d\lambda \\ &\lesssim_M \mathbb{E}_{\mathbb{P}}\left[\exp\left(2\eta \left[\int_0^\cdot Db^1(r, W_r) dr\right]_\gamma^2\right)\right]^{1/2} \leq K(2\eta)^{1/2}. \end{aligned}$$

Putting everything together, we have obtained

$$\begin{aligned} \mathbb{E}_{\mathbb{P}}[\|Y\|_{\gamma;[0,\tau]}^p] &\lesssim \mathbb{E}_{\mathbb{P}}[\exp(pCT\|A\|_\gamma^2)(|Y_0|^p + \|\psi\|_{\gamma;[0,\tau]}^p)] \\ &\lesssim \mathbb{E}_{\mathbb{P}}[\exp(2pCT\|A\|_\gamma^2)]^{1/2} \left(|Y_0|^p + \mathbb{E}_{\mathbb{P}}[\|\psi\|_{\gamma;[0,\tau]}^{2p}]\right)^{1/2} \\ &\lesssim_{p,T,M} |x_0^1 - x_0^2|^p + \|b^1 - b^2\|_{L^{\tilde{q}}(0,\tau;B_{\infty,\infty}^{\alpha})}^p \end{aligned}$$

which proves (3.15) for smooth b^i .

Assume now we are given $x_0 \in \mathbb{R}^d$ and b satisfying Assumption 3.12; we can find $\tilde{q} \leq q$, $\tilde{q} < \infty$ such that (α, \tilde{q}) satisfy (3.12) and a sequence $\{b^n\}_n$ be smooth drifts s.t. $\|b^n\|_E \leq \|b\|_E$ for all $n \geq 1$ and $b^n \rightarrow b$ in $L_T^{\tilde{q}} B_{\infty,\infty}^{\alpha-\varepsilon}$ for any $\varepsilon > 0$ (for instance set $b^n = b * \psi^n$ with $\{\psi^n\}_{n \geq 1}$ a standard family of mollifiers). Let X^n be the unique solutions to (3.14) associated to (x_0, b^n) , then by (3.15) it holds

$$\mathbb{E}[\|X^n - X^m\|_\gamma^p]^{1/p} \lesssim \|b^n - b^m\|_{L_T^{\tilde{q}} B_{\infty,\infty}^{\alpha-1}}$$

showing that the random variables $\theta^n = X^n - W$ are a Cauchy sequence in $L^p_\Omega C_T^\gamma$. Therefore they converge to a unique limit θ , which is adapted to the filtration $\mathcal{F}_t = \sigma\{W_s : s \leq t\}$ since θ^n are so. Similarly the X^n converge to $X = \theta + W$ which is adapted.

The estimates from Corollary 3.9, the linearity of $b \mapsto T^w b$ and the property $b^n \rightarrow b$ in $L^{\tilde{q}}_T B_{p,p}^{\alpha-\varepsilon}$ together imply that \mathbb{P} -a.s. $T^W b^n \rightarrow T^W b$ in $C_T^\gamma C_{loc}^1$. Since we have \mathbb{P} -a.s. $\theta^n \rightarrow \theta$ in C_T^γ as well, we can invoke the closure property of nonlinear Young equations (Lemma 3.3) to deduce that $X = \theta + W$ is a pathwise solution to (3.14) in the sense of Definition 3.4.

Furthermore, by Fatou's lemma

$$\mathbb{E}_\mathbb{P}[\exp(\eta\|\theta\|_{\tilde{\gamma}}^2)] \leq \liminf_{n \rightarrow \infty} \mathbb{E}_\mathbb{P}[\exp(\eta\|\theta^n\|_{\tilde{\gamma}}^2)] \leq K(\eta) < \infty \quad \forall \eta \geq 0;$$

it follows that Girsanov can be applied to $X = \theta + W = x_0 + h + W$ to deduce that X is distributed as $x_0 + W$ under another probability measure equivalent to \mathbb{P} . In particular, \mathbb{P} -a.s. it must hold $T^X b \in C_T^\gamma C_{loc}^1$. To summarise, X is a strong solution (so that a copy of it can be constructed on *any* probability space supporting the measure μ^H) such that $T^X b \in C_T^\gamma C_{loc}^1$, which implies by Lemma 3.7 that pathwise uniqueness must hold. This also implies that the law of any solution coincides with the one constructed by Girsanov theorem, from which uniqueness in law follows.

The extension of inequality (3.14) to any pair of solutions X^i associated to distributional drifts b^i is now a direct consequence of the approximation argument. \square

Remark 3.14. At the price of making the statement of Theorem 3.13 slightly more technical, we have allowed the presence of the additional parameter $\tilde{q} \leq q$ to handle $q = \infty$. Indeed finding approximation sequences in $L_T^\infty B_{p,p}^{\alpha-1}$ can be a hard task since this is not a separable space; the use of $L_T^{\tilde{q}} B_{p,p}^{\alpha-1}$ with $\tilde{q} < \infty$ will also be useful later in the proofs in Section 4.3.

Remark 3.15. Theorem 3.13 gives us the information that, for drifts b satisfying Assumption 3.12, the nonlinear Young interpretation of the SDE is the only physical one. Namely, any other solution concept sharing the fundamental property of being the limit of solutions associated to smooth drifts $b^n \rightarrow b$ will coincide with ours. The statement of Theorem 3.13 can be further strengthened to establish *path-by-path uniqueness*, see [9], however we will not need this for our purposes.

Remark 3.16. Although we have proved the stability estimate (3.15) in order to apply to DDS-DEs, it is of interest on its own. Indeed it can be applied to construct the stochastic flow associated to SDE (3.14), or to develop numerical schemes for distributional drifts b by first approximating them by smoother b^n . We leave both applications for future research.

The next lemmas extend the previous results to the case of random initial data.

Corollary 3.17. *Given $H \in (0, 1)$ and b satisfying Assumption 3.12, strong existence, uniqueness in law and pathwise uniqueness also hold for random initial data $X_0 = \xi$ independent of W . Assume b^i are drifts satisfying the assumptions of Theorem 3.13 and $(\xi^1, \xi^2) \in L^p(\Omega; \mathbb{R}^{2d})$ is independent of W , then the solutions X^i associated to (ξ^i, b^i) satisfy*

$$\mathbb{E}_\mathbb{P}[\|X^1 - X^2\|_{\gamma; [0, \tau]}^p]^{1/p} \leq C \left(\mathbb{E}_\mathbb{P}[|\xi^1 - \xi^2|^p]^{1/p} + \|b^1 - b^2\|_{L^{\tilde{q}}(0, \tau; B_{\infty, \infty}^{\alpha-1})} \right) \quad \forall \tau \in [0, T]. \quad (3.17)$$

where the constant C and the parameters $\gamma, \alpha \tilde{q}$ are the same as in (3.15). Moreover, denoting by $\mu_t^i = \mathcal{L}(X_t^i)$ the laws of the unique solutions X^i associated to (μ_0^i, b^i) with $\mathcal{L}(\xi^i, W) = \mu_0^i \otimes \mathcal{L}(W)$,

it holds

$$\sup_{t \in [0, \tau]} d_p(\mu_t^1, \mu_t^2) \leq C \left(d_p(\mu_0^1, \mu_0^2) + \|b^1 - b^2\|_{L^{\tilde{q}}(0, \tau; B_{\infty, \infty}^{\alpha-1})} \right) \quad \forall \tau \in [0, T]. \quad (3.18)$$

Proof. Strong existence and pathwise uniqueness for random initial data follows from that for deterministic ones by classical arguments. Given a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with (W, ξ^1, ξ^2) defined on it and drifts (b^1, b^2) , we can condition on the variables (ξ^1, ξ^2) independent of W and apply estimate (3.15) to deduce that

$$\mathbb{E}_{\mathbb{P}} [\|X^1 - X^2\|_{\gamma; [0, \tau]}^p | \xi^1, \xi^2]^{1/p} \leq C \left(|\xi^1 - \xi^2| + \|b^1 - b^2\|_{L^{\tilde{q}}(0, \tau; B_{\infty, \infty}^{\alpha-1})} \right);$$

inequality (3.17) follows taking the L_{Ω}^p -norm on both sides, using the tower property of conditional expectation.

Now assume we are given a pair $(\mu_0^1, \mu_0^2) \in \mathcal{P}(\mathbb{R}^d) \times \mathcal{P}(\mathbb{R}^d)$ and let $m \in \Pi(\mu_0^1, \mu_0^2)$ be an optimal coupling for them. On the canonical space $\Omega = \mathbb{R}^{2d} \times C_T$, endowed with $\mathbb{P} = m \otimes \mu^H$, we can construct random variables (ξ^1, ξ^2, W) and solutions X^i associated to (ξ^i, b^i) , in such a way that $\mathcal{L}_{\mathbb{P}}(\xi^1, \xi^2) = m$, $\mathbb{E}[|\xi^1 - \xi^2|^p]^{1/p} = d_p(\mu_0^1, \mu_0^2)$. But then by the definition of d_p it must hold $d_p(\mu_t^1, \mu_t^2) \leq \|X_t^1 - X_t^2\|_{L_{\Omega}^p}$ and so estimate (3.18) follows from (3.17) applied in this setting. \square

Corollary 3.18. *Let $H \in (0, 1)$, b satisfying Assumption 3.12, ξ random initial data independent of W and X be the solution associated to (ξ, b) . Then there exists another probability measure \mathbb{Q} equivalent to \mathbb{P} such that $\mathcal{L}_{\mathbb{Q}}(X) = \mathcal{L}_{\mathbb{P}}(\xi + W)$; moreover*

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

Proof. It suffices to work on the canonical space $(\Omega, \mathcal{F}, \mathbb{P})$ with $\Omega = \mathbb{R}^d \times C_T \ni (x, \omega)$, $\mathbb{P} = \mu_0 \otimes \mu^H$ where $\mu_0 := \mathcal{L}(\xi)$. For any $x \in \mathbb{R}^d$, denote by $\omega \mapsto X^x(\omega)$ the unique strong solution associated to (x, b) , so that $(x, \omega) \mapsto X^x(\omega)$ gives the solution to the SDE with initial distribution μ_0 . Recall from Proposition 3.10 that for any $x \in \mathbb{R}^d$, there exists a probability measure on C_T denoted by \mathbb{Q}^x , equivalent to μ^H , such that $\mathcal{L}_{\mathbb{Q}^x}(X^x) = \mathcal{L}_{\mu^H}(x + W)$; therefore for any measurable $F : \Omega \rightarrow \mathbb{R}$

$$\begin{aligned} \mathbb{E}_{\mathbb{P}}[F(\xi + W)] &= \int_{\mathbb{R}^d} \int_{C_T} F(x + \omega) \mu^H(d\omega) \mu_0(dx) \\ &= \int_{\mathbb{R}^d} \int_{C_T} F(X^x(\omega)) \mathbb{Q}^x(d\omega) \mu_0(dx). \end{aligned}$$

Thus if we define a probability measure \mathbb{Q} on $\Omega = \mathbb{R}^d \times C_T$ by

$$\mathbb{Q}(E_1 \times E_2) = \int_{E_1} \mathbb{Q}^x(E_2) \mu_0(dx) \quad \forall E_1 \in \mathcal{B}(\mathbb{R}^d), E_2 \in \mathcal{B}(C_T),$$

it must hold that $\mathcal{L}_{\mathbb{Q}}(X) = \mathcal{L}_{\mathbb{P}}(\xi + W)$; since $\mu^H \ll \mathbb{Q}^x$ for every x , $\mathbb{P} = \mu^H \otimes \mu^{\xi} \ll \mathbb{Q}$ with Radon-Nikodym derivative given by

$$\frac{d\mathbb{P}}{d\mathbb{Q}}(x, \omega) = \frac{d\mu^H}{d\mathbb{Q}^x}(\omega) \quad \text{for } \mathbb{P}\text{-a.e. } (x, \omega).$$

Exploiting the bounds from Proposition 3.10 (which for given b are uniform in $x \in \mathbb{R}^d$) we find

$$\begin{aligned} \mathbb{E}_{\mathbb{Q}} \left[\left(\frac{d\mathbb{P}}{d\mathbb{Q}} \right)^n + \left(\frac{d\mathbb{Q}}{d\mathbb{P}} \right)^n \right] &= \int_{\mathbb{R}^d} \int_{C_T} \left[\left(\frac{d\mu^H}{d\mu^{\mathbb{Q}^x}}(\omega) \right)^n + \left(\frac{d\mathbb{Q}^x}{d\mu^H}(\omega) \right)^n \right] \mathbb{Q}^x(d\omega) \mu_0(dx) \\ &\lesssim_{n,b} \int_{\mathbb{R}^d} 1 \mu_0(dx) < \infty \end{aligned}$$

providing the conclusion. \square

Remark 3.19. It follows from the above that for any $p \in [1, \infty)$ and any $\varepsilon > 0$

$$\mathbb{E}_{\mathbb{P}} [\|X\|_{H-\varepsilon}^p] \leq \mathbb{E}_{\mathbb{Q}} [\|X\|_{H-\varepsilon}^{2p}]^{1/2} \mathbb{E}_{\mathbb{Q}} \left[\left(\frac{d\mathbb{P}}{d\mathbb{Q}} \right)^{2p} \right]^{1/2} \lesssim \mathbb{E}_{\mathbb{P}} [\|W\|_{H-\varepsilon}^{2p}]^{1/2} < \infty.$$

In particular, if $\xi \in \mathcal{P}_{\tilde{p}}$ for another $\tilde{p} \in [1, \infty)$, then $\mathbb{E}_{\mathbb{P}} [\|X\|_{H-\varepsilon}^{\tilde{p}}] < \infty$. As in the case of Remark 3.11, for fixed ξ the estimate can be performed uniformly over $\|b\|_E \leq M$.

3.4. Regularity of the solution laws. Although our main interest is the study of DDSDEs, our analysis also yields results on the regularity of the law $\mathcal{L}(X_t)$ for the solution to a standard SDE with singular drift. The method is quite simple but appears to be new and does not rely on PDE techniques nor Malliavin calculus; rather we exploit Girsanov transform and the averaging estimates for fBm, in combination with duality arguments.

Proposition 3.20. *Let b satisfy Assumption 3.12, X be the solution associated to (ξ, b) for random initial ξ independent of W . Then $\mathcal{L}(X.) \in L_T^{\bar{q}} B_{1,1}^{\bar{\alpha}}$ for all $(\bar{\alpha}, \bar{q}) \in (0, \infty) \times (1, 2)$ satisfying*

$$\bar{\alpha} < \frac{1}{H} \left(\frac{1}{\bar{q}} - \frac{1}{2} \right). \quad (3.19)$$

Proof. Observe that if $(\bar{\alpha}, \bar{q}, H)$ satisfy (3.19), then we can find $\varepsilon > 0$ small enough so that $(-\bar{\alpha} - 2\varepsilon, \bar{q}', H)$ satisfy the assumptions of Proposition 3.8, where \bar{q}' denotes the conjugate of \bar{q} . By Corollary 3.18, there exists an equivalent measure \mathbb{Q} such that $\mathcal{L}_{\mathbb{Q}}(X) = \mathcal{L}_{\mathbb{P}}(\xi + W)$, therefore for any $f \in L_T^{\bar{q}'} B_{\infty, \infty}^{-\bar{\alpha} - 2\varepsilon}$ it holds

$$\begin{aligned} \left| \int_0^T \langle f_s, \mathcal{L}_{\mathbb{P}}(X_s) \rangle ds \right| &\leq \mathbb{E}_{\mathbb{P}} \left[\left| \int_0^T f(s, X_s) ds \right| \right] \\ &\leq \mathbb{E}_{\mathbb{Q}} \left[\left(\frac{d\mathbb{P}}{d\mathbb{Q}} \right)^2 \right]^{1/2} \mathbb{E}_{\mathbb{Q}} \left[\left| \int_0^T f(s, X_s) ds \right|^2 \right]^{1/2} \\ &\lesssim \left(\int_{\mathbb{R}^d} \mathbb{E}_{\mu^H} \left[\left| \int_0^T f(s, x + \omega_s) ds \right|^2 \right] \mu_0(dx) \right)^{1/2} \lesssim \|f\|_{L_T^{\bar{q}'} B_{\infty, \infty}^{-\bar{\alpha} - 2\varepsilon}}. \end{aligned}$$

where in the last passage we used the fact that estimate (3.6) is uniform in $x \in \mathbb{R}^d$.

Using the embedding $B_{p,p}^{-\bar{\alpha} - \varepsilon} \hookrightarrow B_{\infty, \infty}^{-\bar{\alpha} - \varepsilon - d/p} \hookrightarrow B_{\infty, \infty}^{-\bar{\alpha} - 2\varepsilon}$ for $p < \infty$ big enough, by the duality $(L_T^{\bar{q}'} B_{p,p}^{-s})^* \simeq L_T^{\bar{q}} B_{p', p'}^s$ we deduce that $\mathcal{L}(X.) \in L_T^{\bar{q}} B_{p', p'}^{\bar{\alpha} + \varepsilon}$. Thus $h := (I - \Delta)^{\bar{\alpha}/2} \mathcal{L}(X.) \in L_T^{\bar{q}} B_{p', p'}^{\varepsilon} \hookrightarrow L_T^{\bar{q}} L_x^{p'}$; in order to conclude, it's enough to show that $h \in L_T^{\bar{q}} L_x^1$. Observe that $h \in L_{loc}^1([0, T] \times \mathbb{R}^d)$ and for any $\varphi \in L_T^{\bar{q}'} L_x^{\infty}$ it holds

$$\int_0^T \langle \varphi_s, h_s \rangle ds = \int_0^T \langle (I - \Delta)^{\bar{\alpha}/2} \varphi_s, \mathcal{L}(X_s) \rangle ds \lesssim \|(I - \Delta)^{\bar{\alpha}/2} \varphi\|_{L_T^{\bar{q}'} B_{\infty, \infty}^{-\bar{\alpha}}} \lesssim \|\varphi\|_{L_T^{\bar{q}'} L_x^{\infty}};$$

the conclusion then follows from an application of Lemma A.1 from Appendix A. \square

Proposition 3.21. *Let X, b, ξ be as in Proposition 3.20. Then $\mathcal{L}(X.) \in L_T^q L_x^p$ for all $(q, p) \in (1, \infty)^2$ satisfying*

$$\frac{1}{q} + \frac{Hd}{p} > Hd. \quad (3.20)$$

If in addition $Hd < 1$, then $\mathcal{L}(X.) \in L_T^q L_x^\infty$ for all $q \in (1, \infty)$ satisfying $q < (Hd)^{-1}$.

Proof. Observe that $(q, p) \in (1, \infty)^2$ satisfy (3.20) if and only if the conjugates (q', p') satisfy

$$\frac{1}{q'} + \frac{Hd}{p'} < 1.$$

By [32, Lemma 6.4] (more precisely equation (6.11) right after the proof therein) and estimates based on Girsanov theorem analogous to the proof of Proposition 3.20, we deduce that

$$\left| \int_0^T \langle f_s, \mathcal{L}(X_s) \rangle ds \right| \leq \mathbb{E} \left[\left| \int_0^T f(s, X_s) ds \right| \right] \lesssim \|f\|_{L_T^{q'} L_x^{p'}};$$

therefore by duality $\mathcal{L}(X.) \in L_T^q L_x^p$ if $(q, p) \in (1, \infty)^2$. Taking $p = \infty$, $q < (Hd)^{-1}$ (correspondingly $p' = 1, 1/q' < 1 - Hd$), we obtain

$$\left| \int_0^T \langle f_s, \mathcal{L}(X_s) \rangle ds \right| \lesssim \|f\|_{L_T^{q'} L_x^1} \quad \forall f \in L_T^{q'} L_x^1$$

and so we can conclude by Lemma A.2 in the Appendix that in this case $\mathcal{L}(X.) \in L_T^q L_x^\infty$. \square

4. PROOFS OF THE MAIN RESULTS

We split the proof of Theorem 2.4 into sections, which deal respectively with the cases $H > 1/2$ and $H \leq 1/2$; the proof of Theorem 2.5 is presented in Section 4.3 instead.

We recall to the reader that in this section we will be dealing with DDSDEs of the form

$$X_t = \xi + \int_0^t B_s(X_s, \mathcal{L}(X_s)) ds + W_t \quad \forall t \in [0, T] \quad (4.1)$$

with the drift B belonging to either $\mathcal{G}_p^{q, \alpha}$ or $\mathcal{H}_p^{\beta, \alpha}$ (cf. Definitions 2.2-2.1) depending on the value of $H \in (0, 1)$. The variable ξ is independent of W and with prescribed law $\mu_0 \in \mathcal{P}(\mathbb{R}^d)$, thus depending on the context we will treat both (ξ, B) and (μ_0, B) as the data of the problem.

4.1. The case $H > 1/2$. In this regime we will always consider drifts $B \in \mathcal{H}_p^{\beta, \alpha}$ with $\alpha, \beta > 0$ and $p \in [1, \infty)$. In particular here $B : [0, T] \times \mathbb{R}^d \times \mathcal{P}_p(\mathbb{R}^d) \rightarrow \mathbb{R}^d$ is bounded and uniformly continuous in all of its arguments; in this sense, although the concept of solution introduced in Section 3 does include the standard one by Remark 3.6, we do not employ it here.

Rather, we will simply say that a tuple (X, ξ, W) , defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, such that $\mathcal{L}_\mathbb{P}(\xi, W) = \mu_0 \otimes \mu^H$, is a solution to (4.1) if $\mathcal{L}(X_t) \in \mathcal{P}_p(\mathbb{R}^d)$ for all $t \in [0, T]$ and the integral equation (4.1) holds \mathbb{P} -a.s. The concepts of strong existence, pathwise uniqueness and uniqueness in law immediately carry over from the usual ones for SDEs.

Proposition 4.1. *Suppose $B \in \mathcal{H}_p^{H,\alpha}$ with*

$$H > \frac{1}{2}, \quad \alpha > 1 - \frac{1}{2H}, \quad p \in [1, \infty).$$

Then for any $\mu_0 \in \mathcal{P}_p(\mathbb{R}^d)$ strong existence, pathwise uniqueness and uniqueness in law hold for the DDSDE (4.1) with data (μ_0, B) .

Proof. We divide the proof in several steps.

Step 1: weak existence. By hypothesis $B : [0, T] \times \mathbb{R}^d \times \mathcal{P}_p(\mathbb{R}^d) \rightarrow \mathbb{R}^d$ is a uniformly continuous, bounded map; existence of weak solutions on $[0, T]$ then follows from [19, Proposition 3.10].

Step 2: any weak solution is a strong one. Let X be a weak solution of the DDSDE w.r.t. (ξ, W) on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Then setting $\mu_t = \mathcal{L}(X_t)$, $b^\mu(t, x) = B_t(x, \mu_t)$, X solves the SDE associated to b^μ , which satisfies $|b^\mu(t, x)| \leq \|B\|$ uniformly over (t, x) . As a consequence

$$d_p(\mu_s, \mu_t) \leq \|X_t - X_s\|_{L_\Omega^p} \leq \int_s^t \|b^\mu(r, X_r)\|_{L_\Omega^p} dr + \|W_t - W_s\|_{L_\Omega^p} \lesssim_{T,p} (1 + \|B\|)|t - s|^H \quad (4.2)$$

where we repeatedly applied Minkowski's inequality; by assumption

$$\begin{aligned} |b^\mu(t, x) - b^\mu(s, y)| &= |B_t(x, \mu_t) - B_s(y, \mu_s)| \\ &\leq \|B\| (|t - s|^{\alpha H} + |x - y|^\alpha + d_p(\mu_t, \mu_s)^\alpha) \\ &\lesssim_{T,p} (1 + \|B\|^2) (|t - s|^{\alpha H} + |x - y|^\alpha), \end{aligned}$$

where we applied (4.2) to obtain the last inequality. Namely, b^μ satisfies Assumption 3.12, implying that strong existence and uniqueness in law holds for the associated SDE; therefore X is adapted to (ξ, W) .

Step 3: reduction to the canonical space. As we are dealing with a strong solution X , we can regard it as a random variable on the canonical space $(\Omega, \mathcal{F}, \mathbb{P})$ with $\Omega = \mathbb{R}^d \times C_T$, $\mathbb{P} = \mu_0 \otimes \mu^H$, \mathcal{F} the \mathbb{P} -completion of $\mathcal{B}(\mathbb{R}^d \times C_T)$. Applying the same reasoning to any pair of weak (thus strong) solutions X^1, X^2 , possibly defined on different probability spaces, we can construct a coupling $(\tilde{X}^1, \tilde{X}^2)$ of solutions defined on the canonical space and w.r.t. the same random variables (ξ, W) . If we show that $\tilde{X}^1 \equiv \tilde{X}^2$, then the equality $\mathcal{L}(X^1) = \mathcal{L}(X^2)$ follows.

Step 4: pathwise uniqueness on the canonical space. Let us drop the tilde and adopt the notations $\mu_t^i = \mathcal{L}(X_t^i)$, $b^i(t, x) := B_t(x, \mu_t^i)$. It follows from the computations of Step 2 that we can find $M \sim 1 + \|B\|^2$ such that b^i satisfy Assumption 3.12 with $\|b^i\|_E \leq M$. We can therefore apply estimate (3.17) for the choice $\tilde{q} = q = \infty$, together with $X_0^1 = X_0^2 = \xi$, to find constants $\gamma > 1/2$, $C > 0$ such that

$$\mathbb{E} \left[\|X^1 - X^2\|_{\gamma; [0, \tau]}^p \right]^{1/p} \leq C \sup_{t \in [0, \tau]} \|b^1(t, \cdot) - b^2(t, \cdot)\|_{B_{\infty, \infty}^{\alpha-1}} \quad \forall \tau \in [0, T].$$

By assumption $B \in \mathcal{H}_p^{H,\alpha}$, therefore

$$\|b^1(t, \cdot) - b^2(t, \cdot)\|_{B_{\infty, \infty}^{\alpha-1}} \leq \|B\| d_p(\mu_t^1, \mu_t^2);$$

combining everything, using again $X_0^1 = X_0^2$, for any $\tau \in [0, T]$ it holds

$$\sup_{t \in [0, \tau]} d_p(\mu_t^1, \mu_t^2) \leq \tau^\gamma \mathbb{E} \left[\|X^1 - X^2\|_{\gamma; [0, \tau]}^p \right]^{1/p} \leq C \|B\| \tau^\gamma \sup_{t \in [0, \tau]} d_p(\mu_t^1, \mu_t^2).$$

Choosing $\bar{\tau}$ small enough so that $C\|B\|\bar{\tau}^\gamma < 1$, we conclude that $\mu_t^1 = \mu_t^2$ for all $t \in [0, \bar{\tau}]$ and so that $\mathbb{E}[\|X^1 - X^2\|_{\gamma; [0, \bar{\tau}]}] = 0$, i.e. \mathbb{P} -a.s. $X^1 \equiv X^2$ on $[0, \bar{\tau}]$. In light of this, choosing now $\tau = 2\bar{\tau}$, going through similar computations we have

$$\sup_{t \in [0, \tau]} d_p(\mu_t^1, \mu_t^2) = \sup_{t \in [\bar{\tau}, 2\bar{\tau}]} d_p(\mu_t^1, \mu_t^2) \leq \bar{\tau}^\gamma \mathbb{E} \left[\|X^1 - X^2\|_{\gamma; [\bar{\tau}, 2\bar{\tau}]}^p \right]^{1/p} \leq C\|B\|\bar{\tau}^\gamma \sup_{t \in [0, \tau]} d_p(\mu_t^1, \mu_t^2)$$

implying that the solutions also coincide on $[0, 2\bar{\tau}]$. Iterating the reasoning for $\tau = n\bar{\tau}$ until we cover $[0, T]$ gives the conclusion. \square

4.2. The case $H \leq 1/2$. In this case we can allow the drift to be singular, i.e. take values in $B_{\infty, \infty}^\alpha$ with $\alpha < 0$. We start by defining what we mean by solution to the DDSDE in this case.

Definition 4.2. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, (X, ξ, W) be a $C_T \times \mathbb{R}^d \times C_T$ -valued random variable defined on it with $\mathcal{L}_{\mathbb{P}}(\xi, W) = \mathcal{L}(\xi) \otimes \mu^H$; let $B : [0, T] \times \mathcal{P}_p \rightarrow \mathcal{S}'$ be a measurable map for some $p \in [1, \infty)$. We say that X is a solution to the DDSDE (4.1) associated to (ξ, B) if $\mu_t := \mathcal{L}_{\mathbb{P}}(X_t) \in \mathcal{P}_p$ for all $t \in [0, T]$ and setting $b^\mu(t, \cdot) = B_t(\mu_t)(\cdot)$, X is a pathwise solution to the SDE

$$X_t = \xi + \int_0^t b^\mu(s, X_s) ds + W_t,$$

associated to (b^μ, ξ, W) in the sense of Definition 3.4. All the concepts of strong solution, pathwise uniqueness and uniqueness in law are similarly readapted from those of Definition 3.5.

As before, we will consider both (ξ, B) and (μ_0, B) to be the data of the problem, depending on whether we are focusing on solutions on a prescribed probability space or on their laws.

Assume now $B \in \mathcal{G}_p^{q, \alpha}$ with

$$H \leq \frac{1}{2}, \quad \alpha > 1 + \frac{1}{qH} - \frac{1}{2H}, \quad q \in (2, \infty], \quad p \in [1, \infty); \quad (4.3)$$

then to any $\mu \in C_T \mathcal{P}_p$ we can associated a singular drift $b^\mu(t, \cdot) = B(t, \mu_t)(\cdot)$ such that

$$\|b^\mu(t, \cdot)\|_{B_{\infty, \infty}^\alpha} \leq h_t,$$

where $h \in L_T^q$ is the function associated to B from Definition 2.2. Thus b^μ satisfies Assumption 3.12 and the associated SDE has a unique solution X by Corollary 3.17; if in addition $\mu_0 \in \mathcal{P}_p$, then by Remark 3.19 the map $t \mapsto \mathcal{L}(X_t)$ belongs to $C_T \mathcal{P}_p$.

Thus for fixed μ_0 , setting $\mathcal{I}^{\mu_0}(\mu) = \mathcal{L}(X)$, we can define a map \mathcal{I}^{μ_0} from $C_T \mathcal{P}_p$ to itself; this map comes with an alternative notion of solution to the DDSDE.

Definition 4.3. Assume $B \in \mathcal{G}_p^{q, \alpha}$ with parameters satisfying (4.3), $\mu_0 \in \mathcal{P}_p$; we say that a flow of measures $\mu \in C_T \mathcal{P}_p$ is a solution to the DDSDE associated to (μ_0, B) if it satisfies $\mathcal{I}^{\mu_0}(\mu) = \mu$.

The next lemma clarifies the relation between Definitions 4.2 and 4.3.

Lemma 4.4. Let $B \in \mathcal{G}_p^{q, \alpha}$ with parameters satisfying (4.3), $\mu_0 \in \mathcal{P}_p$. The following hold:

- i. if X is a weak solution to (4.1), then $\mu_t = \mathcal{L}(X_t)$ is a fixed point for \mathcal{I}^{μ_0} ;
- ii. if μ is a fixed point for \mathcal{I}^{μ_0} , then there exists a strong solution X to (4.1);
- iii. if there exists at most one fixed point for \mathcal{I}^{μ_0} , then pathwise uniqueness and uniqueness in law hold for (4.1).

Proof. Point **i.** immediately follows from the definitions. To see Point **ii.**, assume $\mathcal{I}^{\mu_0}(\mu) = \mu$. and set $b^\mu(t, \cdot) = B_t(\mu_t)(\cdot)$; then $b^\mu \in L_T^q B_{\infty, \infty}^\alpha$, so by the results of Section 3, we can construct a strong solution X to the SDE associated to (μ_0, b^μ) . But then by definition of \mathcal{I}^{μ_0} it holds $\mathcal{L}(X_t) = \mu_t$ and so X solves the DDSDE. It remains to show Point **iii.**; assume X^i are two solutions and set $\mu^i = \mathcal{L}(X^i)$. Then by Point **ii.**, μ^i are both fixed points for \mathcal{I}^{μ_0} , so $\mu^1 = \mu^2$ and $b^{\mu^1} = b^{\mu^2}$. But then X^i both solve the SDE associated to b^{μ^1} , for which uniqueness holds both pathwise and in law, so the conclusion follows. \square

It follows from the above that, in order to show strong existence, pathwise uniqueness and uniqueness in law for the DDSDE (4.1) in the sense of Definition 4.2, it's enough to show that there exists exactly one solution $\mu \in C_T \mathcal{P}_p$ in the sense of Definition 4.3.

Proposition 4.5. *Let $B \in \mathcal{G}_p^{q, \alpha}$ with parameters satisfying (4.3); then for any $\mu_0 \in \mathcal{P}_p(\mathbb{R}^d)$ strong existence, pathwise uniqueness and uniqueness in law hold for the DDSDE (4.1) associated to (μ_0, B) .*

Proof. Define the map $\mathcal{I}^{\mu_0} : C_T \mathcal{P}_p \rightarrow C_T \mathcal{P}_p$ associated to (μ_0, B) as before; in order to show that there exists exactly one fixed point to \mathcal{I}^{μ_0} , it's enough to establish its contractivity.

Given $\mu^i \in C_T \mathcal{P}_p$, $i = 1, 2$, set $b^i := b^{\mu^i} = B(t, \mu_t^i)$; denote by X^i two solutions, defined on the same probability space and with respect to the same data (ξ, W) , to the SDEs associated to (ξ, b^i) , where $\mathcal{L}(\xi) = \mu_0$. By definition of $\mathcal{G}_p^{q, \alpha}$, there exists $h \in L_T^q$, such that for any $\tau \in (0, T]$ we have

$$\|b_t^1 - b_t^2\|_{B_{\infty, \infty}^{\alpha-1}} \leq h_t \sup_{t \in [0, \tau]} d_p(\mu_t^1, \mu_t^2).$$

Applying Corollary 3.17, using the fact that $X_0^1 = X_0^2 = \xi$, we can find $\gamma > 1/2$ and $C > 0$ such that for any $\tau \in (0, T]$, we have

$$\begin{aligned} \sup_{t \in [0, \tau]} d_p(\mathcal{I}^{\mu_0}(\mu^1)_t, \mathcal{I}^{\mu_0}(\mu^2)_t) &\leq \sup_{t \in [0, \tau]} \mathbb{E}_{\mathbb{P}}[|X_t^1 - X_t^2|^p]^{\frac{1}{p}} \leq \tau^\gamma \mathbb{E}_{\mathbb{P}}[\|X^1 - X^2\|_{\gamma; [0, \tau]}^p]^{\frac{1}{p}} \\ &\leq C \tau^\gamma \left(\int_0^\tau \|b_t^1 - b_t^2\|_{B_{\infty, \infty}^{\alpha-1}}^q dt \right)^{\frac{1}{q}} \leq C \|h\|_{L_T^q} \tau^\gamma \sup_{t \in [0, \tau]} d_p(\mu_t^1, \mu_t^2). \end{aligned}$$

Choosing $\bar{\tau} > 0$ sufficiently small such that $C \|h\|_{L_T^q} \bar{\tau}^\gamma < 1$, we find that \mathcal{I}^{μ_0} is a contraction from $C([0, \bar{\tau}]; \mathcal{P}_p)$ to itself, so therein there exists a unique fixed point $\bar{\mu} = \mathcal{I}^{\mu_0}(\bar{\mu})$; it remains to show we can extend uniquely this fixed point to the whole interval $[0, T]$.

To do this, the classical argument for SDEs would require to restart the equation at $t = \bar{\tau}$; however we can't perform this, as the fractional Brownian motion is not a Markov process. We can exploit the fact that $\bar{\tau}$ only depends on $C \|h\|_{L_T^q}$ and not the history of the paths X^i nor μ^i to give the following alternative reasoning.

Given $\bar{\tau}, \bar{\mu} \in C([0, \bar{\tau}]; \mathcal{P}_p)$ as above, consider $E := \{\mu \in C([0, 2\bar{\tau}]; \mathcal{P}_p) : \mu|_{[0, \bar{\tau}]} = \bar{\mu}\}$, which is a closed subset of $C([0, 2\bar{\tau}]; \mathcal{P}_p)$ and thus a complete metric space with the same norm. Since $\bar{\mu}$ is a fixed point on $[0, \bar{\tau}]$, \mathcal{I}^{μ_0} leaves E invariant; for any $\mu^i \in E$, arguing as above it holds

$$\begin{aligned} \sup_{t \in [0, 2\bar{\tau}]} d_p(\mathcal{I}^{\mu_0}(\mu^1)_t, \mathcal{I}^{\mu_0}(\mu^2)_t) &= \sup_{t \in [\bar{\tau}, 2\bar{\tau}]} d_p(\mathcal{I}^{\mu_0}(\mu^1)_t, \mathcal{I}^{\mu_0}(\mu^2)_t) \leq \sup_{t \in [\bar{\tau}, 2\bar{\tau}]} \bar{\tau}^\gamma \mathbb{E}_{\mathbb{P}}[\|X^1 - X^2\|_{\gamma; [\bar{\tau}, 2\bar{\tau}]}^p]^{\frac{1}{p}} \\ &\leq \bar{\tau}^\gamma C \|h\|_{L_T^q} \sup_{t \in [0, 2\bar{\tau}]} d_p(\mu_t^1, \mu_t^2). \end{aligned}$$

It follows that \mathcal{I}^{μ_0} is a contraction on E and admits a unique fixed point on it, which is necessarily the only possible extension of $\bar{\mu}$ on $[0, 2\bar{\tau}]$. Repeating the argument on $[0, n\bar{\tau}]$ as many times as necessary to cover $[0, T]$ concludes the proof. \square

4.3. Stability estimates for DDSDEs. The purpose of this section is to provide the proof of Theorem 2.5, which loosely speaking establishes Lipschitz dependence of the solutions $\mu^i \in C_T \mathcal{P}_p$ in terms of the data (μ_0^i, B^i) for $i = 1, 2$.

We assume that we are given drifts B^i belonging to $\mathcal{H}_p^{H, \alpha}$ for parameters satisfying (2.2) when $H > 1/2$, respectively $B^i \in \mathcal{G}_p^{q, \alpha}$ for parameters satisfying (2.3) when $H \leq 1/2$; in both cases we denote the optimal constants by $\|B^i\|$. Given $\mu_0^i \in \mathcal{P}_p$, we denote by $\mu^i \in C_T \mathcal{P}_p$ the unique solutions associated to (μ_0^i, B^i) , whose existence is granted by Theorem 2.4.

Finally, for α, q given as above, let us recall the notation introduced in Theorem 2.5:

$$\|B^1 - B^2\|_\infty := \sup_{(t, \nu) \in [0, T] \times \mathcal{P}_p} \|B_t^1(\nu) - B_t^2(\nu)\|_{B_{\infty, \infty}^{\alpha-1}}$$

and

$$\|B^1 - B^2\|_{q, \infty} := \left(\int_0^T \sup_{\nu \in \mathcal{P}_p} \|B_t^1(\nu) - B_t^2(\nu)\|_{B_{\infty, \infty}^{\alpha-1}}^q dt \right)^{1/q}.$$

Proof of Theorem 2.5. Let μ^i be the solutions as above and set $b_t^i := B^i(t, \mu_t^i)$. Recall from the proofs of Propositions 4.1 and 4.5 that if $\|B^i\| \leq M$, then $\|b^i\|_E \leq C(M)$, E being suitable spaces for which Assumption 3.12 is met; so we are in a position to apply estimates from Section 3.3. First observe that, by addition and subtraction of $B_t^1(\mu_t^2)$, we have

$$\|b_t^1 - b_t^2\|_{B_{\infty, \infty}^{\alpha-1}} \leq \|B_t^1(\mu_t^1) - B_t^1(\mu_t^2)\|_{B_{\infty, \infty}^{\alpha-1}} + \sup_{\nu \in \mathcal{P}_p} \|B_t^1(\nu) - B_t^2(\nu)\|_{B_{\infty, \infty}^{\alpha-1}}. \quad (4.4)$$

The argument slightly differs in the $H > \frac{1}{2}$ and $H \leq \frac{1}{2}$ cases, so we will handle them separately.

We begin with $H > \frac{1}{2}$. Let us choose $\tilde{q} < \infty$ big enough so that (α, \tilde{q}, H) satisfies (3.12); then we can apply estimate (3.18) to obtain

$$\sup_{t \in [0, \tau]} d_p(\mu_t^1, \mu_t^2)^{\tilde{q}} \lesssim_{M, \tilde{q}, T} d_p(\mu_0^1, \mu_0^2)^{\tilde{q}} + \left(\int_0^\tau \|b_s^1 - b_s^2\|_{B_{\infty, \infty}^{\alpha-1}}^{\tilde{q}} ds \right); \quad (4.5)$$

on the other hand, by estimate (4.4) and the assumption $B^1 \in \mathcal{H}_p^{H, \alpha}$ with $\|B^1\| \leq M$, it holds

$$\|b_s^1 - b_s^2\|_{B_{\infty, \infty}^{\alpha-1}} \leq M \sup_{r \in [0, s]} d_p(\mu_r^1, \mu_r^2) + \|B^1 - B^2\|_\infty.$$

Putting everything together, setting $f_t := \sup_{s \in [0, t]} d_p(\mu_s^1, \mu_s^2)^{\tilde{q}}$, we obtain

$$f_t \lesssim_{\tilde{q}} f_0 + \int_0^t M^{\tilde{q}} f_s ds + T \|B^1 - B^2\|_\infty^{\tilde{q}} \quad \forall t \in [0, T];$$

applying Grönwall to f and taking the power $1/\tilde{q}$ on both sides readily gives

$$\sup_{t \in [0, T]} d_p(\mu_t^1, \mu_t^2) \lesssim d_p(\mu_0^1, \mu_0^2) + \|B^1 - B^2\|_\infty$$

which is exactly the desired estimate (2.4).

Suppose now X^i are solutions defined on the same probability space, then combining estimate (3.17) with the ones above we find

$$\begin{aligned} \mathbb{E}[\|X^1 - X^2\|_\gamma^p]^{1/p} &\lesssim \|X_0^1 - X_0^2\|_{L_\Omega^p} + \sup_{t \in [0, T]} \|b_t^1 - b_t^2\| \\ &\lesssim_M \|X_0^1 - X_0^2\|_{L_\Omega^p} + \sup_{t \in [0, T]} d_p(\mu_t^1, \mu_t^2) + \|B^1 - B^2\|_\infty \\ &\lesssim \|X_0^1 - X_0^2\|_{L_\Omega^p} + d_p(\mu_0^1, \mu_0^2) + \|B^1 - B^2\|_\infty \end{aligned}$$

and the conclusion readily follows from $d_p(\mu_0^1, \mu_0^2) \leq \|X_0^1 - X_0^2\|_{L_\Omega^p}$.

We now move on to the case $H \leq \frac{1}{2}$; for $q = \infty$ the proof is the same as above, so we can assume w.l.o.g. $q < \infty$ here. For $B^i \in \mathcal{G}_p^{q, \alpha}$, it follows again by (4.4) that

$$\|b_t^1 - b_t^2\|_{B_{\infty, \infty}^{\alpha-1}} \leq h_t^1 \sup_{r \in [0, t]} d_p(\mu_r^1, \mu_r^2) + \sup_{\nu \in \mathcal{P}_p} \|B_t^1(\nu) - B_t^2(\nu)\|_{B_{\infty, \infty}^{\alpha-1}}, \quad (4.6)$$

where we recall that $h^i \in L_T^q$ are the functions associated to B^i given in Definition 2.2. Following the same strategy as before, by (4.6) and (4.5), $f_t := \sup_{s \in [0, t]} d_p(\mu_s^1, \mu_s^2)^q$ satisfies

$$f_t \leq f_0 + \int_0^t |h_s^1|^q f_s \, ds + \|B^1 - B^2\|_{q, \infty} \quad \forall t \in [0, T];$$

by this inequality and the assumption $\|h^1\|_{L_T^q} = \|B^1\| \leq M$, we conclude again by Grönwall's inequality that

$$\sup_{t \in [0, T]} d_p(\mu_t^1, \mu_t^2) \lesssim_M d_p(\mu_0^1, \mu_0^2) + \|B^1 - B^2\|_{q, \infty}$$

which gives estimate (2.6). The statement for $\mathbb{E}[\|X^1 - X^2\|_\gamma^p]$ now follows exactly as in the case $H > 1/2$. \square

We conclude this section with the application of Theorem 2.5 to a particularly relevant case.

Example 4.6. Let $H \in (0, 1)$, $\alpha > 1 - 1/(2H)$ and set

$$E = \begin{cases} C_T^{\alpha H} C_x^0 \cap C_T^0 C_x^\alpha & \text{for } H > \frac{1}{2} \\ L_T^\infty B_{\infty, \infty}^\alpha & \text{for } H \leq \frac{1}{2} \end{cases}.$$

Let B^i be of the form $B^i(t, \mu) = f_t^i + g_t^i * \mu$ for $f^i, g^i \in E$ with $\|f^i\|_E, \|g^i\|_E \leq M$. Then B^i satisfy the assumptions of Theorem 2.5 in \mathcal{P}_p for any $p \in [1, \infty)$ and estimate (2.4) becomes

$$\sup_{t \in [0, T]} d_p(\mu_t^1, \mu_t^2) \lesssim d_p(\mu_0^1, \mu_0^2) + \|f^1 - f^2\|_{L_T^\infty B_{\infty, \infty}^{\alpha-1}} + \|g^1 - g^2\|_{L_T^\infty B_{\infty, \infty}^{\alpha-1}}.$$

5. REFINED RESULTS IN THE CONVOLUTIONAL CASE

In this section we focus on the case of DDSDEs with convolutional structure, namely

$$X_t = \xi + \int_0^t (b_s * \mu_s)(X_s) \, ds + W_t, \quad \mu_t = \mathcal{L}(X_t) \quad \forall t \in [0, T]. \quad (5.1)$$

They correspond to the case $B_t(\mu) = b_t * \mu$ and can therefore be solved under suitable assumptions on b (e.g. $b \in E$ as in Example 4.6). Due to their specific structure however, as soon as the associated solution X has a regular law μ , its regularity immediately transfers to the drift $b_t^\mu = b_t * \mu_t$, as the next simple lemma shows.

Lemma 5.1. *Let $H \in (0, 1)$, $b \in B_{\infty, \infty}^\alpha$ for $\alpha > 1 - 1/(2H)$, $\mu_0 \in \mathcal{P}_1$ and X denote the unique solution to the DDSDE (5.1) with $\mathcal{L}(\xi, W) = \mu_0 \otimes \mu^H$. Then X also solves an SDE with drift b^μ which belongs to $L_T^1 C_x^1$.*

Proof. Let X be the aforementioned solution, then by the proof of Theorem 2.4 we know that it solves an SDE with drift b^μ satisfying Assumption 3.12; applying Proposition 3.20 for the choice $q = 1$ in (3.19), we deduce that $\mu \in L_T^1 B_{1,1}^{\tilde{\alpha}}$ for any $\tilde{\alpha} < 1/(2H)$. Therefore by the hypothesis and Young's inequality it holds $b^\mu = b * \mu \in L_T^1 B_{\infty, \infty}^{\alpha + \tilde{\alpha}}$ for some $\alpha > 1 - 1/(2H)$ and all $\tilde{\alpha} < 1/(2H)$; choosing $\tilde{\alpha}$ appropriately gives the conclusion. \square

Remark 5.2. Up to technicalities, the proof readapts to the case of time-dependent drifts $B_t(\mu) = b_t * \mu$ with b satisfying Assumption 3.12, with the same conclusion that $b^\mu \in L_T^1 C_x^1$.

Lemma 5.1 shows that in this setting the effective drift b^μ is much more regular than the original b , to the point that the SDE associated to b^μ can be solved classically. However in order to give meaning to the DDSDE, it suffices to know that b^μ satisfies the weaker Assumption 3.12; for this reason we expect the criteria coming from Theorem 2.4 to be suboptimal for convolutional DDSDEs (5.1), as they don't take into account the different regularity of b and b^μ .

A partial improvement of those results is given by Theorems 2.6 and 2.7, whose proofs are presented respectively in Sections 5.1 and 5.2; they rely heavily on general results presented in Appendix A, especially Lemma A.7 and Corollary A.9. The first provides the basic estimate $\|b * (\mu - \nu)\|_{B_{\infty, \infty}^\alpha} \lesssim \|b\|_{B_{\infty, \infty}^\alpha}$, which is true for all measures $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$; while the latter shows a similar result under additional assumptions on μ and ν but requiring less integrability on b (in the sense that b is only required to be in $B_{p,p}^\alpha$ instead of $B_{\infty, \infty}^\alpha$). For example, a special case of Corollary A.9 provides the estimate $\|b * (\mu - \nu)\|_{B_{\infty, \infty}^{\alpha-1}} \lesssim \|b\|_{B_{p,p}^\alpha} d_r(\mu, \nu)(1 + \|\mu\|_{L^q} \|\nu\|_{L^q})$, as long as the relation $1 \leq r/p + 1/q$ holds.

Before moving further, let us rigorously define what we mean by solutions here, although the concept is very similar to that of Definition 4.2.

Definition 5.3. Fix $H \in (0, 1)$; let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, (X, ξ, W) be a $C_T \times \mathbb{R}^d \times C_T$ -valued random variable defined on it with $\mathcal{L}_{\mathbb{P}}(\xi, W) = \mu_0 \otimes \mu^H$ and b be a distributional drift. We say that X is a solution to the DDSDE (5.1) associated to (μ_0, b) if setting $\mu_t := \mathcal{L}_{\mathbb{P}}(X_t)$, $b_t^\mu := b_t * \mu_t$, X satisfies the SDE

$$X_t = \xi + \int_0^t b_s^\mu(X_s) ds + W_t \quad \forall t \in [0, T],$$

where we additionally require that either:

- i. b^μ satisfies Assumption 3.12 and the SDE is interpreted in the sense of Definition 3.4, or
- ii. $b^\mu \in L_T^1 C_x^0$ and the SDE is interpreted in the standard integral sense.

All the concepts of weak solution, strong solution, pathwise uniqueness and uniqueness in law are readapted similarly.

A major role in the proofs of Theorems 2.6 and 2.7 is given by the following conditional uniqueness result.

Proposition 5.4. *Let $H \in (0, 1)$, $p \in [1, \infty)$, p' its conjugate exponent; let b be a distributional drift satisfying one of the following conditions:*

- i. *If $H > 1/2$, then $b \in C_T^{\alpha H} L_x^p$ and $b \in C_T^0 B_{p,p}^\alpha$ with $\alpha > 1 - \frac{1}{2H}$.*

ii. If $H \leq 1/2$ then $b \in L_T^q B_{p,p}^\alpha$, with $\alpha > 1 - \frac{1}{2H} + \frac{1}{Hq}$.

Assume furthermore that for a given $\mu_0 \in L_x^{p'}$ there exists a weak solution X to the DDSDE (5.1) associated to (μ_0, b) , satisfying

$$\sup_{t \in [0, T]} \|\mathcal{L}(X.)\|_{L_x^{p'}} < \infty. \quad (5.2)$$

Then X is a strong solution; moreover it is the unique one (both pathwise and in law) in the class of solutions satisfying condition (5.2)

Proof. We handle the cases $H \leq 1/2$ and $H > 1/2$ slightly differently.

The case $H \leq 1/2$. First observe that, if X satisfies (5.2), then by Young's inequality $b^\mu = b * \mathcal{L}(X.) \in L_T^q B_{\infty, \infty}^\alpha$; in particular b^μ satisfies Assumption 3.12, Definition 5.3 is meaningful and X is necessarily a strong solution.

Now let X^i , $i = 1, 2$, be two solutions to (5.1) satisfying (5.2); as they are both strong solutions, by the usual arguments, we can assume them to be defined on the same probability space, w.r.t. the same (ξ, W) , and we only need to check that $X^1 = X^2$ \mathbb{P} -a.s. Moreover thanks to the strict inequality $\alpha > 1 - \frac{1}{2H} + \frac{1}{Hq}$ here we can assume w.l.o.g. $q < \infty$.

For $i = 1, 2$, set $\mu_t^i = \mathcal{L}(X_t^i)$; as $b * \mu^i$ both satisfy Assumption 3.12, we may apply Corollary 3.17 to find

$$\begin{aligned} \sup_{t \in [0, T]} d_{r'}(\mu_t^1, \mu_t^2)^{r'} &\leq \sup_{t \in [0, T]} \mathbb{E}[\|X_t^1 - X_t^2\|^{r'}] \leq \mathbb{E}[\|X^1 - X^2\|_\gamma^{r'}] \lesssim \|b * (\mu^1 - \mu^2)\|_{L_T^q B_{\infty, \infty}^{\alpha-1}}^{r'} \\ &\lesssim \|b\|_{L_T^q B_{p,p}^\alpha}^{r'} (\|\mu^1\|_{L_T^\infty L_x^{p'}}^{r'} + \|\mu^2\|_{L_T^\infty L_x^q}^{r'}) < \infty. \end{aligned}$$

In particular the quantity $d_{r'}(\mu_t^1, \mu_t^2)$ is finite for any $r' \in [1, \infty)$ and any $t \in [0, T]$.

We now wish to apply Corollary A.9 from Appendix A to obtain better control on the difference of the drifts $b * \mu^1 - b * \mu^2$. To do so observe that, under our assumptions on the parameters (α, q, p) , we can find new parameters $(s, r) \in (1, \infty)^2$ with s large and r close to 1 such that

$$\alpha - \frac{d}{s} > 1 - \frac{1}{2H} + \frac{1}{Hq}, \quad 1 + \frac{r}{s} \leq \frac{r}{p} + \frac{1}{p'}. \quad (5.3)$$

For this choice, set $\tilde{\alpha} := \alpha - d/s$; by construction the parameters (p, p', s, r) satisfy the assumptions of Corollary A.9 from Appendix A; its application, together with standard Besov embeddings, yields

$$\begin{aligned} \|b_t * (\mu_t^1 - \mu_t^2)\|_{B_{\infty, \infty}^{\tilde{\alpha}-1}} &\lesssim \|b_t * (\mu_t^1 - \mu_t^2)\|_{B_{s,s}^{\alpha-1}} \\ &\lesssim \|b_t\|_{B_{p,p}^\alpha} (\|\mu_t^1\|_{L^{p'}}^{1/r} + \|\mu_t^2\|_{L^{p'}}^{1/r}) d_{r'}(\mu_t^1, \mu_t^2). \end{aligned}$$

Now since under (5.3) the triple $(\tilde{\alpha}, q, H)$ also satisfies (3.12), we can again apply estimate (3.17) from Corollary 3.17 to find

$$d_{r'}(\mu_t^1, \mu_t^2)^q \leq \|X_t^1 - X_t^2\|_{L_\Omega^{r'}}^q \lesssim \int_0^t \|b_u * (\mu_u^1 - \mu_u^2)\|_{B_{\infty, \infty}^{\tilde{\alpha}-1}}^q du \lesssim \int_0^t \|b_u\|_{B_{p,p}^\alpha}^q d_{r'}(\mu_u^1, \mu_u^2)^q du.$$

Applying Grönwall's lemma we conclude that $d_{r'}(\mu_t^1, \mu_t^2) = 0$ and so $\mu_t^1 = \mu_t^2$ for all $t \in [0, T]$. Thus, X^i are solutions to the same SDE and therefore $X^1 = X^2$ \mathbb{P} -a.s.

The case $H > 1/2$. We argue essentially in the same way, only this time checking that X is a strong solution starting from the available information on b and $\mathcal{L}(X.)$ is less straightforward.

First observe that Young's inequality still provides $b^\mu \in C_T^0 C_x^\alpha$, so that by Definition 5.3 the DDSDE is meaningful in the classical integral sense. In order to check Assumption 3.12 for b^μ (which implies X being strong), it remains to show that $b * \mu \in C_T^{\tilde{\alpha}H} C_x^0$ for some $\tilde{\alpha}$ such that

$$1 - \frac{1}{2H} < \tilde{\alpha} \leq \alpha.$$

By addition and subtraction, $b_t^\mu - b_s^\mu = (b_t - b_s) * \mu_t + b_s * (\mu_t - \mu_s)$; by the hypothesis on b, μ , we can estimate the first term by

$$\|(b_t - b_s) * \mu_s\|_{C_x^0} \leq \|b_t - b_s\|_{L_x^p} \|\mu\|_{L_x^\infty} \lesssim |t - s|^{\alpha H}. \quad (5.4)$$

Since X is solution to the SDE $X_t = \xi + \int_0^t b_s^\mu(X_s) ds + W_t$, for any $r' \in (1, \infty)$ it holds

$$d_{r'}(\mu_t, \mu_s) \leq \|X_t - X_s\|_{L_\Omega^{r'}} \leq \|b * \mu^i\|_{L_T^\infty C_x^0} |t - s| + \|W_t - W_s\|_{L_\Omega^{r'}} \lesssim_{r', T} |t - s|^H.$$

Now similarly to the case $H \leq 1/2$, choose $(s, r) \in (1, \infty)^2$ such that

$$\tilde{\alpha} := \alpha - \frac{d}{s} > 1 - \frac{1}{2H}, \quad 1 + \frac{r}{s} \leq \frac{r}{p} + \frac{1}{p'};$$

Applying Corollary A.9 and the previous estimate for $d_{r'}(\mu_t, \mu_s)$, we then find

$$\|b_t * (\mu_t^i - \mu_s^i)\|_{B_{\infty, \infty}^{\tilde{\alpha}-1}} \lesssim d_{r'}(\mu_t, \mu_s) \lesssim |t - s|^H.$$

On the other hand, since $b \in C_T^0 B_{p,p}^\alpha$ and $\mu \in L_T^\infty L_x^{p'}$, by Young's inequality, we have $\|b_t * (\mu_t^i - \mu_s^i)\|_{B_{\infty, \infty}^{\tilde{\alpha}-1}} \lesssim 1$. We can now interpolate between the two estimates: choose $\theta = \tilde{\alpha} \in (0, 1)$, so that $\theta(\tilde{\alpha} - 1) + (1 - \theta)\alpha = (1 - \tilde{\alpha})(\alpha - \tilde{\alpha}) > 0$, then by the embedding $B_{\infty, \infty}^\varepsilon \hookrightarrow C_x^0$ for $\varepsilon > 0$, we obtain

$$\|b_t * (\mu_t^i - \mu_s^i)\|_{C_x^0} \lesssim \|b * (\mu_t^i - \mu_s^i)\|_{B_{\infty, \infty}^{\tilde{\alpha}-1}}^{1-\tilde{\alpha}} \|b_t * (\mu_t^i - \mu_s^i)\|_{B_{\infty, \infty}^{\tilde{\alpha}-1}}^{\tilde{\alpha}} \lesssim |t - s|^{\tilde{\alpha}H}.$$

So we conclude that $b * \mu^i \in C_T^0 C_x^{\tilde{\alpha}} \cap C_T^{\tilde{\alpha}H} C_x^0$, where by construction $\tilde{\alpha} > 1 - 1/(2H)$.

The second part of the argument, concerning the comparison of two solutions X^i satisfying (5.2), now proceeds identically as in the case $H \leq 1/2$. \square

5.1. Distributional kernels with bounded divergence. Proposition 5.4 reduces the problem of uniqueness of solutions (in a suitable class) to that of establishing their regularity, in the sense of equation (5.2).

One classical way to show that the condition $\mu_0 \in L_x^{p'}$ is propagated at positive times, which has been exploited systematically after [12], is to impose boundedness of $\operatorname{div} b$; in the setting of DDSDEs with general additive noise and regular drift b , an analogous statement can be found in [19, Proposition 4.3].

Proposition 5.5. *Let $H \in (0, 1)$, $q \in (2, \infty]$, $p \in [1, \infty)$, p' its conjugate exponent. Let b be a distributional drift such that $\operatorname{div} b \in L_T^1 L_x^\infty$ and either:*

- i. *If $H > 1/2$, then $b \in C_T^{\alpha H} L_x^p \cap C_T^0 B_{p,p}^\alpha$ for some $\alpha > 1 - \frac{1}{2H}$.*
- ii. *If $H \leq 1/2$, then $b \in L_T^q B_{p,p}^\alpha$ for some $\alpha > 1 - \frac{1}{2H} + \frac{1}{Hq}$.*

Then for any $\mu_0 \in L_x^{p'}$ there exists a strong solution to the DDSDE (5.1) associated to (μ_0, b) , which moreover satisfies $\mathcal{L}(X) \in L_T^\infty L_x^{p'}$.

Proof. We start by dealing with the case $H \leq 1/2$; at the end of the proof we explain how the reasoning needs to be modified for $H > 1/2$.

The case $H \leq 1/2$. In this case we can assume w.l.o.g. $q < \infty$; recall that if f^n is a bounded sequence in $L_T^q B_{\infty,\infty}^\alpha$, for (α, q) satisfying Assumption 3.12 such that $f^n \rightarrow f$ in $L_T^q B_{\infty,\infty}^{\alpha-1}$, then by Corollary 3.17 the associated solutions

$$X_t^n = \xi + \int_0^t f^n(X_s^n) ds + W_t$$

converge to the unique strong solution X of the SDE associated to (ξ, W, f) (we can assume $\{X^n\}_{n \geq 1}$ and X to be defined on the same probability space for the same (ξ, W)).

Given b as in the hypothesis, consider a sequence of smooth, bounded functions b^n such that $b^n \rightarrow b$ in $L_T^q B_{p,p}^\alpha$ with $\|\operatorname{div} b^n\|_{L_T^1 L_x^\infty} \leq \|\operatorname{div} b\|_{L_T^1 L_x^\infty}$; let X^n be the solutions to

$$X_t^n = \xi + \int_0^t b^n * \mathcal{L}(X_s^n)(X_s^n) ds + W_t,$$

whose existence is granted by classical results (see e.g. [11, Theorem 7]) and set $\mu_t^n = \mathcal{L}(X_t^n)$. By [19, Proposition 4.3], there exists $C = C(\|\operatorname{div} b\|_{L_T^1 L_x^\infty}) > 0$ such that

$$\sup_{n \in \mathbb{N}} \sup_{t \in [0, T]} \|\mu_t^n\|_{L_x^{p'}} \leq C \|\mu_0\|_{L_x^{p'}} < \infty.$$

As a consequence, each X^n solves an SDE with drift $f^n = b^n * \mu^n$ satisfying

$$\sup_{n \in \mathbb{N}} \|f^n\|_{L_T^q B_{\infty,\infty}^\alpha} \lesssim \sup_{n \in \mathbb{N}} \|b^n\|_{L_T^q B_{p,p}^\alpha} \sup_{n \in \mathbb{N}} \sup_{t \in [0, T]} \|\mu_t^n\|_{L_x^{p'}} \lesssim C \|b\|_{L_T^q B_{p,p}^\alpha} \|\mu_0\|_{L_x^{p'}} < \infty.$$

In turn this implies by Remark 3.19 that for any fixed $\varepsilon > 0$ we have the uniform estimate $\sup_n \mathbb{E}[\|X^n\|_{C^{H-\varepsilon}}] < \infty$; since moreover $X_0^n = \xi$ for all $n \in \mathbb{N}$, we can conclude by Ascoli–Arzelà that the sequence $\{X^n\}_{n \geq 1}$ is tight in C_T . We can then extract a (not relabelled) subsequence such that $\mathcal{L}(X^n)$ converge weakly to some $\mu \in \mathcal{P}(C_T)$; consequently $\mu_t^n \rightarrow \mu_t$ in $\mathcal{P}(\mathbb{R}^d)$ for any $t \in [0, T]$, where $\mu_t = e_t \# \mu$ and $e_t : C_T \rightarrow \mathbb{R}^d$ is the evaluation map. It follows from the uniform estimates that $\|\mu_t\|_{L_x^{p'}} \leq C \|\mu_0\|_{L_x^{p'}}$ as well.

We claim that the drifts $f_t^n = b_t^n * \mu_t^n$ converge to $f_t := b_t * \mu_t$ in $L_T^q B_{\infty,\infty}^{\alpha-1}$. Once this is shown, by the initial observation the solutions X^n must converge to the unique solution X associated to (ξ, W, f) ; then it must hold $\mathcal{L}(X_t) = \mu_t$, $f_t = b_t * \mathcal{L}(X_t)$ and so we can conclude that X is a solution to (5.1) with the desired regularity.

It remains to show the claim; to this end, we set

$$f_t^n - f_t = b_t^n * (\mu_t^n - \mu_t) + (b_t^n - b_t) * \mu_t =: g_t^n + h_t^n.$$

By Corollary A.5 in Appendix A, $g_t^n \rightarrow 0$ in $B_{\infty,\infty}^{\tilde{\alpha}}$ for all $\tilde{\alpha} < \alpha$ and a.e. $t \in [0, T]$; the bound $|g_t^n| \leq \|b_t\|_{B_{p,p}^\alpha} (\|\mu_t^n\|_{L_x^{p'}} + \|\mu_t\|_{L_x^{p'}})$ and dominated convergence imply that $g^n \rightarrow g$ in $L_T^q B_{\infty,\infty}^{\tilde{\alpha}}$ for all $\tilde{\alpha} < \alpha$. For h^n we have the estimate

$$\lim_{n \rightarrow \infty} \|h^n\|_{L_T^q B_{\infty,\infty}^\alpha} \leq \sup_{t \in [0, T]} \|\mu_t\|_{L_x^{p'}} \lim_{n \rightarrow \infty} \|b^n - b\|_{L_T^q B_{p,p}^\alpha} = 0.$$

Hence we have shown the claim and thus the conclusion in this case.

The case $H > 1/2$. As in the proof of Proposition 5.4, in this regime $\mathcal{L}(X) \in L_T^\infty L_x^{p'}$ is not enough to deduce straightaway that $b * \mathcal{L}(X)$ satisfies Assumption 3.12; however up to technical details, the proof is almost the same as above.

Specifically, we can consider a sequence $\{b^n\}_n$ of smooth functions, uniformly bounded in $C_T^{\alpha,H}L_x^p \cap C_T^0B_{p,p}^\alpha$, with $\operatorname{div} b^n$ uniformly bounded in $L_T^1L_x^\infty$ and such that $b^n \rightarrow b$ in $L_T^qB_{p,p}^\alpha$ for any $q < \infty$. Then exploiting the a priori bound from [19, Proposition 4.3] and the argument from Proposition 5.4, one can derive uniform estimates for the solutions X^n associated to X and finally pass to the limit with the help of Corollary 3.17.

Alternatively, let us mention that the existence of a weak solution X satisfying $\mathcal{L}(X) \in L_T^\infty L_x^{p'}$ in this setting can be obtained by an application of [19, Proposition 4.4]. \square

Proof of Theorem 2.6. It is now an immediate consequence of Propositions 5.4 and 5.5. \square

5.2. Integrable kernels. We now restrict ourselves to the case $H \leq 1/2$ and drifts $b \in L_T^qL_x^p$; in this setting we can present a second route to establishing existence of a solution with sufficiently regular law, to which we can apply Proposition 5.4.

Before proceeding further, let us explain why it is reasonable to expect so. By the Besov embedding $L_t^p \hookrightarrow B_{\infty,\infty}^{-d/p}$, drifts $b \in L_T^qL_x^p$ satisfy Assumption 3.12 if and only if

$$\frac{1}{q} + \frac{Hd}{p} < \frac{1}{2} - H; \quad (5.5)$$

however, differently from the class $L_T^qB_{p,p}^\alpha$, for $b \in L_T^qL_x^p$ it is known after the works [38, 32] that Girsanov transform (and thus weak existence and uniqueness in law for associated SDEs) is available as soon as

$$\frac{1}{q} + \frac{Hd}{p} < \frac{1}{2}. \quad (5.6)$$

As already seen in Section 3.4, Girsanov transform allows to deduce information on the regularity of $\mathcal{L}(X_t)$, which in turn provides higher regularity of the effective drift b^μ for the convolutional DDSDE. In particular, we may hope that starting from $b \in L_T^qL_x^p$ for (q, p) satisfying (5.6), we end up with $b^\mu \in L_T^{\tilde{q}}L_x^{\tilde{p}}$ with (\tilde{q}, \tilde{p}) satisfying (5.5).

At a technical level, we will proceed similarly as in Section 5.1, first establishing uniform a priori estimates for regular b and then running an approximation procedure. We start by establishing the recalling and improving the available results on Girsanov transform; as we are only interested in smooth approximations, for simplicity we restrict to regular drifts.

Lemma 5.6. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, (ξ, W) a $\mathbb{R}^d \times C_T$ -valued r.v. on it with $\mathcal{L}_{\mathbb{P}}(\xi, W) = \mu_0 \otimes \mu^H$ for some $H \leq 1/2$ and let $f : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ be a globally Lipschitz drift, $f \in L_T^qL_x^p$ for parameters $(p, q) \in [1, \infty]^2$ satisfying (5.6); let X be the unique strong solution to*

$$X_t = \xi + \int_0^t f_s(X_s) ds + W_t \quad \forall t \in [0, T].$$

Then there exists a measure \mathbb{Q} equivalent to \mathbb{P} such that $\mathcal{L}_{\mathbb{Q}}(X) = \mathcal{L}_{\mathbb{P}}(\xi + W)$ and there exists an increasing function F , depending on H, T, p, q , such that

$$\mathbb{E}_{\mathbb{Q}} \left[\left(\frac{d\mathbb{P}}{d\mathbb{Q}} \right)^n \right] + \mathbb{E}_{\mathbb{Q}} \left[\left(\frac{d\mathbb{Q}}{d\mathbb{P}} \right)^n \right] \leq F(n, \|f\|_{L_T^qL_x^p}) < \infty \quad \forall n \in \mathbb{N}$$

where the estimate does not depend on μ_0 nor the specific function f .

Proof. For deterministic initial data $\xi = x_0 \in \mathbb{R}^d$ (equiv. $\mu_0 = \delta_{x_0}$), the statement is a direct consequence of [32, Lemma 6.7], where it is already stressed that the estimates only depend on $\|f\|_{L_T^qL_x^p}$ but not on x_0 nor the specific f . The proof for random initial data ξ independent of

W is now identical to that of Corollary 3.18; the estimate not depending on ξ follows from the property that $\|f(x_0 + \cdot)\|_{L_T^q L_x^p} = \|f\|_{L_T^q L_x^p}$ for all $x_0 \in \mathbb{R}^d$. \square

The next lemma shows that the initial regularity of μ_0 is propagated at positive times, establishing useful a priori estimates; the proof is similar to that of Proposition 3.21.

Lemma 5.7. *Let $\xi, W, X, f, (p, q)$ be as in Lemma 5.6 and assume $\mu_0 \in L_x^r$ for some $r \in (1, \infty)$; then*

$$\sup_{t \in [0, T]} \|\mathcal{L}_{\mathbb{P}}(X_t)\|_{L_x^{\tilde{r}}} < \infty \quad \forall \tilde{r} \in (1, r).$$

Proof. Fix $\tilde{r} < r$ and denote by \tilde{r}' the conjugate exponent of \tilde{r} ; take $\varepsilon > 0$ such that $r'(1+\varepsilon) = \tilde{r}'$. Let \mathbb{Q} be the measure given by Lemma 5.6 such that $\mathcal{L}_{\mathbb{Q}}(X) = \mathcal{L}_{\mathbb{P}}(\xi + W)$; since $d\mathbb{P}/d\mathbb{Q}$ admits moments of any order, for any $g \in C_c^\infty(\mathbb{R}^d)$, by Hölder

$$\begin{aligned} |\langle g, \mathcal{L}(X_t) \rangle| &\leq \mathbb{E}_{\mathbb{P}}[|g|(X_t)] = \mathbb{E}_{\mathbb{Q}}\left[|g|(X_t) \frac{d\mathbb{P}}{d\mathbb{Q}}\right] \\ &\leq \mathbb{E}_{\mathbb{Q}}[|g|^{1+\varepsilon}(X_t)]^{\frac{1}{1+\varepsilon}} \mathbb{E}_{\mathbb{Q}}\left[\left(\frac{d\mathbb{P}}{d\mathbb{Q}}\right)^{1+\frac{1}{\varepsilon}}\right]^{\frac{\varepsilon}{1+\varepsilon}} \\ &\lesssim_{\varepsilon} \mathbb{E}_{\mathbb{P}}[|g|^{1+\varepsilon}(\xi + W_t)]^{\frac{1}{1+\varepsilon}} \\ &= \langle |g|^{1+\varepsilon}, \mu_0 * \mathcal{L}_{\mathbb{P}}(W_t) \rangle^{\frac{1}{1+\varepsilon}} \end{aligned}$$

where in the last passage we used the fact that ξ and W_t are independent under \mathbb{P} . Recalling that $\mathcal{L}(W_t)$ is a probability measure, by Hölder's and then Young's inequality we arrive at

$$|\langle g, \mathcal{L}(X_t) \rangle| \lesssim_{\varepsilon} \| |g|^{1+\varepsilon} \|_{L_x^{r'(1+\varepsilon)}}^{\frac{1}{1+\varepsilon}} \|\rho * \mathcal{L}(W_t)\|_{L_x^r}^{\frac{1}{1+\varepsilon}} \lesssim \|g\|_{L_x^{r'(1+\varepsilon)}} \|\mu_0\|_{L_x^r}^{\frac{1}{1+\varepsilon}} = \|g\|_{L_x^{\tilde{r}'}} \|\mu_0\|_{L_x^r}^{\frac{1}{1+\varepsilon}}$$

As the estimate is uniform over all $g \in C_c^\infty(\mathbb{R}^d)$ and $t \in [0, T]$, by duality we deduce that

$$\sup_{t \in [0, T]} \|\mathcal{L}(X_t)\|_{L_x^{\tilde{r}}} \lesssim_{\varepsilon} \|\mu_0\|_{L_x^r}^{\frac{1}{1+\varepsilon}};$$

as the reasoning holds for all $\tilde{r} < r$, the conclusion follows. \square

We are now ready to prove the existence of solutions to the DDSDE (5.1) for $b \in L_T^q L_x^p$ and sufficiently integrable μ_0 .

Proposition 5.8. *Let $H \leq 1/2$, $(p, r, q) \in [1, \infty) \times (1, \infty) \times (2, \infty]$ such that*

$$\frac{1}{q} + Hd \left(\frac{1}{p} + \frac{1}{r} - 1 \right) < \frac{1}{2} - H, \quad \frac{1}{q} + \frac{Hd}{p} < \frac{1}{2}. \quad (5.7)$$

Then for any $b \in L_T^q L_x^p$ and any $\mu_0 \in L_x^r$ there exists a strong solution X to the associated DDSDE (5.1), which moreover satisfies $\mathcal{L}(X) \in L_T^\infty L_x^{\tilde{r}}$ for any $\tilde{r} \in [1, r)$.

Proof. We pursue the same general strategy as in the proof of Proposition 5.5.

As condition (5.7) only contains strict inequalities, w.l.o.g. we can assume $q < \infty$; consider a sequence $\{b^n\}_n$ of Lipschitz, compactly supported functions such that $b^n \rightarrow b$ in $L_T^q L_x^p$ and

$\|b^n\|_{L_T^q L_x^p} \leq \|b\|_{L_T^q L_x^p}$. It follows from [11, Theorem 7] that for every n there exists a unique solution X^n to the approximating DDSDE

$$X_t^n = \xi + \int_0^t (b_s^n * \mu_s^n)(X_s^n) ds + W_t, \quad \mu_t^n = \mathcal{L}(X_t^n).$$

In particular, each X^n is also a solution to an SDE with drift $f_t^n := b_t^n * \mathcal{L}(X_t^n)$ and by Young's inequality

$$\sup_n \|f^n\|_{L_T^q L_x^p} \leq \sup_n \|b^n\|_{L_T^q L_x^p} \leq \|b\|_{L_T^q L_x^p}.$$

Therefore we may apply Lemmas 5.6 and 5.7 to obtain the uniform bound

$$\sup_n \|\mu^n\|_{L_T^\infty L_x^{\tilde{r}}} < \infty$$

for all $\tilde{r} \in [1, r)$. Applying Hölder's inequality to the integral in time and Young's inequalities to the convolution in space, we find

$$\sup_n \|b^n * \mu^n\|_{L_T^q L_x^{\tilde{p}}} < \infty$$

for any $\tilde{p} < \bar{p}$, where

$$1 + \frac{1}{\tilde{p}} = \frac{1}{r} + \frac{1}{\bar{p}}.$$

Using the fact that \tilde{p} can be chosen arbitrarily close to \bar{p} and that the first inequality in (5.7) is strict, we see that the family $\{b^n * \mu^n\}$ is bounded in $L_T^q L_x^{\tilde{p}}$ for parameters (q, \tilde{p}) satisfying

$$\frac{1}{q} + \frac{Hd}{\tilde{p}} < \frac{1}{2} - H;$$

but this is exactly condition (5.5), i.e. the regularity regime in which we know how to solve the SDE in a strong sense. On the other hand, the uniform bound for $\|b^n\|_{L_T^q L_x^p}$ and the use of Girsanov transform allows to derive a uniform bound for $\mathbb{E}[\|X^n\|_{H-\varepsilon}]$ for any $\varepsilon > 0$; together with $X_0^n = \xi$ for all n this implies tightness of $\{X^n\}_n$, so that we can extract a (not relabelled) subsequence such that $\mathcal{L}(X^n) \rightharpoonup \mu$ in $\mathcal{P}(C_T)$, $\mu_t^n \rightharpoonup \mu_t = e_t \# \mu$ for all $t \in [0, T]$.

From here, the argument is almost identical to that of Proposition 5.5: once we show that $b^n * \mu^n \rightarrow b * \mu$ in a sufficiently strong topology, then by Corollary 3.17 the solutions X^n will converge to the unique strong solution X associated to $(b * \mu, \xi)$, which must therefore be a solution to the DDSDE associated to $\mu_0 = \mathcal{L}(\xi)$ and b . By the uniform bounds on $\{\mu^n\}_n$ and weak convergence $\mu_t^n \rightharpoonup \mu_t = \mathcal{L}(X_t)$, we also deduce that $\mathcal{L}(X) \in L_T^\infty L_x^{\tilde{r}}$ for all $\tilde{r} < r$.

Since by construction $b^n \rightarrow b$ in $L_T^q L_x^p$, $\mu_t^n \rightharpoonup \mu_t$ and $\{b^n * \mu^n\}_n$ is bounded in $L_T^q L_x^{\tilde{p}}$ for some (q, \tilde{p}) satisfying (5.5), we can apply Corollary A.6 from Appendix A to deduce that (up to further relabelling $\tilde{p} - \varepsilon$ into \tilde{p}) $b^n * \mu^n \rightarrow b * \mu$ in $L_T^q L_x^{\tilde{p}}$. In light of the embedding $L_T^q L_x^{\tilde{p}} \hookrightarrow L_T^q B_{\infty, \infty}^{-d/\tilde{p}}$ and Corollary 3.17, this implies the conclusion. \square

We are now ready to prove the main result of this subsection.

Theorem 5.9. *Let $H \leq 1/2$, $(p, r, q) \in [1, \infty) \times (1, \infty) \times (2, \infty]$ satisfy (5.7). Then for any $b \in L_T^q L_x^p$ and $\mu_0 \in L_x^r$, then there exists a strong solution X to (5.1), which satisfies $\mathcal{L}(X) \in L_T^\infty L_x^{\tilde{r}}$ for any $\tilde{r} < r$; pathwise uniqueness and uniqueness in law hold in the class of solutions satisfying this condition.*

Proof. The proof is based on a (non-trivial) combination of Propositions 5.4 and 5.8.

Under our assumptions, the existence of a strong solution such that $\mathcal{L}(X_\cdot) \in L_T^\infty L_x^{\tilde{r}}$ for any $\tilde{r} < r$ is granted; in particular if $r > p'$, then we can choose $\tilde{r} = p'$ and then assumptions of Proposition 5.5 in this case are satisfied thanks to the embedding $L_T^q L_x^p \hookrightarrow L_T^q B_{p,p}^{-\varepsilon}$ for any $\varepsilon > 0$, giving the uniqueness part of the statement. Up to technicalities the borderline case $r = p'$ can be treated similarly, exploiting the embedding $L_T^q L_x^p \hookrightarrow L_T^q B_{\tilde{p},\tilde{p}}^{-\varepsilon}$ for some $\tilde{p} = \tilde{p}(\varepsilon) > p$ chosen so that $1/\tilde{r} + 1/\tilde{p} = 1$.

Thus it remains to study the regime $r < p'$, equivalently $r' > p$; in this case we can choose $\tilde{r} < r$ such that $\tilde{r}' > p$ as well. By Besov embedding it then holds

$$L_T^q L_x^p \hookrightarrow B_{\tilde{r}',\tilde{r}'}^\alpha \quad \text{for } \alpha := -d\left(\frac{1}{p} - \frac{1}{\tilde{r}'}\right) = -d\left(\frac{1}{p} + \frac{1}{\tilde{r}} - 1\right);$$

to verify that b satisfies the assumptions of Proposition 5.4, it then suffices to check that

$$\frac{1}{q} + Hd\left(\frac{1}{p} + \frac{1}{\tilde{r}} - 1\right) < \frac{1}{2} - H;$$

since \tilde{r} can be taken arbitrarily close to r , this follows from the first strict inequality in (5.7). \square

Remark 5.10. For $r > d/(d-1)$, condition (5.6) implies condition (5.7). Therefore Theorem 2.7 is a particular subcase of Theorem 5.9.

APPENDIX A. SOME USEFUL LEMMAS

We collect in this appendix several technical lemmas of analytic nature that have been used throughout the paper. We start with useful facts on the identification of elements of dual spaces.

Lemma A.1. *Suppose $f \in L_{loc}^1([0, T] \times \mathbb{R}^d)$, $q \in [1, \infty)$ and there exists a constant C such that*

$$\int_{[0, T] \times \mathbb{R}^d} f(t, x) \varphi(t, x) \, dt \, dx \leq C \|\varphi\|_{L_T^q L_x^\infty}$$

for all compactly supported $\varphi \in L^\infty([0, T] \times \mathbb{R}^d)$. Then $f \in L_T^q L_x^1$ and $\|f\|_{L_T^q L_x^1} \leq C$.

Proof. Fix $M > 0$ and set

$$g_t := \int_{\mathbb{R}^d} |f(t, x)| \mathbb{1}_{|f(t, \cdot)| \leq M}(x) \mathbb{1}_{|x| \leq M} \, dx, \quad h(t, x) := g_t^{q-1} \operatorname{sgn}(f(t, x)) \mathbb{1}_{|f(t, \cdot)| \leq M}(x) \mathbb{1}_{|x| \leq M};$$

then by assumption

$$\int_{[0, T]} |g_t|^q \, dt = \int_{[0, T] \times \mathbb{R}^d} f(t, x) h(t, x) \, dt \, dx \leq C \|h\|_{L_t^q L_x^\infty} = C \left(\int_{[0, T]} |g_t|^q \, dt \right)^{1-1/q},$$

namely

$$\left(\int_{[0, T]} \left(\int_{\mathbb{R}^d} |f(t, x)| \mathbb{1}_{|f(t, \cdot)| \leq M}(x) \mathbb{1}_{|x| \leq M} \, dx \right)^q \, dt \right)^{1/q} \leq C.$$

Taking the limit $M \rightarrow \infty$ gives the conclusion. \square

Lemma A.2. *Suppose $f \in L^1_{loc}([0, T] \times \mathbb{R}^d)$, $q \in (1, \infty)$ and there exists a constant C such that*

$$\int_{[0, T] \times \mathbb{R}^d} f(t, x) \varphi(t, x) \, dt \, dx \leq C \|\varphi\|_{L^q_T L^1_x}$$

for all compactly supported $\varphi \in L^\infty([0, T] \times \mathbb{R}^d)$. Then $f \in L^q_T L^\infty_x$ and $\|f\|_{L^q_T L^\infty_x} \leq C$.

Proof. Fix $R > 0$ and set $D = \{x \in \mathbb{R}^d : |x| \leq R\}$; we can identify $L^p_D := L^p(D)$ as the subset of $L^p(\mathbb{R}^d)$ made of functions supported on D , similarly $L^q_T L^p_D$ as a subset of $L^q_T L^p_x$. Observe that for $p > 1$, $L^q_T L^p_D \hookrightarrow L^q_T L^1_D$ with $\|f\|_{L^q_T L^1_D} \leq (c_d R^d)^{1/p'} \|f\|_{L^q_T L^p_D}$. By hypothesis f defines an element of the dual of $L^q_T L^p_D$, thus by duality

$$\|f\|_{L^q_T L^{p'}_D} \leq C (c_d R^d)^{1/p'} \quad \forall p' \in (1, \infty).$$

Since D has finite measure, $\|g\|_{L^\infty_D} = \lim_{p \rightarrow \infty} \|g\|_{L^p_D}$ for all measurable g ; taking $p' \rightarrow \infty$ in the above by Fatou's lemma we obtain

$$\|f\|_{L^q_T L^\infty_D} \leq C;$$

as the reasoning holds for any $R > 0$, the conclusion follows talking $R \rightarrow \infty$. \square

The next statements concern the compactness properties of convolutions, specifically how weak convergence of measures is enhanced to strong convergence of associated functionals $b * \mu$.

Lemma A.3. *Let $b \in L^p_x$ with $p \in [1, \infty)$ and $\{\mu^n\}_n \subset \mathcal{P}(\mathbb{R}^d)$ such that $\mu^n \rightharpoonup \mu$ weakly, then*

$$\lim_{n \rightarrow \infty} \|b * \mu^n - b * \mu\|_{L^p_x} = 0.$$

If moreover $\{\mu^n\}_n$ is bounded in L^r_x with $r > 1$, then $b * \mu^n \rightarrow b * \mu$ in $L^{\tilde{p}}_x$ for any $\tilde{p} \in [p, \infty)$ s.t.

$$1 + \frac{1}{\tilde{p}} > \frac{1}{p} + \frac{1}{r}. \quad (\text{A.1})$$

Proof. Given $h \in \mathbb{R}^d$, define the translation operator $\tau_h : f \mapsto f(\cdot + h)$ acting on L^p_x . Recall that any given $b \in L^p_x$ is equicontinuous, in the sense that $\tau_{h^n} b \rightarrow \tau_h b$ in L^p_x for $h^n \rightarrow h$ in \mathbb{R}^d .

Since $\mu^n \rightharpoonup \mu$, by Skorokhod's representation theorem we can construct a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a family of r.v.s $\{X^n\}_n, X$ on it such that $\mathcal{L}_{\mathbb{P}}(X^n) = \mu^n$, $\mathcal{L}_{\mathbb{P}}(X) = \mu$ and $X^n \rightarrow X$ \mathbb{P} -a.s.; it then holds

$$\begin{aligned} \|b * \mu^n - b * \mu\|_{L^p_x}^p &= \int_{\mathbb{R}^d} |\mathbb{E}[b(x - X^n) - b(x - X)]|^p \, dx \\ &\leq \mathbb{E} \left[\int_{\mathbb{R}^d} |b(x - X^n) - b(x - X)|^p \, dx \right] = \mathbb{E} \left[\|\tau_{X^n} b - \tau_X b\|_{L^p_x}^p \right] \end{aligned}$$

where in the second passage we used Jensen's inequality. By the aforementioned equicontinuity it holds $\|\tau_{X^n} b - \tau_X b\|_{L^p_x} \rightarrow 0$ \mathbb{P} -a.s. and we have the uniform bound $\|\tau_{X^n} b - \tau_X b\|_{L^p_x} \leq 2\|b\|_{L^p_x}$, thus the first claim follows from dominated convergence.

Regarding the second claim, Young's inequality gives a uniform bound for $\{b * \mu^n\}$ in $L^{\tilde{p}}_x$ for $1 + 1/\tilde{p} = 1/p + 1/r$; combined with convergence in L^p_x and interpolation estimates, we deduce convergence in $L^{\tilde{p}}_x$ for any $\tilde{p} \in [p, \bar{p})$. \square

Remark A.4. In the borderline case of $b \in L_x^p$ and $\{\mu^n\}_n$ bounded in L_x^p , then $\{b * \mu^n\}_n$ is a bounded sequence in $C_0(\mathbb{R}^d)$, which denotes the Banach space of continuous functions vanishing at infinity, endowed with the supremum norm. In this case it can be shown that $\{b * \mu^n\}_n$ is also equicontinuous, so by Ascoli–Arzelà it converges to $b * \mu$ uniformly on compact sets.

Corollary A.5. *Let $b \in B_{p,p}^\alpha$ with $p \in [1, \infty)$, $\alpha \in \mathbb{R}$, and $\{\mu^n\}_n \subset \mathcal{P}(\mathbb{R}^d)$ such that $\mu^n \rightharpoonup \mu$ weakly, then $b * \mu^n \rightarrow b * \mu$ in $B_{p,p}^\alpha$; if moreover $\{\mu^n\}_n$ is bounded in L_x^r with $r > 1$, then $b * \mu^n \rightarrow b * \mu$ in $B_{\tilde{p},\tilde{p}}^\alpha$ for any $\tilde{p} \in [p, \infty)$ satisfying (A.1). Finally, if*

$$\frac{1}{p} + \frac{1}{r} < 1,$$

*then $b * \mu^n \rightarrow b * \mu$ in $B_{\infty,\infty}^{\tilde{\alpha}}$ for any $\tilde{\alpha} < \alpha$.*

Proof. Observe that $\Delta_j(b * \mu^n - b * \mu) = (\Delta_j b) * (\mu^n - \mu)$, so by Lemma A.3 $\|\Delta_j(b * \mu^n - b * \mu)\|_{L_x^p} \rightarrow 0$ for any fixed $j \in \mathbb{N}$. Moreover

$$\|(\Delta_j b) * (\mu^n - \mu)\|_{L_x^p}^p \lesssim_p \|\Delta_j b\|_{L_x^p}^p \quad \text{with} \quad \sum_j 2^{\alpha j p} \|\Delta_j b\|_{L_x^p}^p < \infty$$

therefore by dominated convergence

$$\lim_i \|b * \mu^n - b * \mu\|_{B_{p,p}^\alpha}^p = \lim_i \sum_j 2^{\alpha j p} \|(\Delta_j b) * (\mu^n - \mu)\|_{L_x^p}^p = 0.$$

The second statement follows as before by interpolation between convergence in $B_{p,p}^\alpha$ and boundedness in $B_{\tilde{p},\tilde{p}}^\alpha$ with \tilde{p} satisfying the equality in (A.1).

Finally, if $1 > 1/p + 1/r$, then condition (A.1) is satisfied for all \tilde{p} big enough; the conclusion then follows from convergence in $B_{\tilde{p},\tilde{p}}^\alpha$ and the Besov embedding $B_{\tilde{p},\tilde{p}}^\alpha \hookrightarrow B_{\infty,\infty}^{\alpha-d/\tilde{p}}$. \square

Corollary A.6. *Let $p, q \in [1, \infty)$, $\{b^n\}_n \subset L_T^q L_x^p$ be a sequence such that $b^n \rightarrow b$ in $L_T^q L_x^p$; moreover let $\{\mu^n, \mu\} \subset C_T \mathcal{P}(\mathbb{R}^d)$ be such that $\mu_t^n \rightharpoonup \mu_t$ weakly for every $t \in [0, T]$ and such that*

$$\sup_{n \in \mathbb{N}} \sup_{t \in [0, T]} \|\mu_t^n\|_{L_x^r} + \sup_{t \in [0, T]} \|\mu_t\|_{L_x^r} < \infty$$

*for some $r \in [1, \infty]$. Then $b^n * \mu^n \rightarrow b * \mu$ in $L_T^q L_x^{\tilde{p}}$ for all $\tilde{p} \in [p, \infty)$ satisfying (A.1).*

Proof. It suffices to show that $b^n * \mu^n \rightarrow b * \mu$ in $L_T^q L_x^{\tilde{p}}$; once this is done, convergence in $L_T^q L_x^{\tilde{p}}$ follows as usual by interpolation and boundedness in $L_T^q L_x^{\tilde{p}}$, which comes from the assumptions and Young’s inequality. It holds

$$\|b^n * \mu^n\|_{L_T^q L_x^{\tilde{p}}} \leq \|(b^n - b) * \mu^n\|_{L_T^q L_x^{\tilde{p}}} + \|b * \mu^n - b * \mu\|_{L_T^q L_x^{\tilde{p}}}$$

where we can estimate the first term by $\|(b^n - b) * \mu^n\|_{L_T^q L_x^{\tilde{p}}} \leq \|b^n - b\|_{L_T^q L_x^p} \rightarrow 0$. For the second term, by Lemma A.3 and the assumptions it holds $\|b_t * (\mu_t^n - \mu_t)\|_{L_x^{\tilde{p}}} \rightarrow 0$ for Lebesgue a.e. $t \in [0, T]$, as well as $\|b_t * (\mu_t^n - \mu_t)\|_{L_x^p} \leq 2\|b_t\|_{L_x^p}$; thus by dominated convergence we infer $\|b * \mu^n - b * \mu\|_{L_T^q L_x^{\tilde{p}}} \rightarrow 0$ as well. \square

Lemma A.7. *For any $p \in [1, \infty)$ and $\alpha \in \mathbb{R}$ there exists a constant $C = C(\alpha)$ such that*

$$\|b * (\mu - \nu)\|_{B_{\infty,\infty}^{\alpha-1}} \leq C \|b\|_{B_{\infty,\infty}^\alpha} d_p(\mu, \nu)$$

for all $b \in B_{\infty,\infty}^\alpha$ and $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$.

Proof. It's enough give the proof for $p = 1$, as the general case follows from $d_1(\mu, \nu) \leq d_p(\mu, \nu)$; we can assume $d_1(\mu, \nu) < \infty$, otherwise the inequality is trivial. By Bernstein estimates, reasoning on Littlewood-Paley blocks, we have

$$\begin{aligned} \|b * (\mu - \nu)\|_{B_{\infty, \infty}^{\alpha-1}} &= \sup_n \{2^{n(\alpha-1)} \|(\Delta_n b) * (\mu - \nu)\|_{L_x^\infty}\} \\ &\leq \sup_n \{2^{n(\alpha-1)} \|\Delta_n b\|_{W_x^{1, \infty}}\} d_1(\mu, \nu) \\ &\lesssim \sup_n \{2^{n\alpha} \|\Delta_n b\|_{L_x^\infty}\} d_1(\mu, \nu) = \|b\|_{B_{\infty, \infty}^\alpha} d_1(\mu, \nu) \end{aligned}$$

which gives the claim. \square

The last statements we are going to provide concern the continuity of the map $\mu \mapsto b * \mu$ in suitable topologies. Their proof require the use of maximal functions and their basic properties, which we recall first; we refer the interested reader to [43] for their proofs.

Given $b \in L^p(\mathbb{R}^d)$, $p \in [1, \infty]$, its maximal function Mb is defined by

$$Mb(x) := \sup_{r>0} \frac{1}{\lambda_d r^d} \int_{B(x,r)} |b(y)| dy$$

where λ_d stands for the Lebesgue measure of $B(0, 1)$ in \mathbb{R}^d . It is well known that if $p \in (1, \infty]$, then $Mf \in L^p(\mathbb{R}^d)$ and

$$\|Mb\|_p \leq c_{d,p} \|b\|_p$$

for some constant $c_{d,p} > 0$; similar definitions and properties hold in the case of vector-valued drifts $b \in L^p(\mathbb{R}^d; \mathbb{R}^m)$ (in which case $c = c_{d,p,m}$).

If $b \in W^{1,p}(\mathbb{R}^d; \mathbb{R}^d)$, then there exists a Lebesgue-negligible set $N \subset \mathbb{R}^d$ and a constant $c_d > 0$ such that the Hajlasz inequality holds:

$$|b(x) - b(y)| \leq c_d |x - y| (MDb(x) + MDb(y)) \quad \forall x, y \in \mathbb{R}^d \setminus N. \quad (\text{A.2})$$

The above and similar inequalities (see [8, 7] for recent asymmetric extensions) allow to control the map $\mu \mapsto b * \mu$ in Wasserstein spaces.

Lemma A.8. *Let $(p, q, r, s) \in (1, \infty)^4$ be such that $r \leq p \wedge s$ and*

$$1 + \frac{r}{s} \leq \frac{r}{p} + \frac{1}{q}. \quad (\text{A.3})$$

Then there exists a constant C , depending on d and the above parameters, such that for any $b \in W_x^{1,p}$ and $\mu, \nu \in L_x^q$ it holds

$$\|b * (\mu - \nu)\|_{L_x^s} \leq C \|b\|_{W_x^{1,p}} (\|\mu\|_{L_x^q}^{1/r} + \|\nu\|_{L_x^q}^{1/r}) d_{r'}(\mu, \nu)$$

Proof. If $d_{r'}(\mu, \nu) = +\infty$ the inequality is trivially true, so we can assume $d_{r'}(\mu, \nu) < \infty$; moreover since $\mu, \nu \in L_x^{\tilde{q}}$ for any $\tilde{q} \in [1, q]$, w.l.o.g. we can assume equality holds in (A.3).

Let $m \in \Pi(\mu, \nu)$ be an optimal coupling of (μ, ν) for $d_{r'}(\mu, \nu)$ and let $N \subset \mathbb{R}^d$ be as in (A.2); since μ, ν are absolutely continuous w.r.t. Lebesgue, it holds that $m((N \times \mathbb{R}^d) \cup (\mathbb{R}^d \times N)) = 0$.

Therefore we can apply (A.2) for any fixed $x \in \mathbb{R}^d$ to find

$$\begin{aligned}
 |b * (\mu - \nu)(x)| &= \left| \int_{\mathbb{R}^{2d}} [b(x-y) - b(x-z)] m(dy, dz) \right| \\
 &\lesssim \int_{\mathbb{R}^{2d}} |y-z| (MDb(x-y) + MDb(x-z)) m(dy, dz) \\
 &\leq \left(\int_{\mathbb{R}^{2d}} |y-z|^{r'} m(dy, dz) \right)^{1/r'} \left(\int_{\mathbb{R}^{2d}} |MDb(x-y) + MDb(x-z)|^r m(dy, dz) \right)^{1/r} \\
 &\leq d_{r'}(\mu, \nu) \left[\left(\int_{\mathbb{R}^d} |MDb(x-y)|^r \mu(dy) \right)^{1/r} + \left(\int_{\mathbb{R}^d} |MDb(x-z)|^r \nu(dz) \right)^{1/r} \right].
 \end{aligned}$$

Observe that

$$\left(\int_{\mathbb{R}^d} |MDb(\cdot - y)|^r \mu(dy) \right)^{1/r} = (|MDb|^r * \mu)^{1/r}(\cdot)$$

and by assumption (A.3) and Young's inequality it holds

$$\|(|MDb|^r * \mu)^{1/r}\|_{L_x^s} = \| |MDb|^r * \mu \|_{L_x^{s/r}}^{1/r} \lesssim \| |MDb|^r \|_{L_x^{p/r}}^{1/r} \| \mu \|_{L_x^q}^{1/r} = \| MDb \|_{L_x^p} \| \mu \|_{L_x^q}^{1/r}.$$

Together with a similar estimate for $|MDb| * \nu$ and the property $\|MDb\|_{L_x^p} \lesssim \|b\|_{W_x^{1,p}}$, combining everything we arrive at

$$\|b * (\mu - \nu)\|_{L_x^s} \lesssim \|b\|_{W_x^{1,p}} (\| \mu \|_{L_x^q}^{1/r} + \| \nu \|_{L_x^q}^{1/r}) d_{r'}(\mu, \nu)$$

which is the claim. \square

Corollary A.9. *Let $\alpha \in \mathbb{R}$, (p, q, r, s) as in Lemma A.8. Then there exists a constant C , depending on d and the above parameters, such that for any $b \in B_{p,p}^\alpha$ and $\mu, \nu \in L_x^q$*

$$\|b * (\mu - \nu)\|_{B_{s,s}^{\alpha-1}} \leq C \|b\|_{B_{p,p}^\alpha} (\| \mu \|_{L_x^q}^{1/r} + \| \nu \|_{L_x^q}^{1/r}) d_{r'}(\mu, \nu)$$

Proof. Let $\Delta_j b$ denote the Littlewood–Paley blocks of b , then $\Delta_j(b * (\mu - \nu)) = (\Delta_j b) * (\mu - \nu)$. By the previous lemma and Bernstein estimates, one has that

$$\begin{aligned}
 \|(\Delta_j b) * (\mu - \nu)\|_{L_x^s} &\lesssim \|\Delta_j b\|_{W_x^{1,p}} (\| \mu \|_{L_x^q}^{1/r} + \| \nu \|_{L_x^q}^{1/r}) d_{r'}(\mu, \nu) \\
 &\lesssim 2^j \|\Delta_j b\|_p (\| \mu \|_{L_x^q}^{1/r} + \| \nu \|_{L_x^q}^{1/r}) d_{r'}(\mu, \nu);
 \end{aligned}$$

moreover assumption (A.3) implies $s \geq p$. Therefore we have

$$\begin{aligned}
 \|b * (\mu - \nu)\|_{B_{s,s}^\alpha} &= \left(\sum_j \left[2^{(\alpha-1)j} \|(\Delta_j b) * (\mu - \nu)\|_{L_x^s} \right]^s \right)^{1/s} \\
 &\lesssim \left(\sum_j \left[2^{\alpha j} \|\Delta_j b\|_{L_x^s} \right]^s \right)^{1/s} (\| \mu \|_{L_x^q}^{1/r} + \| \nu \|_{L_x^q}^{1/r}) d_{r'}(\mu, \nu) \\
 &\lesssim \left(\sum_j \left[2^{\alpha j} \|\Delta_j b\|_{L_x^s} \right]^p \right)^{1/p} (\| \mu \|_{L_x^q}^{1/r} + \| \nu \|_{L_x^q}^{1/r}) d_{r'}(\mu, \nu)
 \end{aligned}$$

which gives the conclusion. \square

REFERENCES

- [1] Oussama Amine, David Baños, and Frank Proske. C-infinity-regularization by noise of singular ODEs. *arXiv preprint arXiv:1710.05760*, 2017.
- [2] Hajer Bahouri, Jean-Yves Chemin, and Raphaël Danchin. *Fourier analysis and nonlinear partial differential equations*, volume 343 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer, Heidelberg, 2011.
- [3] David Baños, Torstein Nilssen, and Frank Proske. Strong existence and higher order Fréchet differentiability of stochastic flows of fractional Brownian motion driven SDEs with singular drift. *Journal of Dynamics and Differential Equations*, pages 1–48, 2019.
- [4] Martin Bauer and Thilo Meyer-Brandis. McKean–Vlasov equations on infinite-dimensional Hilbert spaces with irregular drift and additive fractional noise. *arXiv preprint arXiv:1912.07427*, 2019.
- [5] Martin Bauer, Thilo Meyer-Brandis, and Frank Proske. Strong solutions of mean-field stochastic differential equations with irregular drift. *Electron. J. Probab.*, 23:Paper No. 132, 35, 2018.
- [6] Didier Bresch, Pierre-Emmanuel Jabin, and Zhenfu Wang. On mean-field limits and quantitative estimates with a large class of singular kernels: application to the Patlak-Keller-Segel model. *C. R. Math. Acad. Sci. Paris*, 357(9):708–720, 2019.
- [7] Elia Brué, Maria Colombo, and Camillo De Lellis. Positive solutions of transport equations and classical nonuniqueness of characteristic curves. *Arch. Ration. Mech. Anal.*, 240(2):1055–1090, 2021.
- [8] Laura Caravenna and Gianluca Crippa. A directional Lipschitz extension lemma, with applications to uniqueness and Lagrangianity for the continuity equation. *Comm. Partial Differential Equations*, 46(8):1488–1520, 2021.
- [9] Rémi Catellier and Massimiliano Gubinelli. Averaging along irregular curves and regularisation of ODEs. *Stochastic Process. Appl.*, 126(8):2323–2366, 2016.
- [10] Paul-Eric Chaudru de Raynal. Strong well posedness of McKean-Vlasov stochastic differential equations with Hölder drift. *Stochastic Process. Appl.*, 130(1):79–107, 2020.
- [11] Michele Coghi, Jean-Dominique Deuschel, Peter K. Friz, and Mario Aureli. Pathwise McKean-Vlasov theory with additive noise. *Ann. Appl. Probab.*, 30(5):2355–2392, 2020.
- [12] Ronald J. DiPerna and Pierre-Louis Lions. Ordinary differential equations, transport theory and Sobolev spaces. *Inventiones mathematicae*, 98(3):511–547, 1989.
- [13] Franco Flandoli. *Random Perturbation of PDEs and Fluid Dynamic Models: École d’été de Probabilités de Saint-Flour XL–2010*, volume 2015. Springer Science & Business Media, 2011.
- [14] Nicolas Fournier, Maxime Hauray, and Stéphane Mischler. Propagation of chaos for the 2D viscous vortex model. *J. Eur. Math. Soc. (JEMS)*, 16(7):1423–1466, 2014.
- [15] Lucio Galeati. Nonlinear Young differential equations: a review. *J. Dyn. Diff. Equat.*, available at <https://doi.org/10.1007/s10884-021-09952-w>, 2021.
- [16] Lucio Galeati and Massimiliano Gubinelli. Prevalence of ρ -irregularity and related properties. *arXiv preprint arXiv:2004.00872*, 2020.
- [17] Lucio Galeati and Massimiliano Gubinelli. Noiseless regularisation by noise. *Revista Matemática Iberoamericana*, 38(2):433–502, 2022.
- [18] Lucio Galeati and Fabian A. Harang. Regularization of multiplicative SDEs through additive noise. *to appear on Ann. Appl. Prob.*, 2020.
- [19] Lucio Galeati, Fabian A. Harang, and Avi Mayorcas. Distribution dependent SDEs driven by additive continuous noise. *Electron. J. Probab.*, 27:Paper No. 37, 38, 2022.
- [20] Jürgen Gärtner. On the McKean–Vlasov limit for interacting diffusions. *Math. Nachr.*, 137:197–248, 1988.
- [21] Yi Han. Solving McKean–Vlasov SDEs via relative entropy. *arXiv preprint arXiv:2204.05709*, 2022.
- [22] Fabian Harang and Avi Mayorcas. Pathwise regularisation of singular interacting particle systems and their mean field limits. *arXiv preprint arXiv:2010.15517*, 2020.
- [23] Fabian A. Harang and Nicolas Perkowski. C^∞ -regularization of ODEs perturbed by noise. *Stochastics and Dynamics*, 2021.
- [24] Jasper Hoeksema, Thomas Holding, Mario Aureli, and Oliver Tse. Large deviations for singularly interacting diffusions. *arXiv preprint arXiv:2002.01295*, 2020.
- [25] Xing Huang and Feng-Yu Wang. Distribution dependent SDEs with singular coefficients. *Stochastic Processes and their Applications*, 129(11):4747–4770, 2019.

- [26] Xing Huang and Feng-Yu Wang. McKean-Vlasov SDEs with drifts discontinuous under Wasserstein distance. *Discrete Contin. Dyn. Syst.*, 41(4):1667–1679, 2021.
- [27] Pierre-Emmanuel Jabin and Zhenfu Wang. Mean field limit for stochastic particle systems. In *Active particles. Vol. 1. Advances in theory, models, and applications*, Model. Simul. Sci. Eng. Technol., pages 379–402. Birkhäuser/Springer, Cham, 2017.
- [28] Pierre-Emmanuel Jabin and Zhenfu Wang. Quantitative estimates of propagation of chaos for stochastic systems with $W^{-1,\infty}$ kernels. *Invent. Math.*, 214(1):523–591, 2018.
- [29] Jean-Francois Jabir. Rate of propagation of chaos for diffusive stochastic particle systems via Girsanov transformation. *arXiv preprint arXiv:1907.09096*, 2019.
- [30] Daniel Lacker. On a strong form of propagation of chaos for McKean-Vlasov equations. *Electron. Commun. Probab.*, 23:Paper No. 45, 11, 2018.
- [31] Jean-Michel Lasry and Pierre-Louis Lions. Mean field games. *Jpn. J. Math.*, 2(1):229–260, 2007.
- [32] Khoa Lê. A stochastic sewing lemma and applications. *Electronic Journal of Probability*, 25, 2020.
- [33] Giovanni Leoni. *A first course in Sobolev spaces*, volume 181 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, second edition, 2017.
- [34] Henry P. McKean, Jr. A class of Markov processes associated with nonlinear parabolic equations. *Proc. Nat. Acad. Sci. U.S.A.*, 56:1907–1911, 1966.
- [35] Yuliya S. Mishura and Alexander Yu Veretennikov. Existence and uniqueness theorems for solutions of McKean–Vlasov stochastic equations. *arXiv preprint arXiv:1603.02212*, 2016.
- [36] David Nualart. *The Malliavin calculus and related topics*, volume 1995. Springer, 2006.
- [37] David Nualart and Youssef Ouknine. Regularization of differential equations by fractional noise. *Stochastic Processes and their Applications*, 102(1):103–116, 2002.
- [38] David Nualart and Youssef Ouknine. Stochastic differential equations with additive fractional noise and locally unbounded drift. In Evariste Giné, Christian Houdré, and David Nualart, editors, *Stochastic Inequalities and Applications*, pages 353–365, Basel, 2003. Birkhäuser Basel.
- [39] Jean Picard. Representation formulae for the fractional Brownian motion. In *Séminaire de Probabilités XLIII*, pages 3–70. Springer, 2011.
- [40] Michael Röckner and Xicheng Zhang. Well-posedness of distribution dependent SDEs with singular drifts. *Bernoulli*, 27(2):1131–1158, 2021.
- [41] Sylvia Serfaty. Mean field limits of the Gross-Pitaevskii and parabolic Ginzburg-Landau equations. *J. Amer. Math. Soc.*, 30(3):713–768, 2017.
- [42] Sylvia Serfaty. Mean field limit for Coulomb-type flows. *Duke Math. J.*, 169(15):2887–2935, 2020. With an appendix by Mitia Duerinckx and Serfaty.
- [43] Elias M. Stein. *Singular integrals and differentiability properties of functions*, volume 2 of *Princeton Mathematical Series, No. 30*. Princeton University Press, Princeton, N.J., 1970.
- [44] Alain-Sol Sznitman. Topics in propagation of chaos. In *École d’Été de Probabilités de Saint-Flour XIX—1989*, volume 1464 of *Lecture Notes in Math.*, pages 165–251. Springer, Berlin, 1991.
- [45] Hiroshi Tanaka. Limit theorems for certain diffusion processes with interaction. In *Stochastic analysis (Katata/Kyoto, 1982)*, volume 32 of *North-Holland Math. Library*, pages 469–488. North-Holland, Amsterdam, 1984.
- [46] Milica Tomasevic. Propagation of chaos for stochastic particle systems with singular mean-field interaction of $L^q - L^p$ type. *HAL-03086253 available at https://hal.inria.fr/hal-03086253*, 2020.
- [47] Alexander Yu. Veretennikov. On strong solution and explicit formulas for solutions of stochastic integral equations. *Math. USSR Sb.*, 39:387–403, 1981.
- [48] Cédric Villani. *Optimal transport: old and new*, volume 338. Springer Science & Business Media, 2008.
- [49] A. K. Zvonkin. A transformation of the phase space of a diffusion process that will remove the drift. *Mat. Sb. (N.S.)*, 93(135):129–149, 152, 1974.

LUCIO GALEATI: INSTITUTE OF APPLIED MATHEMATICS, UNIVERSITY OF BONN, 53115 ENDENICHER ALLEE 60, BONN, GERMANY.

Email address: lucio.galeati@iam.uni-bonn.de

FABIAN A. HARANG: DEPARTMENT OF MATHEMATICS, UNIVERSITY OF OSLO, P.O. BOX 1053, BLINDERN, 0316, OSLO, NORWAY, AND DEPARTMENT OF ECONOMICS, BI NORWEGIAN BUSINESS SCHOOL, NYDALSVEIEN 37, OSLO, NORWAY.

Email address: `fabian.a.harang@bi.no`

AVI J. MAYORCAS: DPMMS, UNIVERSITY OF CAMBRIDGE, CMS, WILBERFORCE ROAD, CB3 0WB, UK.

Email address: `am3015@cam.ac.uk`