Stochastic ODEs and stochastic linear PDEs with critical drift: regularity, duality and uniqueness

Lisa Beck[1], Franco Flandoli[2], Massimiliano Gubinelli[3], Mario Maurelli[4]

Abstract

In this paper linear stochastic transport and continuity equations with drift in critical L^p spaces are considered. In this situation noise prevents shocks for the transport equation and singularities in the density for the continuity equation, starting from smooth initial conditions. Specifically, we first prove a result of Sobolev regularity of solutions, which is false for the corresponding deterministic equation. The technique needed to reach the critical case is new and based on parabolic equations satisfied by moments of first derivatives of the solution, opposite to previous works based on stochastic flows. The approach extends to higher order derivatives under more regularity of the drift term. By a duality approach, these regularity results are then applied to prove uniqueness of weak solutions to linear stochastic continuity and transport equations and certain well-posedness results for the associated stochastic differential equation (sDE) (roughly speaking, existence and uniqueness of flows and their C^α regularity, strong uniqueness for the sDE when the initial datum has diffuse law). Finally, we show two types of examples: on the one hand, we present well-posed sDEs, when the corresponding ODEs are ill-posed, and on the other hand, we give a counterexample in the supercritical case.

MSC (2010): 60H10, 60H15 (primary); 35A02, 35B65 (secondary)

1 Introduction

Let $b\colon [0,T]\times\mathbb{R}^d\to\mathbb{R}^d$, for $d\in\mathbb{N}$, be a deterministic, time-dependent vector field, that we call drift. Let $(W_t)_{t\geq 0}$ be a Brownian motion in $\mathbb{R}^d$, defined on a probability space $(\Omega,\mathcal{A},P)$ with respect to a filtration $(\mathcal{G}_t)_{t\geq 0}$ and let σ be a real number. The following three stochastic equations are (at least formally) related:

1. the stochastic differential equation (*sDE*)

$$dX = b(t,X)dt + \sigma dW_t, \qquad X_0 = x, \tag{sDE}$$

where $x\in\mathbb{R}^d$; the unknown $(X_t)_{t\in[0,T]}$ is a stochastic process in $\mathbb{R}^d$;

2. the stochastic transport equation (*sTE*)

$$du + b\cdot\nabla u dt + \sigma\nabla u\circ dW_t = 0, \qquad u|_{t=0} = u_0, \tag{sTE}$$

where $u_0\colon\mathbb{R}^d\to\mathbb{R}$, $b\cdot\nabla u = \sum_{i=1}^d b_i\partial_{x_i}u$, $\nabla u\circ dW_t = \sum_{i=1}^d \partial_{x_i}u\circ dW_t^i$, and Stratonovich multiplication is used (precise definitions will be given below); the unknown $(u(t,x))_{t\in[0,T],x\in\mathbb{R}^d}$ is a scalar random field,

3. the stochastic continuity equation (*sCE*)

$$d\mu + \operatorname{div}(b\mu)dt + \sigma\operatorname{div}(\mu\circ dW_t) = 0, \qquad \mu|_{t=0} = \mu_0, \tag{sCE}$$

[1]Institut für Mathematik, Universität Augsburg, Universitätsstrasse 14, 86159 Augsburg, Germany. E-mail address: lisa.beck@math.uni-augsburg.de.

[2]Scuola Normale Superiore, Piazza dei Cavalieri 7, 56126 Pisa, Italy. E-mail address franco.flandoli@sns.it.

[3]Hausdorff Center for Mathematics and Institute for Applied Mathematics, Universität, Bonn, Germany. E-mail address: gubinelli@iam.uni-bonn.de

[4]Dipartimento di Matematica 'Federigo Enriques', Università degli Studi di Milano, via Saldini 50, 20133 Milano, Italy. E-mail address: mario.maurelli@unimi.it

where μ_0 is a measure, $\mathrm{div}(\mu \circ dW_t)$ stands for $\sum_{i=1}^d \partial_{x_i}\mu \circ dW_t^i$, the unknown $(\mu_t)_{t\in[0,T]}$ is a family of random measures on $\mathbb{R}^d$, and thus the differential operations have to be understood in the sense of distributions.

The aim of this paper is to investigate several questions for these equations in the case when the drift is in a subcritical or even critical space, a case not reached by any approach until now.

1.1 The deterministic case $\sigma = 0$

For comparison with the results for the stochastic equations presented later on (due to the presence of noise), we first address the deterministic case $\sigma = 0$. We start by explaining the link between the three equations and recall some classical results – in the positive and in the negative direction – under various regularity assumptions on the drift b. When b is smooth enough, then:

(i) the sDE generates a flow $\Phi_t(x)$ of diffeomorphisms;

(ii) the sTE is uniquely solvable in suitable spaces, and for the solution we have the representation formula $u(t,x) = u_0(\Phi_t^{-1}(x))$;

(iii) the sCE is uniquely solvable in suitable spaces, and the solution μ_t is the push forward of μ_0 under Φ_t, $\mu_t = (\Phi_t)_\sharp \mu_0$.

These links between the three equations can be either proved a posteriori, after the equations have been solved by their own arguments, or they can be used to solve one equation by means of the other.

Well-posedness of the previous equations and links between them have been explored also when b is less regular. To simplify the exposition, let us summarize with the term "weakly differentiable" the classes of non-smooth b considered in [29, 2]. In these works it has been proved that, whenever b is weakly differentiable, sTE and sCE are well-posed in classes of weak solutions; moreover, a generalized or Lagrangian flow for the sDE exists. Remarkable is the fact that the flow is obtained by a preliminary solution of the sTE or of the sCE, see [29, 2] (later on in [23], similar results have been obtained directly on sDE). However, when the regularity of b is too poor, several problems arise, for which, at the level of the sDE and its flow, we want to mention two types:

1) non-uniqueness for the sDE, and, more generally, presence of discontinuities in the flow;

2) non-injectivity of the flow (two trajectories can coalesce) and, more generally, mass concentration.

These phenomena have counterparts at the level of the associated sCE and sTE:

1) non-uniqueness for the sDE leads to non-uniqueness for sCE and sTE;

2) non-injectivity of the flow leads to shocks in the sTE (i.e. absence of continuous solutions, even starting from a continuous initial datum), while mass concentration means that a measure-valued solution of the sCE does not remain distributed.

Elementary examples can be easily constructed by means of continuous drifts in dimension 1; more sophisticated examples in higher dimension, with bounded measurable and divergence free drift, can be found in [1]. Concerning regularity, let us briefly give some details for an easy example: Consider, in $d = 1$, the drift $b(x) := -\mathrm{sign}(x)|x|^\alpha$ for some $\alpha \in (0,1)$. All trajectories of the ODE coalesce at $x = 0$ in finite time; the solution to the deterministic TE develops a shock (discontinuity) in finite time, at $x = 0$, from every smooth initial condition u_0 such that $u_0(x) \neq u_0(-x)$ for some $x \neq 0$; the deterministic CE concentrates mass at $x = 0$ in finite time, if the initial mass is not zero. See also Section 7 for similar examples of drift terms leading to non-uniqueness or coalescence of trajectories for the deterministic ODE (which in turn results in non-uniqueness and discontinuities/mass concentration for the PDEs).

Notice that the outstanding results of [29, 2] (still in the deterministic case) are concerned only with uniqueness of weak solutions. The only results to our knowledge about regularity of solutions with rough drifts are those of [5, Section 3.3] relative to the loss of regularity of solutions

to the transport equation when the vector field satisfies a log-Lipschitz condition, which is a far better situation than those considered in this paper. We shall prove below that these phenomena disappear in the presence of noise. Of course they also disappear in the presence of viscosity, but random perturbations of transport type $\nabla u \circ dW_t$ and viscosity Δu are completely different mechanisms. The sTE remains an hyperbolic equation, in the sense that the solution follows the characteristics of the single particles (so we do not expect regularization of an irregular initial datum); on the contrary, the insertion of a viscous term corresponds to some average, making the equation of parabolic type. One could interpret transport noise as a turbulent motion of the medium where transport of a passive scalar takes place, see [18], which is different from a dissipative mechanism, although some of the consequences on the passive scalar may have similarities.

1.2 Stochastic case $\sigma \neq 0$

In the stochastic case, $\sigma \neq 0$, when b is smooth enough, the existence of a stochastic flow of diffeomorphisms Φ for the sDE, the well-posedness of sTE and the relation $u(t,x) = u_0(\Phi_t^{-1}(x))$ are again known results, see [55, 56, 57]; moreover, the link with sCE could be established as well. However, the stochastic case offers a new possibility, namely that due to nontrivial *regularization effects of the noise*, well-posedness of sDE, sTE and sCE remains true even if the drift b is quite poor, opposite to the deterministic case. Notice that we are not talking about the well-known regularization effect of a Laplacian or an expected value. By regularization we mean that some of the pathologies mentioned above about the deterministic case (non-uniqueness and blow-up) may disappear even at the level of a single trajectory ω; we do not address any regularization of solutions in time, i.e. that solutions become more regular than the initial conditions, a fact that is clearly false when we expect relations like $u(t,x) = u_0(\Phi_t^{-1}(x))$.

1.3 Aim of this paper

The aim of this work is to prove several results in this direction and develop a sort of comprehensive theory on this topic. The results in this paper are considerably advanced and are obtained by means of new powerful strategies, which give a more complete theory. The list of our main results is described in the next three subsections; in a few sentences, we are concerned with:

(i) regularity for the transport (and continuity) equation;

(ii) uniqueness for the continuity (and transport) equation;

(iii) uniqueness for the sDE and regularity for the flow.

In the following subsections, we will explain the results in more detail and give precise references to previous works on the topics. Moreover, we will also analyze the crucial regularity assumptions on the drift term (discussing its criticality in a heuristic way and via appropriate examples, which are either classical or elaborated at the end of the paper).

1.4 Regularity assumptions on b

As already highlighted before, the key point for the question of existence, uniqueness and regularity of the solutions to the relevant equations is the regularity assumption on the drift b. In particular, we will not work with any kind of differentiability or Hölder condition, but merely with an integrability condition. We say that a vector field $f\colon [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ satisfies the Ladyzhenskaya–Prodi–Serrin condition (LPS) with exponents $p, q \in (2, \infty)$ if

$$f \in L^q([0,T]; L^p(\mathbb{R}^d, \mathbb{R}^d)), \qquad \frac{d}{p} + \frac{2}{q} \leq 1.$$

We shall write $f \in \mathcal{LPS}(p,q)$ (the precise definition will be given in Section 2.1), and we use the norm

$$\|f\|_{L^q([0,T];L^p)} := \left(\int_0^T \left(\int_{\mathbb{R}^d} |f(t,x)|^p dx \right)^{q/p} dt \right)^{1/q}.$$

We may extend the definition to the limit case $(p,q) = (\infty, 2)$ in the natural way: we say that $f \in \mathcal{LPS}(\infty, 2)$ if $f \in L^2(0,T; L^\infty(\mathbb{R}^d, \mathbb{R}^d))$ and we use the norm

$$\|f\|^2_{L^2(0,T;L^\infty)} = \int_0^T \|f(t,\cdot)\|^2_\infty dt$$

with the usual meaning of $\|\cdot\|_\infty$ as the essential supremum norm. The extension to the other limit case $(p,q) = (d,\infty)$ is more critical (similarly to the theory of 3D Navier–Stokes equations, see below). The easy case is when $q = \infty$ is interpreted as continuity in time: $C([0,T]; L^d(\mathbb{R}^d, \mathbb{R}^d))$; on the contrary, $L^\infty(0,T; L^d(\mathbb{R}^d, \mathbb{R}^d))$ is too difficult in general and we shall impose an additional smallness assumption (which shall be understood implicitly whenever we address the case (d,∞) in this introduction).

Roughly speaking, our results will hold for a drift b which is the sum of a Lipschitz function of space (with some integrability in time) plus a vector field of LPS class. In the sequel of the introduction, for simplicity of the exposition, we shall not mention the Lipschitz component anymore, which is however important to avoid that the final results are restricted to drift with integrability (or at least boundedness) at infinity.

Let us note that when $p, q \in (2,\infty)$ the space $L^q([0,T]; L^p(\mathbb{R}^d, \mathbb{R}^d))$ is the closure in the topology $\|\cdot\|_{L^q([0,T];L^p)}$ of smooth functions with compact support. The same is true for $C([0,T]; L^d(\mathbb{R}^d, \mathbb{R}^d))$. In the limit cases $(p,q) = (\infty, 2)$ and $(p,q) = (d,\infty)$, using classical mollifiers, there exists a sequence of smooth functions with compact support which converges almost surely and has uniform bound in the corresponding norm. This fact will allow us to follow an approach of a priori estimates, i.e. perform all computations for solutions to the equation with smooth coefficients, obtain uniform estimates for the associated solutions, and then deduce the statement after passage to the limit.

We further want to comment on the significance of the LPS condition in fluid dynamics. The name LPS comes from the authors Ladyzhenskaya, Prodi and Serrin who identified this condition as a class where regularity and well-posedness of 3D Navier–Stokes equations hold, see [52, 59, 71, 73, 58] and [46]. The limit case $(p,q) = (d,\infty)$ generated a lot of research and can be treated almost as the other cases if there is continuity in time or some smallness condition, see for instance [10, 60], but the full $L^\infty(0,T; L^d(\mathbb{R}^d, \mathbb{R}^d))$ case is very difficult, see [32] and related works. It has been solved only recently, at the price of a very innovative and complex proof. A similar result for our problem is unknown. The deep connection of the LPS class, especially when $\frac{d}{p} + \frac{2}{q} = 1$, with the theory of 3D Navier–Stokes equations is one of our main motivations to analyze stochastic transport under such conditions.

We finally note that, while preparing the second version of this work (after the first version appeared on arXiv), one article [66] and two preprints [79, 64] have appeared on the topic of this paper. In the article [66] pathwise (but not path-by-path) uniqueness is shown for the sCE under Krylov–Röckner conditions in the subcritical case. The preprints [79] and [64] go almost up to the critical case for weak and strong solution to the SDEs, the latter showing also Sobolev regularity of the stochastic flow. Respectively, the former preprint, while staying within the subcritical case in the interior $[0,T)$, allows a singularity at time T which matches, and actually goes slightly beyond, the critical case. In the latter preprint [64], the limiting case is $d/p + 2/q = 1$ is reached when replacing the L^q integrability condition by a $L^{q,1}$ condition (for Lorenz space $L^{q,1} \subsetneq L^q$).

1.5 Criticality

We now show that the LPS condition is subcritical with strict inequality and critical with equality in the condition $d/p + 2/q \leq 1$. We have already emphasized that we treat the critical case because no other approach is known to attack this case, but the paper includes also the subcritical case. The general intuitive idea is that, near the singularities of b, the Gaussian velocity field is "stronger" than (or comparable to) b, which results in avoiding non-uniqueness or blow-up of solutions. The name "critical" comes from the following scaling argument (done only in a heuristic way since it only serves as motivation).

Let $u\colon [0,T] \times \mathbb{R}^d \to \mathbb{R}$ be a solution to the sTE. For $\lambda > 0$ and $\alpha \in \mathbb{R}$, we introduce the scaled function $u_\lambda\colon [0, T/\lambda^\alpha] \times \mathbb{R}^d \to \mathbb{R}$ defined as $u_\lambda(t,x) \coloneqq u(\lambda^\alpha t, \lambda x)$. We denote by $\partial_t u$ and ∇u the derivative of u in the first argument and the gradient in the second one (similarly for u_λ). Since

$\partial_t u_\lambda(t,x) = \lambda^\alpha \partial_t u(\lambda^\alpha t, \lambda x)$ and $\nabla u_\lambda(t,x) = \lambda \nabla u(\lambda^\alpha t, \lambda x)$, we get that u_λ satisfies formally

$$\partial_t u_\lambda(t,x) + b_\lambda(t,x) \cdot \nabla u_\lambda(t,x) + \lambda^{\alpha-1} \nabla u_\lambda(t,x) \circ W'(\lambda^\alpha t) = 0,$$

where $b_\lambda(t,x) = \lambda^{\alpha-1} b(\lambda^\alpha t, \lambda x)$ is the rescaled drift and $W'(\lambda^\alpha t)$ formally denotes the derivative of W at time $\lambda^\alpha t$. We now want to write the stochastic part in terms of a new Brownian motion. For this purpose, we define a process $(W_\lambda(t))_{t\geq 0}$, via $W_\lambda(t) \coloneqq \lambda^{-\alpha/2} W(\lambda^\alpha t)$ and notice that W_λ is a Brownian motion with $W'_\lambda(t) = \lambda^{\alpha/2} W'(\lambda^\alpha t)$. Thus the previous equation becomes

$$\partial_t u_\lambda(t,x) + b_\lambda(t,x) \cdot \nabla u_\lambda(t,x) + \lambda^{\alpha/2-1} \nabla u_\lambda(t,x) \circ W'_\lambda(t) = 0.$$

We first choose $\alpha = 2$ such that the stochastic part $\lambda^{\alpha/2-1} \nabla u_\lambda(t,x) \circ W'_\lambda(t)$ is comparable to the derivative in time $\partial_t u_\lambda$. Notice that this is the parabolic scaling, although sTE is not parabolic (but as we will see below, a basic idea of our approach is that certain expected values of the solution satisfy parabolic equations for which the above scaling is the relevant one). Next we require that, for small λ, the rescaled drift b_λ becomes small (or at least controlled) in some suitable norm (in our case, $L^q(0,T;L^p(\mathbb{R}^d,\mathbb{R}^d))$). It is easy to see that

$$\|b_\lambda\|_{L^q(0,T/\lambda^2;L^p)} = \lambda^{1-(2/q+d/p)} \|b\|_{L^q(0,T;L^p)}$$

(here, the exponent d comes from rescaling in space and the exponent 2 from rescaling in time and the choice $\alpha = 2$). In conclusion, we find that

- if LPS holds with strict inequality, then $\|b_\lambda\|_{L^q(0,T/\lambda^2;L^p)} \to 0$ as $\lambda \to 0$: the stochastic term dominates and we expect a regularizing effect (subcritical case);
- if LPS holds with equality, then $\|b_\lambda\|_{L^q(0,T/\lambda^\alpha;L^p)} = \|b\|_{L^q(0,T;L^p)}$ remains constant: the deterministic drift and the stochastic forcing are comparable (critical case).

This intuitively explains why the analysis of the critical case is more difficult. Notice that, if LPS does not hold, then we expect the drift to dominate, so that a general result for regularization by noise is probably false. In this sense, LPS condition should be regarded as an optimal condition for expecting regularity of solutions.

1.6 Regularity results for the SPDEs

Concerning regularity, we proceed in a unified approach to attack the sTE and the sCE simultaneously (but for the sCE we have to assume the LPS condition also on $\operatorname{div} b$). In fact, we shall treat a generalized stochastic equation of transport type which contains both the sTE and the sCE as special cases. For this equation we prove a regularity result which contains as a particular case the following:

Theorem 1.1. *Assume the LPS condition on b (and also on $\operatorname{div} b$ for the sCE). If u_0 is of class $\cap_{r\geq 1} W^{1,r}(\mathbb{R}^d)$, then there exists a solution to the sTE (similarly for the sCE) which is of class $\cap_{m\geq 1} W^{1,m}_{\mathrm{loc}}(\mathbb{R}^d)$.*

A more detailed statement is given in Section 2.9 below. This result is false for $\sigma = 0$, as we mentioned in Section 1.1. Referring to some of the pathologies which may happen in the deterministic case, we may say that, under regular initial conditions, noise prevents the emergence of *shocks* (discontinuities) for the sTE, and *singularities* of the density for the sCE (the mass at time t has a locally bounded density with respect to Lebesgue measure).

The method of proof is completely new. It is of analytic nature, based on PDE manipulations and estimates, opposite to the methods used before in [39, 33, 63] and which are based on a preliminary construction of the stochastic flow for the sDE. We believe that, apart from the result, this new method of proof is the first important technical achievement of this paper (see Section 2.2 for a detailed description of the central ingredients of our method).

We now want to give some details on the precise statements, the regularity assumptions on drift and the strategy of proof for some known regularity results for the sTE, for the purpose of comparison with the results presented here. The paper [39] deals with the case of *Hölder continuous*

bounded drift and is based on the construction of the stochastic flow from [40]. The paper [33] is concerned with the class called in the sequel as *Krylov-Röckner class*, after [54], where pathwise uniqueness and other results are proved for the sDE. We say that a vector field $f\colon [0,T]\times\mathbb{R}^d \to \mathbb{R}^d$ satisfies the Krylov-Röckner (KR) condition if the LPS condition holds with strict inequality

$$\frac{d}{p}+\frac{2}{q}<1$$

and we shall write $f \in \mathcal{KR}(p,q)$. The improvement from $\frac{d}{p}+\frac{2}{q}<1$ to $\frac{d}{p}+\frac{2}{q}=1$ appeared also in the theory of 3D Navier–Stokes equations and required new techniques (which in turn opened new research directions on $L^\infty(0,T;L^d(\mathbb{R}^d,\mathbb{R}^d))$ regularity). Also here it requires a completely new approach. Under the condition $\frac{d}{p}+\frac{2}{q}=1$, we do not know how to solve the sDE directly (see however the recent preprints [64, 79] mentioned above); even in a weak sense, by Girsanov theorem, the strict inequality seems to be needed ([54, 69, 51]). Similarly as in [39], the proof of regularity of solutions of the sTE from [33] is based on the construction of stochastic flows for the sDE and their regularity in terms of weak differentiability. Finally, [63] and [67] treat the case of *bounded measurable drift* and, in [67], fractional Brownian motion (the classical work under this condition on pathwise uniqueness for the sDE is [78]), again starting from a weak differentiability result for stochastic flows, proved however with methods different from [33].

Let us mention that proving that noise prevents blow-up (in cases where blow-up phenomena are possible in the deterministic situation) is an intriguing problem that is under investigation also for other equations, different from transport ones, see e g. [7, 21, 26, 28, 31, 42, 44, 50].

1.7 Uniqueness results for the sPDEs

The second issue of our work is uniqueness of weak solutions to equations of continuity (and transport) type. More precisely, we prove a *path-by-path uniqueness* result via a duality approach, which relies on the regularity results described in Section 1.6. When uniqueness is understood in a class of weak solutions, then the adjoint existence result must be in a class of sufficiently regular solutions (which is why the assumption for path-by-path uniqueness for the sCE will be the assumption for regularity for the sTE and vice versa); for this reason this approach cannot be applied in the deterministic case, when b is not sufficiently regular.

By path-by-path uniqueness we mean something stronger than pathwise uniqueness, namely that given ω a.s., the deterministic PDE corresponding to that particular ω has a unique weak solution (note that our sPDE can be reformulated as a random PDE, which then can be read in a proper sense at ω fixed). Instead, pathwise uniqueness means that two processes, hence families indexed by ω, both solutions of the equation, coincide for a.e. ω. We prove:

Theorem 1.2. *Assume the LPS condition on b (and also on $\operatorname{div} b$ for the sTE). Then the sCE (similarly the sTE) has path-by-path uniqueness of weak L^m-solutions, for every finite m.*

A more precise statement is given in Section 3.4 below. No other method is known to produce such a strong result of uniqueness. This duality method in the stochastic setting is the second important technical achievement of this paper.

The intuitive reason why, by duality, one can prove *path-by-path* uniqueness (usually so difficult to be proven) is the following one. The duality approach gives us an identity of the form

$$\langle \rho_t, \varphi\rangle = \left\langle \rho_0, u_0^{t,\varphi}\right\rangle \tag{1.1}$$

where ρ_t is any weak solution of the sCE (ρ_t is the density of μ_t) with initial condition ρ_0 and $(u_s^{t,\varphi})_{s\in[0,t]}$ is any regular solution of the sTE rewritten in backward form with final condition φ at time t. As we shall see below, we use an approximate version of (1.1), but the idea we want to explain here is the same. Identity (1.1) holds a.s. in Ω, for any given φ and t. But taking a dense (in a suitable topology) countable set $\mathcal{D}$ of φ's, we have (1.1) for a.e. ω, uniformly on $\mathcal{D}$ and thus we may identify ρ_t. This is the reason why this approach is so powerful to prove path-by-path results. Of course behind this simple idea, the main technical point is the regularity of the solutions to the sTE, which makes it possible to prove an identity of the form (1.1) for weak solutions ρ_t, for all those ω's such that $u_s^{t,\varphi}$ is regular enough.

Concerning other uniqueness results for the sTE with poor drift, let us briefly comment on a few of them. In [41] the case of Hölder continuous bounded drift is treated, by means of the differentiable flow associated to the sDE; [65] extends the result and the approach to drifts in KR class with zero divergence. The paper [16] extends the results to the sTE with Hölder continuous drift but driven by fractional Brownian motion, relying again on the flow; the technique used there for the analysis of the sDE itself is instead different from [41] and leads to path-by-path uniqueness. The paper [4] assumes weakly differentiable drift but relaxes the assumption on the divergence of the drift, with respect to the deterministic works [29, 2]. The papers [61, 36] use Wiener chaos expansion techniques to obtain uniqueness for the sTE for drifts close to KR class, see [61], or even beyond, see [36], at the price of uniqueness in a smaller class (namely among solutions adapted to the *Brownian* filtration). A full solution of the uniqueness problem in the KR class was still open (apart from the paper [66] and the recent preprints [64, 79] mentioned above) and this is a by-product of this paper, which solves the problem in a stronger sense in two directions:

i) path-by-path uniqueness instead of pathwise uniqueness;

ii) drift in the LPS class instead of only KR class.

Let us mention that the approach to uniqueness of [4] shares some technical steps with the results described in Section 1.6: renormalization of solutions (in the sense of [29]), Itô reformulation of the Stratonovich equation and then expected value (a Laplacian arises from this procedure). However, in [4] this approach has been applied directly to uniqueness of weak solutions so the renormalization step required weak differentiability of the drift. Instead, here we deal with regularity of solutions and thus the renormalization is applied to regular solutions of approximate problems and no additional assumption on the drift is needed.

Finally, we comment on some related uniqueness results in the nonlinear case. The duality technique has been used in [48], for scalar conservation laws with linear transport noise, and in [49], for nonlinear transport noise, but in a different way and without producing a path-by-path uniqueness result. Other results on uniqueness by noise are available with different techniques and/or different choices of noise, e.g. [11] for a dyadic model of turbulence and [6] for a parabolic model.

1.8 Results for the sDE

The last issue of our paper is to provide existence, uniqueness and regularity of stochastic flows for the sDE, imposing merely the LPS condition. The strategy here is to deduce such results from the path-by-path uniqueness result established in Section 1.7. To understand the novelties, let us recall that the more general strong well-posedness result for the sDE is due to [54] under the KR condition on b. To simplify the exposition and unify the discussion of the literature, let us consider the autonomous case $b(t,x) = b(x)$ and an assumption of the form $b \in L^p(\mathbb{R}^d, \mathbb{R}^d)$ (depending on the reference, various locality conditions and behavior at infinity are assumed). The condition $p > d$ seems to be the limit case for solvability in all approaches, see for instance [54, 51, 69, 19, 75], whether they are based on Girsanov theorem, Krylov estimates, parabolic theory or Dirichlet forms. There are some results on weak well-posedness for measure-valued drifts, see [9], and distribution-valued drifts, see [43, 27, 14], but it is unclear whether they apply to the limit case $p = d$: for example, the result in [9], when restricted to measures with density b with respect to the Lebesgue measure, requires $p > d$, see [9, Example 2.3]. The present paper is the first one to give information on sDEs in the limit case $p = d$ (apart from [64, 79]).

Since path-by-path solvability is another issue related to our results, let us mention [25], when the drift is bounded measurable: for a.e. ω, there exists one and only one solution. New results for several classes of noise and drift have been obtained by [15]. In general, the problem of path-by-path solvability of an sDE with poor drift is extremely difficult, compared to pathwise uniqueness which is already nontrivial. Thus it is remarkable that the approach by duality developed here gives results in this direction.

Our contribution on the sDE is threefold: existence, uniqueness and regularity of Lagrangian flows, pathwise uniqueness from a diffuse initial datum and path-by-path uniqueness from given initial condition. The following subsections detail these three classes of results.

1.8.1 Lagrangian flows

We prove a well-posedness result among Lagrangian flows (see below for more explanations) under the LPS condition on the drift:

Theorem 1.3. *Under LPS condition, for a.e. ω, there exists a unique Lagrangian flow Φ^ω solving the sDE at ω fixed. This flow is, at fixed time, $W^{1,m}_{\rm loc}(\mathbb{R}^d,\mathbb{R}^d)$-regular for every finite m, in particular it has a $C^\alpha(\mathbb{R}^d,\mathbb{R}^d)$ version (at fixed time) for every $\alpha<1$.*

Uniqueness will follow from uniqueness of the sCE, regularity from regularity of the solution to the sTE. The result is new because our uniqueness result is path-by-path: for a.e. ω, two Lagrangian flows solving the sDE with that ω fixed must coincide (notice that the sDE has a clear path-by-path meaning). A Lagrangian flow Φ, solving a given ODE, is a generalized flow, in the sense of [29, 2]: a measurable map $x\mapsto\Phi_t(x)$ with a certain non-contracting property, such that $t\mapsto\Phi_t(x)$ verifies that ODE for a.e. initial condition x. However, in general, we do not construct solutions of the sDE in a classical sense, corresponding to a given initial condition $X_0=x$. In fact we do not know whether or not strong solutions exist and are unique under the LPS condition with $\frac{d}{p}+\frac{2}{q}=1$ (while for $\frac{d}{p}+\frac{2}{q}<1$ strong solutions do exist, see [54]).

Let us mention that regularity under more restrictive assumptions was already proved, see for example [34, 63] or [15] (also for fractional Brownian motion). However, these results do not cover the full LPS condition and their proofs are based on the sDE, rather than on the sTE.

1.8.2 Pathwise uniqueness from a diffuse initial datum

We also prove a (classical) pathwise uniqueness result under the LPS condition, when the initial datum has a diffuse law. This is done by exploiting the regularity result of the sTE and by using a duality technique similar to the one mentioned before.

Theorem 1.4. *If X_0 is a diffuse random variable (not a single x) on $\mathbb{R}^d$, then pathwise uniqueness holds among solutions having diffuse marginal laws (more precisely, such that the law of X_t has a density in $L^\infty([0,T];L^m(\mathbb{R}^d))$, for a suitable m).*

Finally we notice that uniqueness of the law of solutions (or at least of their one-dimensional marginals, namely the solutions of Fokker-Planck equations) may hold true for very irregular drift, i.e. $b\in L^2$, if diffuse initial conditions with suitable density are considered; see [37, 12].

1.8.3 Results of path-by-path uniqueness from given initial condition

When the regularity results for the stochastic equation of transport type is improved from $W^{1,p}$ to C^1-regularity, then the uniqueness results of Section 1.7 for the sCE holds in the very general class of finite measures and it is a path-by-path uniqueness result. As a consequence, we get an analogous path-by-path uniqueness result for the sDE with classical given initial conditions, a result competitive with [25] and [15]. The main problem is to find assumptions, as weak as possible, on the drift b which are sufficient to guarantee C^1 regularity of the solutions. We describe two cases. The first one, which follows the strategy described in Section 1.1, is when the weak derivatives of b (instead of only b itself) belongs to the LPS class, that is $\partial_i b\in\mathcal{LPS}(p,q)$ for $i=1,\dots,d$. However, since this is a weak differentiability assumption, it is less general than expected. The second case is when b is Hölder continuous (in space) and bounded, but here we have to refer to [41, 39] for the proof of the main regularity results.

Theorem 1.5. *If Db belongs to the LPS class or if b is Hölder continuous (in space), then, for a.e. ω, for every x in $\mathbb{R}^d$, there exists a unique solution to the sDE, starting from x, at ω fixed.*

Notice that the "good subset" of Ω is independent of the initial condition x; this is not obvious from the approaches of [15, 25], due to the application of Girsanov transformation for a given initial condition.

Let us mention that in [74], generalized in [70] to the case of Lévy noise, path-by-path uniqueness is shown, from a fixed initial condition, for a Hölder continuous drift, using the regularity of the flow. This approach is the translation at the sDE level of the duality technique for the sPDE.

1.8.4 Summary on uniqueness results

Since the reader might not be acquainted with the various types of uniqueness, we resume here the possible path-by-path uniqueness results and their links.

path-by-path uniqueness among trajectories	$\Rightarrow$	pathwise uniqueness for deterministic initial data
$\Downarrow$		$\Downarrow$
path-by-path uniqueness among flows	"$\Rightarrow$"	pathwise uniqueness for diffuse initial data

Let us explain more in detail these implications (this is in parts heuristics and must not be taken as rigorous proofs):

- Path-by-path uniqueness among trajectories implies path-by-path uniqueness among flows: Assume path-by-path (or pathwise) uniqueness among trajectories and let Φ, Ψ be two flows solving the sDE. Then, for a.e. x in $\mathbb{R}^d$, $\Phi(x)$ and $\Psi(x)$ are solutions to the sDE, starting from x, so, by uniqueness, they must coincide and hence $\Phi = \Psi$ a.e..

- Path-by-path uniqueness among trajectories implies pathwise uniqueness for deterministic initial data: Assume path-by-path uniqueness among trajectories and let X, Y be two adapted processes which solve the sDE. Then, for a.e. ω, $X(\omega)$ and $Y(\omega)$ solve the sDE for that fixed ω, so they must coincide and hence $X = Y$ a.e..

- Pathwise uniqueness for deterministic initial data implies pathwise uniqueness for diffuse initial data: Assume pathwise uniqueness for deterministic initial data and let X, Y be two solutions, on a probability space $(\Omega, \mathcal{A}, P)$, starting from a diffuse initial datum X_0. For x in $\mathbb{R}^d$, define the set $\Omega_x = \{\omega \in \Omega \colon X_0(\omega) = x\}$. Then, for $(X_0)_{\#}P$-a.e. x, X and Y, restricted to Ω_x, solve the sDE starting from x, so they must coincide and hence $X = Y$ a.e..

- Path-by-path uniqueness among flows (with non-concentration properties) "implies" pathwise uniqueness for diffuse initial data: The quotation marks are here for two reasons: because the general proof is more complicated than the idea below and because the pathwise uniqueness is not among all the processes (with diffuse initial data), but a restriction is needed to transfer the non-concentration property. Assume path-by-path uniqueness among flows and let X, Y be two solutions on a probability space $(\Omega, \mathcal{A}, P)$, starting from a diffuse initial datum X_0. We give the idea in the case $\Omega = C([0,T];\mathbb{R}^d) \times B_R \ni \omega = (\gamma, x)$ (the Wiener space times the space of initial datum), $P = Q \times \mathcal{L}^d$, where Q is the Wiener measure, $W(\gamma, x) = \gamma$, $X_0(\gamma, x) = x$. In this case (which is a model for the general case), for Q-a.e. γ, $\Phi(\gamma, \cdot)$ and $\Psi(\gamma, \cdot)$ are flows solving the SDE for that fixed ω. If they have the required non-concentration properties, then, by uniqueness, they must coincide. Hence uniqueness holds among processes X, with diffuse X_0, such that $X(\gamma, \cdot)$ has a certain non-concentration property; this is the restriction we need.

We will prove: path-by-path uniqueness among Lagrangian flows, when b is in LPS class; path-by-path uniqueness among solutions starting from a fixed initial point, when b and Db are in LPS class or when b is Hölder continuous (in space). We will develop in detail pathwise uniqueness from a diffuse initial datum in Section 5.1 (where the last implication will be proved) and in Section 5.2 (where a somehow more general result will be given).

1.8.5 Examples

In Section 7 we give several examples of equations with irregular drift of two categories:

i) on one side, several examples of drifts which in the deterministic case give rise to non-uniqueness, discontinuity or shocks in the flow, while in the stochastic case our results apply and these problems disappear;

ii) on the other side, a counterexample of a drift outside of the LPS class, for which even the sDE is ill-posed.

1.9 Concluding remarks and generalizations

The three classes of results described above are listed in logical order: we need the regularity results of Section 1.6 for sTE and sCE to prove the uniqueness results of Section 1.7 for sCE and sTE by duality; then we deduce the results of Section 1.8 for sDE from such uniqueness results. The fact that regularity for transport equations (with poor drift) is the starting point marks the difference with the deterministic theory, where such kinds of results are absent. Hence, the results and techniques of the present paper are not generalizations of deterministic ones.

The two most innovative technical tools developed in this work are the analytic proof of regularity (Section 1.6) and the path-by-path duality argument yielding uniqueness in this very strong sense. The generality of LPS condition seems to be unreachable with more classical tools, based on a direct analysis of the sDE. Moreover, in principle some of the analytic steps of Section 1.6 and the duality argument could be applied to other classes of stochastic equations; however, the renormalization step in the regularity proof is quite peculiar of transport equations.

The noise considered in this work is the simplest one, in the class of multiplicative noise of transport type. The reason for this choice is that it suffices to prove the regularization phenomena and the exposition will not be obscured by unnecessary details. However, for nonlinear problems it seems that more structured noise is needed, see [42, 28]. So it is natural to ask whether the results of this paper extend to such noise. Let us briefly discuss this issue. The more general sDE takes the form

$$dX = b(t,X)dt + \sum_{k=1}^{\infty} \sigma_k(X) \circ dW_t^k, \qquad X_0 = x \tag{1.2}$$

where $\sigma_k : \mathbb{R}^d \to \mathbb{R}^d$ and W^k are independent Brownian motions, and sTE, sCE are now

$$du + b \cdot \nabla u dt + \sum_{k=1}^{\infty} \sigma_k \cdot \nabla u \circ dW_t^k = 0, \qquad u|_{t=0} = u_0 \tag{1.3}$$

$$d\mu + \operatorname{div}(b\mu)dt + \sum_{k=1}^{\infty} \operatorname{div}(\sigma_k \mu) \circ dW_t^k = 0, \qquad \mu|_{t=0} = \mu_0 \tag{1.4}$$

Concerning the assumptions on σ_k, for simplicity, think of the case when they are of class C_b^4 with proper summability in k. In order to generalize the regularity theory (Section 1.6) it is necessary to be able to perform parabolic estimates and thus the generator associated to this sDE must be strongly elliptic; a simple sufficient condition is that the covariance matrix function $Q(x,y) \coloneqq \sum_{k=1}^{\infty} \sigma_k(x) \otimes \sigma_k(y)$ of the random field $\eta(t,x) = \sum_{k=1}^{\infty} \sigma_k(x) W_t^k$ depends only on $x-y$ (namely $\eta(t,x)$ is space-homogeneous), $\operatorname{div} \sigma_k(x) = 0$ (this simplifies several lower order terms) and for the function $Q(x) = Q(x-y)$ we have

$$\det Q(0) \neq 0.$$

This replaces the assumption $\sigma \neq 0$.

The duality argument (Section 1.7) is very general and in principle it does not require any special structure except the linearity of the equations. However, in the form developed here, we use auxiliary random PDEs associated to the sPDEs via the simple transformation $v(t,x) \coloneqq u(t, x + \sigma W_t)$; we do this in order to avoid troubles with backward and forward sPDEs at the same time. But this simple transformation requires additive noise. In the case of multiplicative noise, one has to consider the auxiliary stochastic equation

$$dY = \sum_{k=1}^{\infty} \sigma_k(Y) \circ dW_t^k, \qquad Y_0 = y \tag{1.5}$$

and its stochastic flow of diffeomorphisms $\psi_t(x)$, and use the transformation $v(t,x) \coloneqq u(t, \psi_t(x))$. This new random field satisfies

$$\partial_t v + \widetilde{b} \cdot \nabla v = 0$$

where

$$\tilde{b}(t,x) \coloneqq D\psi_t^{-1}(\psi_t(x)) b(t, \psi_t(x)),$$

and the duality arguments can be repeated, in the form developed here. The uniqueness results mentioned in Section 1.7 then extend to this case.

The path-by-path analysis of the SDE (1.2) may look a priori less natural, since this equation does not have a pathwise interpretation. However, when the coefficients σ_k are sufficiently regular to generate, for the auxiliary equation (1.5), a stochastic flow of diffeomorphisms $\psi_t(x)$, then we may give a (formally) alternative formulation of (1.2) as a random differential equation, of the form

$$dZ = \tilde{b}(t, Z)dt,$$

in analogy with the random PDE for the auxiliary variable $v(t, x)$ above. This equation can be studied pathwise, with the techniques of Section 1.8. Here, however, we feel that more work is needed in order to connect the results with the more classical viewpoint of equation (1.2) and thus we refrain to express strong claims here.

Concerning the path-by-path uniqueness, say for the sDE, note that this issue can be studied for any *deterministic* path W, not necessarily the sample paths from the Brownian motion. Hence, one can ask which conditions on a single, deterministic path W ensure uniqueness of (sDE), which is now a deterministic ODE. This is investigated in [15], where the concept of (ρ, γ)-irregular paths is given by means of Fourier analysis and it is shown that such paths provide uniqueness for certain classes of non-Lipschitz drifts b (in particular if W is a sample path of the Brownian motion, uniqueness is shown for Hölder continuous drifts). In contrast to the present paper, the techniques used in [15] are based on Young integration, and the results, when specialized to Brownian sample paths, are mostly concerned with Hölder continuous drifts. While for a general path it is not easy to verify the (ρ, γ)-irregularity condition, one can prove, see [21, Proposition 1.4], that this condition implies that the path must be irregular (non-Lipschitz in time): this corresponds to the fact that a regular path does not regularize an ill-posed ODE, in general. It would be interesting to compare the (ρ, γ)-irregularity notion with the concept of truly rough paths (e.g. [45]), which also quantifies the irregularity of a path. Another, somehow more explicit, sufficient condition on deterministic paths is given in [20, equation (3.3)], though it is used for regularization of scalar conservation laws rather than ODEs. Here the regularization of non-linear PDEs was achieved by means of a noise, that is here the derivative of the regularizing path, which is itself non-linear and precisely multiplies the non-linearity: see e.g. [20, 21, 22], and [47] for other pathwise arguments.

Throughout the paper, the drift b is assumed to be deterministic. In view of applications especially to non-linear equations, it would be very important to extend the result to random drifts. While we do not see obstacles for the extension of the the duality technique, being path-by-path in nature, the first step, namely the proof of regularity of solutions for the sTE, does not allow for such generalization: if the drift were random, then the equations for the moments of the derivative of the solution would not form a closed system. This is not simply a limitation of the techniques: there are in fact simple counterexamples to regularization by noise for general random drifts. Let us mention that, in some cases, it is possible to have regularization by noise even for random drifts, see [15] and related work, assuming a suitable Hölder continuity of the drift, or [30, 68], assuming Malliavin differentiability of the drift.

Finally, let us note that throughout the rest of the paper, concerning function spaces, we shall use for simplicity the same notation for scalar-valued and vector-valued functions (but it will be always clear from the context if the functions used have values in $\mathbb{R}^d$, like b, X, or Φ, or in $\mathbb{R}$, like c or solutions u, v).

2 Regularity for sTE and sCE

In order to unify the analysis of the sTE and sCE we introduce the stochastic generalized transport equation ($sgTE$) in $\mathbb{R}^d$

$$du + (b \cdot \nabla u + cu)dt + \sigma \nabla u \circ dW = 0, \qquad u|_{t=0} = u_0 \tag{sgTE}$$

where b, σ, u and u_0 are as above, $c\colon [0, T] \times \mathbb{R}^d \to \mathbb{R}$ and W is a Brownian motion with respect to a given filtration $(\mathcal{G}_t)_t$. We shall prove regularity results for solution to (sgTE).

Remark 2.1. *We note that the case $c = 0$ corresponds to* (sTE)*, while the case $c = \operatorname{div} b$ corresponds to* (sCE)*, with*

$$du + \operatorname{div}(bu)dt + \sigma \operatorname{div}(u \circ dW_t) = 0, \qquad u|_{t=0} = u_0, \tag{2.1}$$

where u stands for the density of the measure μ_t with respect to the Lebesgue measure.

2.1 Assumptions

Throughout all the paper, we assume that $(\Omega, \mathcal{A}, P)$ is a probability space, $(\mathcal{G}_t)_{t\in[0,T]}$ is a filtration satisfying the standard assumptions, that is, it is complete and right-continuous. The process W denotes a Brownian motion with respect to $(\mathcal{G}_t)_t$, unless differently specified.

Concerning the general equation (sgTE) we will always assume that we are in the purely stochastic case with $\sigma \neq 0$ and that the coefficients b and c satisfy the following decomposition and regularity condition.

Condition 2.2 (LPS+reg)**.** *The fields b and c can be written as $b = b^{(1)} + b^{(2)}$, $c = c^{(1)} + c^{(2)}$, where*

1. LPS-condition: *$b^{(1)}$, $c^{(1)}$ satisfy one of the following three assumptions:*
   a) *$b^{(1)}$, $c^{(1)}$ are in $\mathcal{LPS}(p,q)$ for some p, q in $(2,\infty)$ (with $\frac{2}{q} + \frac{d}{p} \leq 1$) or $p = \infty$, $q = 2$;*
   b) *$b^{(1)}$, $c^{(1)}$ are in $C([0,T]; L^d(\mathbb{R}^d))$ with $d \geq 3$;*
   c) *$b^{(1)}$, $c^{(1)}$ are in $L^\infty([0,T]; L^d(\mathbb{R}^d))$ with $d \geq 3$ and there hold*

   $$\|b^{(1)}\|_{L^\infty([0,T];L^d(\mathbb{R}^d))} \leq \delta \quad \text{and} \quad \|c^{(1)}\|_{L^\infty([0,T];L^d(\mathbb{R}^d))} \leq \delta,$$

   with δ small enough; precisely, given an exponent m as in Theorem 2.7*, δ depends on m, σ, d, as given by inequality* (2.18)*;*
2. Regularity condition: *$b^{(2)}$ is in $L^2([0,T]; C^1_{lin}(\mathbb{R}^d))$ and $c^{(2)}$ is in $L^2([0,T]; C^1_b(\mathbb{R}^d))$, i.e., for a.e. $t \in [0,T]$, $b^{(2)}(t,\cdot)$ and $c^{(2)}(t,\cdot)$ are in $C^1(\mathbb{R}^d)$ and*

$$\|b^{(2)}\|^2_{L^2([0,T];C^1_{lin}(\mathbb{R}^d))} \coloneqq \int_0^T \left(\left\| \frac{b^{(2)}(s,\cdot)}{1+|\cdot|} \right\|_\infty + \|Db^{(2)}(s,\cdot)\|_\infty \right)^2 ds < \infty,$$

$$\|c^{(2)}\|^2_{L^2([0,T];C^1_b(\mathbb{R}^d))} \coloneqq \int_0^T \left(\|c^{(2)}(s,\cdot)\|_\infty + \|Dc^{(2)}(s,\cdot)\|_\infty \right)^2 ds < \infty.$$

(*The expression "b is in a certain class A" must be understood componentwise.*)

Remark 2.3. *The hypotheses on $b^{(2)}$ and $c^{(2)}$ are slightly stronger than the natural ones, namely $b^{(2)}$ in $L^1([0,T]; C^1_{lin}(\mathbb{R}^d))$, $c^{(2)}$ in $L^1([0,T]; C^1_b(\mathbb{R}^d))$: we require L^2 integrability in time instead of L^1. This is mainly due to a technical point which will appear in Section* 3*. However, with minor modifications, this assumption could be relaxed to L^1 integrability throughout this section.*

Remark 2.4. *A simple extension of Condition* 2.2 *is to ask that $b = \sum_{j=1}^N \tilde{b}^{(j)}$, where, for every j, $\tilde{b}^{(j)}$ is a vector field satisfying Condition* 2.2 *with exponents p_j, q_j that can depend on j; similarly for c. This extension is quite easy and we refrain to discuss it explicitly.*

Remark 2.5. *The sTE is just equation* (sgTE) *with $c = 0$ and thus we do write explicitly the assumptions for sTE. The sCE instead corresponds to* (sgTE) *with $c = \operatorname{div} b$ and for completeness let us note that we hence need to assume for sCE that we have $b = b^{(1)} + b^{(2)}$, with*

- (i_0) *for some $p, q \in (2,\infty)$, or $(p,q) = (\infty, 2)$, $b^{(1)}, \operatorname{div} b^{(1)} \in \mathcal{LPS}(p,q)$;*
- (i_1) *for $(p,q) = (d,\infty)$, $d \geq 3$, either we assume $b^{(1)} \in C([0,T]; L^d(\mathbb{R}^d))$ or we require the smallness assumption in Condition* 2.2*,* 1c*);*
- (ii) *$b^{(2)}, \operatorname{div} b^{(2)} \in L^2(0,T; C^1(\mathbb{R}^d))$, with*

$$\int_0^T \left(\left\| \frac{b^{(2)}(s,\cdot)}{1+|\cdot|} \right\|_\infty + \left\| Db^{(2)}(s) \right\|_\infty + \left\| D \operatorname{div} b^{(2)}(s) \right\|_\infty \right)^2 ds < \infty.$$

2.2 Strategy of proof

In order to prove the regularity results, we follow the approach of a priori estimates: we prove regularity estimates for the smooth solutions of approximate problems with smooth coefficients, be careful to show that the regularity estimates have constants independent of the approximation; then we deduce the regularity for the solution of the limit problem by passing to the limit.

The strategy of proof is made of several steps which bear some similarities with the computations done in literature of theoretical physics of passive scalars, see for instance [18].

First, we differentiate the sgTE (which is possible because we deal with smooth solutions of regularized problems), with the purpose of estimating the derivatives of the solutions. However, terms like $\partial_i b_k$ appear. In the deterministic case, *unless b is Lipschitz*, these terms spoil any attempt to prove differentiability of solutions by this method. In the stochastic case, we shall integrate these bad terms by parts at the price of a second derivative of the solution, which however will be controlled, as it will be explained below.

Second, we use the very important property of transport type equations of being invariant under certain transformations of the solution. For the classical sTE, the typical transformation is $u \mapsto \beta(u)$ where $\beta \in C^1(\mathbb{R})$: if u is a solution, then $\beta(u)$ is (at least formally) again a solution. For regular solutions, as in our case, this can be made rigorous; let us only mention that, for weak solutions, this is a major issue, which gives rise to the concept of renormalized solutions [29] (namely those for which $\beta(u)$ is again a weak solution) and the so called commutator lemma; we do not meet these problems here, in the framework of regular solutions. Nevertheless, to recall the issue, we shall call this step *renormalization*, namely that suitable transformations of the solution lead to solutions. In our case, since we consider the differentiated sgTE, we work on the level of derivatives of the solution u and therefore we apply transformations to $\partial_i u$. In order to find a closed system, we have to consider, as transformations, all possible products of $\partial_i u$, and u itself. This leads to some complications in the book-keeping of indices, but the essential idea is still the renormalization principle.

Third, we reformulate the sPDE from the Stratonovich to the Itô form. The corrector is a second order differential operator. It is strongly elliptic in itself, but combined with the Itô term (containing first derivatives of solutions), it does not give a parabolic character to the equation. The equation is indeed equivalent to the original, hyperbolic (time-reversible) formulation.

Forth, we take the expectation. This projection annihilates the Itô term and gives a true parabolic equation. The expected value of powers of $\partial_i u$ (or any product of them) solves a parabolic equation, and, as a system in all possible products, it is a closed system. For other functionals of the solution, as the two-point correlation function $E[u(t,x)u(t,y)]$, the fact that a closed parabolic equation arises was a basic tool in the theory of passive scalars [18].

Finally, on the parabolic equation we perform energy-type estimates. The elliptic term puts into play, on the positive side of the estimates, terms like $\nabla E[(\partial_i u)^m]$. They are the key tool to estimate those terms coming from the partial integration of $\partial_i b_k$ (see the comments above). The good parabolic terms $\nabla E[(\partial_i u)^m]$ come from the Stratonovich-Itô corrector, after projection by the expected value. This is the technical difference to the deterministic case.

2.3 Preparation

The following preliminary lemma is essentially known, although maybe not explicitly written in all details in the literature; we shall therefore sketch the proof. As explained in the last section, given non-smooth coefficients, we shall approximate them with smooth ones. Their role is only to allow us to perform certain computations on the solutions (such as Itô formula, finite expected values, finite integrals on $\mathbb{R}^d$ and so on). More precisely, the outcome of the next lemma are C_c^∞-estimates (infinitely differentiable with compact support in all variables) in space for all times, for the solutions corresponding to the equation with smooth (regularized) coefficients. However, we emphasize that these estimates are not uniform in the approximations, in contrast to our main regularity estimates concerning Sobolev-type regularity established later on in Theorem 2.7.

Lemma 2.6. *If $b, c \in C_c^\infty([0,T] \times \mathbb{R}^d)$, $u_0 \in C_c^\infty(\mathbb{R}^d)$, then there exists an adapted solution u of equation* (sgTE) *with paths of class $C([0,T]; C_c^\infty(\mathbb{R}^d))$ (the support of u depends on the path). We*

have

$$\sup_{(t,x)\in[0,T]\times\mathbb{R}^d} E\big[|D^\alpha u(t,x)|^r\big] < \infty \tag{2.2}$$

for every $\alpha \geq 0$ and $r \geq 1$. Moreover, we have

$$\sup_{(t,x,\omega)\in[0,T]\times\mathbb{R}^d\times\Omega} |u(t,x,\omega)| \leq \|u_0\|_\infty e^{\int_0^T \|c(s,\cdot)\|_\infty ds} \tag{2.3}$$

and for every $r, R \geq 1$

$$\sup_{t\in[0,T]} \int_{\mathbb{R}^d} \big(1+|x|^R\big) E\big[|D^\alpha u(t,x)|^r\big] dx < \infty, \qquad \text{for } \alpha = 0,1,2. \tag{2.4}$$

Proof. Step 1: Existence of a solution. Under the assumption $b \in C_c^\infty([0,T]\times\mathbb{R}^d)$, equation (sDE) has a pathwise unique strong solution X_t^x for every given $x \in \mathbb{R}^d$. As proved in [56], the random field X_t^x has a modification $\Phi_t(x)$ which is a stochastic flow of diffeomorphisms of class C^∞ (since b is infinitely differentiable with bounded derivatives). Moreover, in view of [57, Theorem 6.1.9] we know that, given $u_0 \in C_c^\infty(\mathbb{R}^d)$, the process

$$u(t,x) \coloneqq u_0\big(\Phi_t^{-1}(x)\big) e^{\int_0^t c(s,\Phi_s^{-1}(x))ds} \tag{2.5}$$

(which has paths of class $C([0,T]; C_c^\infty(\mathbb{R}^d))$ by the properties of Φ_t^{-1}) is an adapted strong solution to (sgTE). Inequality (2.3) then follows from (2.5).

Step 2: Regularity of the solution. For the flow $\Phi_t(x)$ we have the simple inequality

$$|\Phi_t(x)| \leq |x| + T\|b\|_\infty + |\sigma||W_t|$$

and thus, for every $R > 0$ there exists a constant $C_R > 0$ such that

$$E\big[|\Phi_t(x)|^R\big] \leq C_R\big(|x|^R + T^R\|b\|_\infty^R + |\sigma|^R T^{R/2}\big). \tag{2.6}$$

This bound will be used below. For the derivative of the flow with respect to the initial condition in the direction h, $D_h\Phi_t(x) = \lim_{\varepsilon\to 0} \varepsilon^{-1}(\Phi_t(x+\varepsilon h) - \Phi_t(x))$, one has

$$\frac{d}{dt} D_h\Phi_t(x) = Db(t,\Phi_t(x)) D_h\Phi_t(x), \qquad D_h\Phi_0(x) = h$$

and thus, since Db is bounded,

$$|D_h\Phi_t(x)| \leq C_1|h| \qquad \text{for } t \in [0,T], \tag{2.7}$$

where $C_1 \geq 1$ is a deterministic constant. The same is true for higher derivatives and for the inverse flow. This proves inequality (2.2) for $\alpha > 0$, while the inequality for $\alpha = 0$ comes from (2.3).

Concerning the claim (2.4), for $\alpha = 0$ and $t \in [0,T]$ we have

$$\begin{aligned}
E\Big[\int_{\mathbb{R}^d} \big(1+|x|^R\big)|u(t,x)|^r dx\Big] &\leq e^{r\int_0^T \|c(s,\cdot)\|_\infty ds} E\Big[\int_{\mathbb{R}^d} \big(1+|x|^R\big)\big|u_0\big(\Phi_t^{-1}(x)\big)\big|^r dx\Big] \\
&= e^{r\int_0^T \|c(s,\cdot)\|_\infty ds} E\Big[\int_{\mathbb{R}^d} \big(1+|\Phi_t(y)|^R\big)|u_0(y)|^r |\det D\Phi_t(y)| dy\Big] \\
&\leq C_{2,r} \int_{\mathbb{R}^d} \big(1+E\big[|\Phi_t(y)|^R\big]\big)|u_0(y)|^r dy
\end{aligned}$$

by (2.7), where $C_{2,r} = C_1^d e^{r\int_0^T \|c(s,\cdot)\|_\infty ds}$. Combined with (2.6) this implies (2.4) for $\alpha = 0$, since u_0 has compact support. The proof of (2.4) for $\alpha = 1, 2$ is similar: we first differentiate u by using the explicit formula (2.5) and get several terms, then we control them by means of boundedness of c and its derivatives, boundedness of derivatives of direct and inverse flow, and the change of variable formula used above for $\alpha = 0$. The computation is lengthy but elementary. For instance we have

$$D_h u(t,x) = e^{\int_0^t c(s,\Phi_s^{-1}(x))ds} Du_0\big(\Phi_t^{-1}(x)\big) D_h\Phi_t^{-1}(x)$$

$$+ u_0\big(\Phi_t^{-1}(x)\big)e^{\int_0^t c(s,\Phi_s^{-1}(x))ds}\int_0^t Dc\big(s,\Phi_s^{-1}(x)\big)D_h\Phi_s^{-1}(x)ds,$$

$$|D_hu(t,x)|^r \le C_rC_{2,r}C_1^r\big|Du_0\big(\Phi_t^{-1}(x)\big)\big|^r|h|^r + C_rC_{2,r}T^r\|Dc\|_\infty^rC_1^r\big|u_0\big(\Phi_t^{-1}(x)\big)\big|^r|h|^r.$$

Hence, we obtain

$$E\Big[\int_{\mathbb{R}^d}\big(1+|x|^R\big)|D_hu(t,x)|^rdx\Big] \le C_rC_{2,r}C_1^{r+d}|h|^r\int_{\mathbb{R}^d}\big(1+E\big[|\Phi_t(y)|^R\big]\big)|Du_0(y)|^rdy$$
$$+ C_rC_{2,r}T^r\|Dc\|_\infty^2C_1^{r+d}|h|^r\int_{\mathbb{R}^d}\big(1+E\big[|\Phi_t(y)|^R\big]\big)|u_0(y)|^rdy$$

which implies (2.4) for $\alpha = 1$. The proof is complete. □

2.4 Main result on a priori estimates

In the sequel we take the regular solution given by Lemma 2.6 and prove a priori estimates. For the formulation of the result, let us introduce a C^1-function $\chi\colon \mathbb{R}^d \to [0,\infty)$ such that

$$|\nabla\chi(x)| \le C_\chi\frac{\chi(x)}{1+|x|} \qquad \text{for all } x\in\mathbb{R}^d \tag{2.8}$$

for some constant $C_\chi > 0$. For example, we might take $\chi(x) = (1+|x|^2)^{s/2}$ which satisfies $|\nabla\chi(x)| \le 2|s|\chi(x)/(1+|x|)$, for every $s\in\mathbb{R}$ (all cases $s<0$, $s=0$, and $s>0$ are of interest). The associated norm $\|u_0\|_{W^{1,r}_\chi(\mathbb{R}^d)}$ is defined by

$$\|u_0\|^r_{W^{1,r}_\chi(\mathbb{R}^d)} = \sum_{i=0}^d\int_{\mathbb{R}^d}|\partial_iu_0(x)|^r\chi(x)dx$$

where we have used the notation $\partial_0 f = f$.

Theorem 2.7. *Let p,q be in $(2,\infty)$ satisfying $\frac{2}{q}+\frac{d}{p}\le 1$ or $(p,q)=(\infty,2)$, let m be a positive integer, let $\sigma\neq 0$, and let χ be a function satisfying* (2.8)*. Assume that b and c are a vector field and a scalar field, respectively, such that $b = b^{(1)}+b^{(2)}$, $c = c^{(1)}+c^{(2)}$, with $b^{(i)}$, $c^{(i)}$ in $C_c^\infty([0,T]\times\mathbb{R}^d)$ for $i=1,2$. Then there exists a constant C such that, for every u_0 in $C_c^\infty(\mathbb{R}^d)$, the smooth solution u of equation* (sgTE) *starting from u_0, given by Lemma* 2.6*, verifies*

$$\sup_{t\in[0,T]}\sum_{i=0}^d\int_{\mathbb{R}^d}E\big[(\partial_iu(t,x))^m\big]^2\chi(x)dx \le C\|u_0\|^{2m}_{W^{1,2m}_\chi(\mathbb{R}^d)}.$$

Moreover, the constant C can be chosen to have continuous dependence on m,d,σ,χ,p,q and on the $L^q([0,T];L^p(\mathbb{R}^d))$ norms of $b^{(1)}$ and $c^{(1)}$, on the $L^1([0,T];C^1_{lin}(\mathbb{R}^d))$ norm of $b^{(2)}$, and on the $L^1([0,T];C^1_b(\mathbb{R}^d))$ norm of $c^{(2)}$.

The result holds also for $(p,q)=(d,\infty)$ with the additional hypothesis that the $L^\infty([0,T];L^d(\mathbb{R}^d))$ norms of $b^{(1)}$ and $c^{(1)}$ are smaller than δ, see Condition 2.2*,* 1c*) (in this case the continuous dependence of C on these norms is up to δ).*

Corollary 2.8. *With the same notations of the previous theorem, if m is an even integer, then for every $s\in\mathbb{R}$ there exists a constant C depending also on s (in addition to the dependencies from the theorem) such that*

$$\sup_{t\in[0,T]}E\Big[\|u(t,\cdot)\|^m_{W^{1,m}_{(1+|\cdot|)^s}(\mathbb{R}^d)}\Big] \le C\|u_0\|^m_{W^{1,2m}_{(1+|\cdot|)^{2s+d+1}}(\mathbb{R}^d)}.$$

Proof. Via Hölder's inequality we have

$$\int_{\mathbb{R}^d}\big(1+|x|\big)^sE\big[|\partial_iu(t,x)|^m\big]dx = \int_{\mathbb{R}^d}\big(1+|x|\big)^{-\frac{d+1}{2}}\big(1+|x|\big)^{s+\frac{d+1}{2}}E\big[|\partial_iu(t,x)|^m\big]dx$$

$$\leq \Big(\int_{\mathbb{R}^d} \big(1+|x|\big)^{-d-1} dx \Big)^{1/2} \Big(\int_{\mathbb{R}^d} \big(1+|x|\big)^{2s+d+1} E\big[|\partial_i u(t,x)|^m\big]^2 dx \Big)^{1/2}$$
$$\leq C \|u_0\|^m_{W^{1,2m}_{(1+|\cdot|)^{2s+d+1}}(\mathbb{R}^d)}.$$

for a suitable constant $C > 0$. □

Remark 2.9. *Such power-type weights play a crucial role for later applications. Therefore, let us note that, for every $s \in \mathbb{R}$ and $m \in (1,\infty)$, $W^{1,m}_{(1+|\cdot|)^s}(\mathbb{R}^d)$ is a reflexive Banach space. We can show this, for instance, by observing that the dual of $L^m_{(1+|\cdot|)^s}(\mathbb{R}^d)$ is isomorphic to $L^{m'}_{(1+|\cdot|)^{sm'/m}}(\mathbb{R}^d)$ with $1/m+1/m' = 1$. Hence, the L^m spaces with these weights are reflexive, which directly carries over to the weighted Sobolev spaces since they are closed subspaces via the mapping $f \mapsto (f, \partial_1 f, \dots, \partial_d f)$. The same holds for spaces like $L^m([0,T]\times\Omega; W^{1,m}_{(1+|\cdot|)^s}(\mathbb{R}^d))$ and $L^m([0,T]\times\Omega; L^m_{(1+|\cdot|)^s}(\mathbb{R}^d))$. In particular, the Banach–Alaoglu theorem is at our disposal.*

The next subsections are devoted to the proof of the a priori estimate of the theorem. At the end, they will be used to construct a (weaker) solution corresponding to non-smooth data. Thus, in the sequel, u refers to a smooth solution, with smooth and compactly supported data b, u_0.

2.5 Formal computation

This section serves as a formal explanation of the first main steps of the proof, those based on renormalization, passage from Itô to Stratonovich formulation and taking the expectation. A precise statement and proof is given in the next Section 2.6.

The aim of the following computations is to write, given any positive integer m, a closed system of parabolic equations for the quantities $E[\prod_{i\in I} \partial_i u]$, where I varies in the finite multi-indices with elements in $\{0,1,\dots,d\}$ of length at most m. In principle, we need only the quantities $E[(\partial_i u)^m]$ for $i = 1,\dots,d$, but they do not form a closed system.

Equation (sgTE) is formally of the form

$$\mathcal{L}u + cu = 0$$

where $\mathcal{L}$ is the differential operator

$$\mathcal{L}f = \partial_t f + b\cdot\nabla f + \sigma\nabla f \circ \dot{W}.$$

Being a first order differential operator, it formally satisfies the Leibniz rule

$$\mathcal{L}\Big(\prod_{j=1}^m f_j\Big) = \sum_{i=1}^m \prod_{j\neq i} f_j \mathcal{L} f_i. \tag{2.9}$$

This is the step that we call renormalization, following [29]: in the language of that paper, if $\beta\colon \mathbb{R}\to\mathbb{R}$ is a C^1-function and v is a solution of $\mathcal{L}v = 0$, then formally $\mathcal{L}\beta(v) = 0$, and solutions which satisfy this rule rigorously are called renormalized solutions. Property (2.9) is a variant of this idea. We apply the renormalization to first derivatives of u. Precisely, if u is a solution of $\mathcal{L}u + cu = 0$, we set

$$v_i \coloneqq \partial_i u, \qquad \text{for } i = 1,\dots,d.$$

One has $\partial_i(\mathcal{L}u + cu) = 0$ and thus

$$\mathcal{L}v_i = -(\partial_i b\cdot\nabla u + u\partial_i c + cv_i), \qquad i = 1,\dots,d.$$

With the notation $v_0 = u$ we also have

$$\mathcal{L}v_0 = -cv_0.$$

In the sequel, I will be a finite multi-index with elements in $\{0,1,\dots,d\}$, namely an element of $\cup_{m\in\mathbb{N}}\{0,1,\dots,d\}^m$. If $I\in\{0,1,\dots,d\}^m$ we set $|I| = m$. Given a function $h\colon\{0,1,\dots,d\}\to\mathbb{R}$, $\sum_{i\in I} h(i)$ means the sum over all the components of I (counting repetitions), and similarly for the

product. The multi-index $I \setminus i$ means that we drop in I a component of value i; the multi-index $I \setminus i \cup k$ means that we substitute in I a component of value i with a component of value k; which component i is dropped or replaced does not matter because we consider only expressions of the form $\sum_{i \in I} h(i)$ and similar ones. Let us set

$$v_I \coloneqq \prod_{i \in I} v_i \quad \text{which satisfies} \quad \mathcal{L} v_I = \sum_{i \in I} v_{I \setminus i} \mathcal{L} v_i$$

in view of the Leibniz rule (2.9). Now, the equations for v_i differ depending on whether $i = 0$ or $i \in \{1, \dots, d\}$. The term $c v_i$ appears in all of them, but not $\partial_i b \cdot \nabla u + u \partial_i c$. Hence, we find

$$\begin{aligned} \mathcal{L} v_I &= -c\,|I|\, v_I - \sum_{i \in I, i \neq 0} v_{I \setminus i} (\partial_i b \cdot \nabla u + u \partial_i c) \\ &= -c\,|I|\, v_I - \sum_{i \in I, i \neq 0} \sum_{k=1}^{d} v_{I \setminus i \cup k} \partial_i b_k - \sum_{i \in I, i \neq 0} v_{I \setminus i \cup 0} \partial_i c. \end{aligned}$$

Next we want to take the expected value. The problem is the Stratonovich term $\sigma \nabla v_I \circ \dot{W}$ in $\mathcal{L} v_I$. Rewriting it as an Itô term with correction, we get

$$\sigma \nabla v_I \circ dW = \sigma \nabla v_I dW + \frac{1}{2} \sum_{k=1}^{d} d\big[\sigma \partial_k v_I, W^k\big]$$

where the brackets $[\cdot, \cdot]$ denote the joint quadratic variation. Since dv_I has $-\sigma \sum_{k'=1}^{d} \partial_{k'} v_I dW^{k'}$ as local martingale term, we have $d[\sigma \partial_k v_I, W^k] = -\sigma^2 \partial_k^2 v_I dt$, and thus, we find

$$\sigma \nabla v_I \circ dW = \sigma \nabla v_I dW - \frac{\sigma^2}{2} \Delta v_I dt.$$

Taking (formally) the expectation in the equation for v_I, we arrive at

$$\partial_t w_I + b \cdot \nabla w_I + c\,|I|\, w_I + \sum_{i \in I, i \neq 0} \sum_{k=1}^{d} w_{I \setminus i \cup k} \partial_i b_k + \sum_{i \in I, i \neq 0} w_{I \setminus i \cup 0} \partial_i c = \frac{\sigma^2}{2} \Delta w_I \tag{2.10}$$

where

$$w_I \coloneqq E[v_I].$$

This is the first half of the proof of Theorem 2.7, which will be carried out rigorously in Section 2.6. The second half is the estimate on w coming from the parabolic nature of this equation, which will be established in Section 2.7.

2.6 Rigorous proof of (2.10)

We work with the regular solution u given by Lemma 2.6 and we use the notations I, $I \setminus i \cup k$, v_i, v_I, w_I as in the previous section. We observe that, since u is smooth in x, the v_i's and their spatial derivatives are well-defined. Moreover, due to inequality (2.2), also the expected values w_I's are well-defined and smooth in x.

Lemma 2.10. *The function $w_I(t, x)$ is continuously differentiable in time and satisfies the (pointwise) equation (2.10).*

Proof. The solution provided by Lemma 2.6 is a pointwise regular solution to (sgTE), namely it satisfies with probability one the identity

$$u(t, x) + \int_0^t \big(b \cdot \nabla u + c u\big)(s, x) ds + \int_0^t \sigma \nabla u(s, x) \circ dW_s = u_0(x)$$

for every $(t,x) \in [0,T] \times \mathbb{R}^d$. Since $\nabla u(t,x)$ is a semimartingale (from the definition of u), the Stratonovich integral is well-defined. Using [56, Theorem 7.10] of differentiation under stochastic integrals one deduces

$$\partial_i u(t,x) + \int_0^t \big(\partial_i b \cdot \nabla u + b \cdot \nabla \partial_i u + u \partial_i c + c \partial_i u\big)(s,x) ds + \int_0^t \sigma \nabla \partial_i u(s,x) \circ dW_s = \partial_i u_0(x),$$

hence, for $i = 1, \dots, d$, we have

$$v_i(t,x) + \int_0^t \Big(\sum_{k=1}^d \partial_i b_k v_k + b \cdot \nabla v_i + u \partial_i c + c v_i \Big)(s,x) ds + \int_0^t \sigma \nabla v_i(s,x) \circ dW_s = v_i(0,x),$$

while for $i = 0$, we obtain just from the solution property

$$v_0(t,x) + \int_0^t \big(b \cdot \nabla v_0 + c v_0\big)(s,x) ds + \int_0^t \sigma \nabla v_0(s,x) \circ dW_s = v_0(0,x).$$

Setting

$$r_0 = 0 \quad \text{and} \quad r_i = \sum_{k=1}^d \partial_i b_k v_k + u \partial_i c \ \text{ for } i = 1, \dots, d,$$

we may write for all $i = 0, 1, \dots, d$

$$v_i(t,x) + \int_0^t \big(b \cdot \nabla v_i + c v_i + r_i\big)(s,x) ds + \int_0^t \sigma \nabla v_0(s,x) \circ dW_s = v_0(0,x).$$

Now we use Itô formula in Stratonovich form, see [56, Theorem 8.3], to get

$$\begin{aligned} v_I(t,x) &= v_I(0,x) + \sum_{i \in I} \int_0^t v_{I \setminus i}(s,x) \circ dv_i(s,x) \\ &= v_I(0,x) - \sum_{i \in I} \int_0^t v_{I \setminus i}(s,x) \big(b \cdot \nabla v_i + c v_i + r_i\big)(s,x) ds \\ &\quad - \sum_{i \in I} \int_0^t v_{I \setminus i}(s,x) \sigma \nabla v_i(s,x) \circ dW_s. \end{aligned}$$

Moreover, we have $\partial_j v_I = \sum_{i \in I} v_{I \setminus i} \partial_j v_i$, and thus we may rewrite the previous identity as

$$\begin{aligned} v_I(t,x) + \int_0^t \big(b \cdot \nabla v_I + |I|\, c v_I\big)(s,x) ds + \sum_{i \in I} \int_0^t (v_{I \setminus i} r_i)(s,x) ds \\ = v_I(0,x) - \sigma \int_0^t \nabla v_I(s,x) \circ dW_s. \end{aligned} \tag{2.11}$$

By the definition of r_i, for the last integral on the left-hand side of (2.11), it holds

$$\sum_{i \in I} \int_0^t (v_{I \setminus i} r_i)(s,x) ds = \sum_{i \in I, i \neq 0} \sum_{k=1}^d \int_0^t (v_{I \setminus i \cup k} \partial_i b_k)(s,x) ds + \sum_{i \in I, i \neq 0} \int_0^t (v_{I \setminus i \cup 0} \partial_i c)(s,x) ds.$$

Furthermore, before taking expectations, we want to pass in (2.11) from the Stratonovich to the Itô formulation. To this end, we first note (again by [56, Theorem 7.10]) that

$$\partial_j v_I(t,x) + \int_0^t g(s,x) ds = \partial_j v_I(0,x) - \sigma \sum_{k=1}^d \int_0^t \partial_j \partial_k v_I(s,x) \circ dW_s^k$$

for a bounded process g. Hence, for the stochastic integral in (2.11) we find

$$\sigma \sum_{j=1}^d \int_0^t \partial_j v_I(s,x) \circ dW_s^j = \sigma \int_0^t \nabla v_I(s,x) \cdot dW_s + \frac{\sigma}{2} \sum_{j=1}^d \big[\partial_j v_I(\cdot,x), W^j\big]_t$$

$$= \sigma \int_0^t \nabla v_I(s,x) \cdot dW_s - \frac{\sigma^2}{2} \sum_{j=1}^d \int_0^t \partial_j^2 v_I(s,x) ds.$$

We have proved so far

$$\begin{aligned} v_I(t,x) + \int_0^t \big(b \cdot \nabla v_I + |I|\, c v_I\big)(s,x) ds \\ + \sum_{i \in I, i \neq 0} \sum_{k=1}^d \int_0^t (v_{I \setminus i \cup k} \partial_i b_k)(s,x) ds + \sum_{i \in I, i \neq 0} \int_0^t (v_{I \setminus i \cup 0} \partial_i c)(s,x) ds \\ = v_I(0,x) - \sigma \int_0^t \nabla v_I(s,x) dW_s + \frac{\sigma^2}{2} \int_0^t \Delta v_I(s,x) ds. \end{aligned}$$

The process $\nabla v_I(s,x)$ is bounded (via Lemma 2.6), thus $\int_0^t \nabla v_I(s,x) dW_s$ is a martingale. All other terms have also finite expectation, due to estimate (2.2) of Lemma 2.6. Hence, taking expectation, we have

$$\begin{aligned} w_I(t,x) + \int_0^t \big(b \cdot \nabla w_I + |I|\, c w_I\big)(s,x) ds \\ + \sum_{i \in I, i \neq 0} \sum_{k=1}^d \int_0^t (w_{I \setminus i \cup k} \partial_i b_k)(s,x) ds + \sum_{i \in I, i \neq 0} \int_0^t (w_{I \setminus i \cup 0} \partial_i c)(s,x) ds \\ = v_I(0,x) + \frac{\sigma^2}{2} \int_0^t \Delta w_I(s,x) ds. \end{aligned}$$

This identity implies that $w_I(t,x)$ is continuously differentiable in t and that equation (2.10) holds. The proof of the lemma is complete. □

2.7 Estimates for the parabolic deterministic equation

The system for the w_I's is a parabolic deterministic linear system of partial differential equations. In this section we will obtain energy estimates for w_I which will allow us to obtain the desired a priori bounds. The fact that we have a system instead of a single equation will not affect the estimate (to have an idea of what the final parabolic estimate should be, one could think that w_I is independent of I).

For every smooth function $\chi : \mathbb{R}^d \to [0, \infty)$ as in the statement of the Theorem 2.7 we multiply the identity (2.10) by χw_I and get

$$\partial_t(\chi w_I^2) + 2\chi w_I b \cdot \nabla w_I + 2\chi c |I| w_I^2 + 2 \sum_{i \in I, i \neq 0} \sum_{k=0}^d \chi w_I w_{I \setminus i \cup k} \partial_i b_k = \sigma^2 \chi w_I \Delta w_I,$$

where, for a shorter notation, we have set $b_0 \coloneqq c$. From estimate (2.4) of Lemma 2.6 we know that all terms in this identity are integrable on $\mathbb{R}^d$, uniformly in time. Hence,

$$\begin{aligned} &\int_{\mathbb{R}^d} \chi w_I^2(t,x) dx + \sigma^2 \int_0^t \int_{\mathbb{R}^d} \chi |\nabla w_I|^2(s,x) dx ds \\ &= -2\sigma^2 \int_0^t \int_{\mathbb{R}^d} w_I \nabla w_I \cdot \nabla \chi(s,x) dx ds \\ &\quad - 2 \int_0^t \int_{\mathbb{R}^d} \big(\chi w_I b \cdot \nabla w_I + \chi c |I| w_I^2\big)(s,x) dx ds \\ &\quad - 2 \sum_{i \in I, i \neq 0} \sum_{k=0}^d \int_0^t \int_{\mathbb{R}^d} \chi \big(w_I w_{I \setminus i \cup k} \partial_i b_k\big)(s,x) dx ds. \end{aligned}$$

The term with $\partial_i b_k$ would spoil all our efforts of proving estimates depending only on the LPS norm of the coefficients, but fortunately we may integrate by parts that term. This is the first

key ingredient of this second half of the proof of Theorem 2.7. The second key ingredient is the presence of the term $\sigma^2 \int_0^t \int_{\mathbb{R}^d} \chi |\nabla w_I|^2 dxds$, ultimately coming from the passage Stratonovich to Itô formulation plus taking expectation; this allows us to ask as little as possible on b to close the estimates: we may keep first derivatives of w_I on the right-hand-side of the previous identity, opposite to the deterministic case.

Before starting with the estimates, we recall that $b = b^{(1)} + b^{(2)}$ and $c = c^{(1)} + c^{(2)}$ are assumed to be the sum of two smooth vector fields. Since the desired estimates in Theorem 2.7 differ for the rough part $b^{(1)}$ and the regular (but possibly with linear growth) part $b^{(2)}$, we now split b and c and use the integration by parts formula, in the following way: when a term with $\partial_i b^{(1)}$ appears, we bring the derivative on the other terms; when we have $b^{(2)}$ multiplied by the derivative of some object, we bring the derivative on $b^{(2)}$. In this way we obtain

$$\int_{\mathbb{R}^d} \chi w_I^2(t,x)dx + \sigma^2 \int_0^t \int_{\mathbb{R}^d} \chi |\nabla w_I|^2(s,x)dxds = A_{I,0} + A_{I,1}^{(1)} + A_{I,1}^{(2)} + A_{I,2}^{(1)} + A_{I,2}^{(2)},$$

where we have defined

$$\begin{aligned}
A_{I,0} &:= -2\sigma^2 \int_0^t \int_{\mathbb{R}^d} \big(w_I \nabla w_I \cdot \nabla \chi\big)(s,x)dxds \\
A_{I,1}^{(1)} &:= -2 \int_0^t \int_{\mathbb{R}^d} \big(\chi w_I b^{(1)} \cdot \nabla w_I + \chi c^{(1)} |I| w_I^2\big)(s,x)dxds \\
A_{I,1}^{(2)} &:= \int_0^t \int_{\mathbb{R}^d} \big(\chi \operatorname{div} b^{(2)} w_I^2 + \nabla \chi \cdot b^{(2)} w_I^2 - 2\chi c^{(2)} |I| w_I^2\big)(s,x)dxds \\
A_{I,2}^{(1)} &:= 2 \sum_{i \in I, i \neq 0} \sum_{k=0}^d \int_0^t \int_{\mathbb{R}^d} \big(\big(\partial_i \chi w_I w_{I \setminus i \cup k} + \chi \partial_i w_I w_{I \setminus i \cup k} + \chi w_I \partial_i w_{I \setminus i \cup k}\big) b_k^{(1)}\big)(s,x)dxds \\
A_{I,2}^{(2)} &:= -2 \sum_{i \in I, i \neq 0} \sum_{k=0}^d \int_0^t \int_{\mathbb{R}^d} \big(\chi w_I w_{I \setminus i \cup k} \partial_i b_k^{(2)}\big)(s,x)dxds.
\end{aligned}$$

To estimate these terms we essentially use the following consequence of Hölder's inequality

$$\int_0^t \int_{\mathbb{R}^d} \big(fgh\big)(s,x)dxds \leq \int_0^t \|f\|_\infty(s) \int_{\mathbb{R}^d} \big(|g|^2 + |h|^2\big)(s,x)dxds \tag{2.12}$$

for functions f, g, h defined over $[0,T] \times \mathbb{R}^d$ such that the relevant integrals are well-defined. Moreover, we shall use at several instances the estimate (2.8) on $|\nabla \chi|$, and we further drop the notation (s,x) inside the integrals. For the first term we now employ (2.12) with $f \equiv 1$ (the special case of Hölder's inequality), $g = \sqrt{\varepsilon \chi} |\nabla w_I|$ and $h = 2\sigma^2 \sqrt{\varepsilon^{-1} \chi} |w_I|$ for an arbitrary positive number $\varepsilon > 0$ to find

$$A_{I,0} \leq \varepsilon \int_0^t \int_{\mathbb{R}^d} \chi |\nabla w_I|^2 dxds + C_{\varepsilon,\sigma,\chi} \int_0^t \int_{\mathbb{R}^d} \chi |w_I|^2 dxds.$$

Similarly for the second term, we have

$$A_{I,1}^{(1)} \leq \varepsilon \int_0^t \int_{\mathbb{R}^d} \chi |\nabla w_I|^2 dxds + C_\varepsilon \int_0^t \int_{\mathbb{R}^d} (|b^{(1)}|^2 + |c^{(1)}||I|) \chi w_I^2 dxds.$$

Next, with $g = h = \sqrt{\chi} w_I$ and f chosen as $\operatorname{div} b^{(2)}$, $b^{(2)} C_\chi (1 + |x|)^{-1}$ and $2|c^{(2)}||I|$, respectively, we obtain via (2.8) the estimate

$$\begin{aligned}
A_{I,1}^{(2)} &\leq \int_0^t \int_{\mathbb{R}^d} \big(\big(\operatorname{div} b^{(2)} - 2c^{(2)}|I|\big) \chi w_I^2\big) dxds + C_\chi \int_0^t \int_{\mathbb{R}^d} w_I^2 \frac{|b^{(2)}|}{1+|x|} \chi dxds \\
&\leq \int_0^t \Big(\|\operatorname{div} b^{(2)}(s)\|_\infty + C_\chi \Big\| \frac{b^{(2)}(s,\cdot)}{1+|\cdot|} \Big\|_\infty + 2|I| \|c^{(2)}(s)\|_\infty \Big) \int_{\mathbb{R}^d} \chi w_I^2 dxds.
\end{aligned}$$

Similarly as for the terms $A^{(2)}_{I,1}$ and $A^{(2)}_{I,1}$, we now proceed for the remaining terms $A^{(2)}_{I,2}$ and $A^{(2)}_{I,2}$, with the main difference that w_I eventually needs to be replaced with $w_{I\setminus i\cup k}$. In this way, we get

$$\begin{aligned}
A^{(1)}_{I,2} &\le 2\sum_{i\in I, i\neq 0}\sum_{k=0}^{d}\int_0^t\int_{\mathbb{R}^d}\big(\chi|w_{I\setminus i\cup k}||\partial_i w_I| + \chi|w_I||\partial_i w_{I\setminus i\cup k}| + C_\chi|w_{I\setminus i\cup k}||w_I|\chi\big)|b^{(1)}_k|dxds\\
&\le \varepsilon\sum_{i\in I, i\neq 0}\sum_{k=0}^{d}\int_0^t\int_{\mathbb{R}^d}\big(\chi|\partial_i w_I|^2 + \chi|\partial_i w_{I\setminus i\cup k}|^2\big)dxds\\
&\quad + C_{\varepsilon,\chi}\sum_{i\in I, i\neq 0}\sum_{k=0}^{d}\int_0^t\int_{\mathbb{R}^d}\big(\chi w^2_{I\setminus i\cup k} + \chi w^2_I\big)|b^{(1)}_k|^2dxds
\end{aligned}$$

and finally

$$A^{(2)}_{I,2} \le \sum_{i\in I, i\neq 0}\sum_{k=0}^{d}\int_0^t \|\nabla b^{(2)}_k(s)\|_\infty\int_{\mathbb{R}^d}\chi\big(w_I^2 + w^2_{I\setminus i\cup k}\big)dxds.$$

Given $m\in\mathbb{N}$, we abbreviate $\theta_m = (\sum_{|I|=m}\chi w_I^2)^{1/2}$ and $\rho_m = (\sum_{|I|=m}\sum_{i=1}^d\chi|\partial_i w_I|^2)^{1/2}$. Collecting the previous estimates and summing over $|I|=m$, we have proved so far

$$\begin{aligned}
&\int_{\mathbb{R}^d}\theta_m^2(t,x)dx + \sigma^2\int_0^t\int_{\mathbb{R}^d}\rho_m^2dxdr \le \sum_{|I|=m}\big(A_{I,0} + A^{(1)}_{I,1} + A^{(2)}_{I,1} + A^{(1)}_{I,2} + A^{(2)}_{I,2}\big)\\
&\le \varepsilon\int_0^t\int_{\mathbb{R}^d}\rho_m^2dxds + C_\varepsilon\int_0^t\int_{\mathbb{R}^d}\big(|b^{(1)}|^2 + m|c^{(1)}|\big)\theta_m^2dxds\\
&\quad + \int_0^t\Big(\|\operatorname{div} b^{(2)}(s)\|_\infty + C_\chi\Big\|\frac{b^{(2)}(s,\cdot)}{1+|\cdot|}\Big\|_\infty + 2m\|c^{(2)}(s)\|_\infty\Big)\int_{\mathbb{R}^d}\theta_m^2dxds\\
&\quad + 2\varepsilon C_{m,d}\int_0^t\int_{\mathbb{R}^d}\rho_m^2dxds + 2C_{\varepsilon,\chi}C_{m,d}\int_0^t\int_{\mathbb{R}^d}\theta_m^2\big(|b^{(1)}|^2 + (c^{(1)})^2\big)dxds\\
&\quad + 2C_{m,d}\int_0^t\big(\|\nabla b^{(2)}(s)\|_\infty + \|\nabla c^{(2)}(s)\|_\infty\big)\int_{\mathbb{R}^d}\theta_m^2dxds,
\end{aligned}$$

where we have repeatedly employed the identities $\sum_{|I|=m}\sum_{i\in I}\sum_{k=0}^d\chi|w_{I\setminus i\cup k}|^2 = m(d+1)\theta_m$ and $\sum_{|I|=m}\sum_{i\in I}\sum_{k=0}^d\chi|\partial_i w_{I\setminus i\cup k}|^2 = m(d+1)\rho_m$ (since every J of length m is counted $m(d+1)$ times in the previous sum); so here $C_{m,d} = m(d+1)$. We can then continue to estimate (using Hölder inequality for $m|c^{(1)}|$)

$$\begin{aligned}
&\le \varepsilon(1+4C_{m,d})\int_0^t\int_{\mathbb{R}^d}\rho_m^2dxds\\
&\quad + C_{\varepsilon,m,d,\chi}\int_0^t\int_{\mathbb{R}^d}\big(|b^{(1)}|^2 + (c^{(1)})^2 + 1\big)\theta_m^2dxds\\
&\quad + C_{m,d}\int_0^t\Big(\|\nabla b^{(2)}(s)\|_\infty + C_\chi\Big\|\frac{b^{(2)}(s,\cdot)}{1+|\cdot|}\Big\|_\infty + \|c^{(2)}(s)\|_\infty + \|\nabla c^{(2)}(s)\|_\infty\Big)\int_{\mathbb{R}^d}\theta_m^2dxds
\end{aligned}$$

for new positive constants $C_{\varepsilon,m,d,\chi}, C_{m,d}$ (which incorporate the m inside the integrals). We choose ε so small that $\varepsilon(1+4C_{m,d})\le\sigma^2$ and rename $C_{\varepsilon,m,d,\chi}$ by $C_{m,d,\sigma}$. Therefore, we end up with the preliminary estimate

$$\begin{aligned}
&\int_{\mathbb{R}^d}\theta_m^2(t,x)dx + \frac{\sigma^2}{2}\int_0^t\int_{\mathbb{R}^d}\rho_m^2dxdr \qquad (2.13)\\
&\le C_{m,d,\sigma,\chi}\int_0^t\int_{\mathbb{R}^d}\big(|b^{(1)}|^2 + (c^{(1)})^2 + 1\big)\theta_m^2dxds\\
&\quad + C_{m,d}\int_0^t\Big(\|\nabla b^{(2)}(s)\|_\infty + C_\chi\Big\|\frac{b^{(2)}(s,\cdot)}{1+|\cdot|}\Big\|_\infty + \|c^{(2)}(s)\|_\infty + \|\nabla c^{(2)}(s)\|_\infty\Big)\int_{\mathbb{R}^d}\theta_m^2dxds.
\end{aligned}$$

2.8 End of the proof of Theorem 2.7

Starting from the previous inequality (2.13), we can now continue to estimate its right-hand side by taking into account the LPS-condition on b and c. To this end, we need to distinguish the three cases $(p,q) = (\infty, 2)$, $(p,q) \in (2,\infty)$ and $(p,q) = (d,\infty)$. The main difficulty will be to estimate the term

$$\int_0^t \int_{\mathbb{R}^d} \big(|b^{(1)}|^2 + (c^{(1)})^2\big)\theta_m^2 dxds.$$

From the resulting inequality we can then conclude the proof of Theorem 2.7 via the Gronwall lemma. For the sake of simplicity, let us first restrict ourselves to the important particular case where $b^{(1)}$ and $c^{(1)}$ can be estimated in the L^∞-topology.

Proof of Theorem 2.7 *in the case* $(p,q) = (\infty, 2)$. Here we have

$$\begin{aligned}
&\int_{\mathbb{R}^d} \theta_m^2(t,x)dx + \frac{\sigma^2}{2}\int_0^t \int_{\mathbb{R}^d} \rho_m^2 dxdr \\
&\leq C_{m,d,\sigma,\chi} \int_0^t \big(\|b^{(1)}(s)\|_\infty^2 + \|c^{(1)}(s)\|_\infty^2 + 1\big) \int_{\mathbb{R}^d} \theta_m^2 dxds \\
&\quad + C_{m,d} \int_0^t \Big(\|\nabla b^{(2)}(s)\|_\infty + C_\chi \Big\| \frac{b^{(2)}(s,\cdot)}{1+|\cdot|}\Big\|_\infty + \|c^{(2)}(s)\|_\infty + \|\nabla c^{(2)}(s)\|_\infty\Big) \int_{\mathbb{R}^d} \theta_m^2 dxds.
\end{aligned}$$

Thus, we deduce via Gronwall's lemma that there is a constant $C_0 = C_0(m,d,\sigma,b^{(1)},b^{(2)},c^{(1)},c^{(2)})$, which depends on $b^{(1)}, b^{(2)}, c^{(1)}, c^{(2)}$ through the norms

$$\begin{aligned}
&\int_0^T \|b^{(1)}(s)\|_\infty^2 ds, \qquad &&\int_0^T \Big(\Big\| \frac{b^{(2)}(s,\cdot)}{1+|\cdot|}\Big\|_\infty + \|\nabla b^{(2)}(s)\|_\infty\Big) ds, \\
&\int_0^T \|c^{(1)}(s)\|_\infty^2 ds, \qquad &&\int_0^T \big(\|c^{(2)}(s)\|_\infty + \|\nabla c^{(2)}(s)\|_\infty\big) ds
\end{aligned}$$

such that

$$\sup_{t\in[0,T]} \int_{\mathbb{R}^d} \theta_m^2(t,x)dx \leq C_0 \int_{\mathbb{R}^d} \theta_m^2(0,x)dx. \tag{2.14}$$

We then notice that by the definition of θ_m and by Young's inequality, there holds

$$\sum_{i=0}^d \int_{\mathbb{R}^d} E[(\partial_i u(t,x))^m]^2 \chi(x)dx \leq \int_{\mathbb{R}^d} \theta_m^2(t,x)dx \leq C_{m,d} \sum_{i=0}^d \int_{\mathbb{R}^d} E[|\partial_i u(t,x)|^m]^2 \chi(x)dx$$

for some constant $C_{m,d} > 0$, hence

$$\sup_{t\in[0,T]} \sum_{i=1}^d \int_{\mathbb{R}^d} E[(\partial_i u(t,x))^m]^2 \chi(x)dx \leq C_0 \int_{\mathbb{R}^d} \theta_m^2(0,x)dx \leq C_0 C_{m,d} \|u_0\|^{2m}_{W^{1,2m}_\chi(\mathbb{R}^d)}.$$

This finishes the proof of Theorem 2.7 in the case $(p,q) = (\infty, 2)$. □

Let us come to the general case. Notice that it is only here, for the first and only time, that the exponents (p,q) of the LPS condition enter. By $\|\cdot\|_{W^{1,2}}$ and $\|\cdot\|_{L^p}$ we denote the usual norms in $W^{1,2}(\mathbb{R}^d)$ and $L^p(\mathbb{R}^d)$ respectively. We first prove the following technical lemma, which will be relevant to continue with the estimate for the terms on the right-hand side of inequality (2.13) for the general case $p \neq \infty$.

Lemma 2.11. *If $p > d \vee 2$, then for every $\varepsilon > 0$ there is a constant $C_\varepsilon > 0$, depending only on p, d and ε, such that for all $f, g \in C_c^\infty(\mathbb{R}^d)$ we have*

$$\int_{\mathbb{R}^d} |f(x)g(x)|^2 dx \leq \varepsilon \|g\|^2_{W^{1,2}} + C_\varepsilon \|f\|^{\frac{2p}{p-d}}_{L^p} \|g\|^2_{L^2}. \tag{2.15}$$

If $p = d \geq 3$, we have

$$\int_{\mathbb{R}^d} |f(x)g(x)|^2 dx \leq C_d \|f\|^2_{L^d} \|\nabla g\|^2_{L^2} \tag{2.16}$$

with a constant $C_d > 0$ depending only on d.

Proof. Let us start by recalling Gagliardo–Nirenberg interpolation inequality on $\mathbb{R}^d$, $d \neq 2$: for every $0 \leq \beta \leq 1$ and $\alpha \geq 2$ which satisfy

$$\frac{1}{\alpha} = \frac{1}{2} - \frac{\beta}{d}$$

the following holds: there exists a constant $C > 0$ depending only on β and d such that every $g \in W^{1,2}(\mathbb{R}^d)$ belongs to $L^\alpha(\mathbb{R}^d)$ with

$$\|g\|_{L^\alpha} \leq C \|g\|_{L^2}^{1-\beta} \|\nabla g\|_{L^2}^{\beta}.$$

The result is true also for $d = 2$ but requires the additional condition $\beta < 1$. We apply this inequality with $\beta = \frac{d}{p}$, $\alpha = \frac{2p}{p-2}$. The assumptions of Gagliardo–Nirenberg inequality are satisfied because $\beta \leq 1$ for $d \neq 2$ and $\beta < 1$ for $d = 2$. Then

$$\|g\|_{L^{\frac{2p}{p-2}}} \leq C \|g\|_{L^2}^{1-\frac{d}{p}} \|\nabla g\|_{L^2}^{\frac{d}{p}}.$$

Now we come the the proof of the lemma. We first apply Hölder's inequality with exponents $\frac{p}{2}$ and $\frac{p}{p-2}$ and then the previous inequality to find

$$\int_{\mathbb{R}^d} |fg|^2 dx \leq \|f\|_{L^p}^2 \|g\|_{L^{\frac{2p}{p-2}}}^2 \leq C \|f\|_{L^p}^2 \|g\|_{L^2}^{2(1-\frac{d}{p})} \|\nabla g\|_{L^2}^{2\frac{d}{p}},$$

which is the claim (2.16) for $p = d$. For $p > d$, we use Young's inequality

$$ab \leq \frac{a^r}{r} + \frac{b^{r'}}{r'}, \qquad r, r' > 1, \qquad \frac{1}{r} + \frac{1}{r'} = 1$$

with

$$r = \frac{p}{d}, \qquad a = (r\varepsilon)^{\frac{d}{p}} \|\nabla g\|_{L^2}^{2\frac{d}{p}}, \qquad b = (r\varepsilon)^{-\frac{d}{p}} C \|f\|_{L^p}^2 \|g\|_{L^2}^{2(1-\frac{d}{p})}.$$

With $r' = \frac{p}{p-d}$ we get

$$\begin{aligned}\int_{\mathbb{R}^d} |fg|^2 dx &\leq \varepsilon \|\nabla g\|_{L^2}^2 + \frac{(r\varepsilon)^{-\frac{d}{p}r'} C^{r'}}{r'} \|f\|_{L^p}^{2r'} \|g\|_{L^2}^{2(1-\frac{d}{p})r'} \\ &= \varepsilon \|\nabla g\|_{L^2}^2 + p^{-1}(p-d)(p\varepsilon/d)^{-\frac{d}{p-d}} C^{\frac{p}{p-d}} \|f\|_{L^p}^{\frac{2p}{p-d}} \|g\|_{L^2}^2,\end{aligned}$$

and thus, we have found C_ε such that (2.15) holds. This concludes the proof. □

Lemma 2.12. *If $b \in \mathcal{LPS}(p,q)$ with $q < \infty$ (hence $p > d$), then*

$$\int_0^T \Big(\int_{\mathbb{R}^d} |b(s,x)|^p dx \Big)^{\frac{2}{p-d}} ds \leq \|b\|_{L^q([0,T];L^p)}^q. \tag{2.17}$$

Proof. From $\frac{2}{q} \leq 1 - \frac{d}{p} = \frac{p-d}{p}$ we see $\frac{2}{p-d} \leq \frac{q}{p}$. Therefore, the assumption $b \in \mathcal{LPS}(p,q)$ with $q < \infty$ implies

$$\int_0^T \Big(\int_{\mathbb{R}^d} |b(s,x)|^p dx \Big)^{\frac{2}{p-d}} ds \leq T^{1-\frac{2p}{q(p-d)}} \Big(\int_0^T \Big(\int_{\mathbb{R}^d} |b(s,x)|^p dx \Big)^{\frac{q}{p}} ds \Big)^{\frac{2p}{q(p-d)}} < \infty. \qquad \square$$

The previous interpolation Lemma 2.11 now allows us to continue with the proof of Theorem 2.7 in the remaining cases.

Proof of Theorem 2.7 *in the case* $(p,q) \in (2,\infty)$. We start by observing

$$|\partial_i \theta_m| \leq \frac{1}{2\theta_m} \sum_{|I|=m} |\partial_i \chi| w_I^2 + \frac{1}{\theta_m} \sum_{|I|=m} \chi |w_I| |\partial_i w_I|$$

$$\leq \frac{C_\chi}{2\theta_m} \sum_{|I|=m} \chi w_I^2 + \sqrt{\sum_{|I|=m} \chi |\partial_i w_I|^2} \leq \frac{C_\chi}{2} \sqrt{\theta_m} + \sqrt{\rho_m},$$

and thus

$$\|\theta_m\|_{W^{1,2}}^2 = \|\theta_m\|_{L^2}^2 + \sum_{i=1}^d \|\partial_i \theta_m\|_{L^2}^2 \leq C\|\theta_m\|_{L^2}^2 + C\|\rho_m\|_{L^2}^2$$

for some constant C depending only on χ, d, m. Therefore, the application of inequality (2.15) to the terms of the second line of inequality (2.13) shows

$$\begin{aligned}\int_0^t \int_{\mathbb{R}^d} |b^{(1)}|^2 \theta_m^2 dx ds &\leq \int_0^t \Big(\varepsilon \|\theta_m\|_{W^{1,2}}^2 + C_\varepsilon \|b^{(1)}\|_{L^p}^{\frac{2p}{p-d}} \|\theta_m\|_{L^2}^2 \Big) ds \\ &\leq \varepsilon C \int_0^t \|\rho_m\|_{L^2}^2 ds + C'_\varepsilon \int_0^t \Big(1 + \|b^{(1)}\|_{L^p}^{\frac{2p}{p-d}} \Big) \|\theta_m\|_{L^2}^2 ds\end{aligned}$$

for some constant $C'_\varepsilon > 0$. We use this inequality and the similar one for $c^{(1)}$ in the second line of inequality (2.13) and get, for ε small enough and by means of the Gronwall lemma (applicable because of the inequality (2.17)), a bound of the form (2.14). With the final arguments used above in the case $(p, q) = (\infty, 2)$, this completes the proof of Theorem 2.7 in the case $p, q \in (2, \infty)$. □

Proof of Theorem 2.7 *in the case* $(p, q) = (d, \infty)$. In this case we apply inequality (2.16) to the terms of the second line of inequality (2.13) to find

$$\int_0^t \int_{\mathbb{R}^d} |b^{(1)}|^2 \theta_m^2 dx ds \leq C_d \int_0^t \|b^{(1)}\|_{L^d}^2 \|\nabla \theta_m\|_{L^2}^2 ds$$

and an analogous inequality for the term with $c^{(1)}$. We then estimate $\|\nabla \theta_m\|_{L^2}^2$ as above by $C^*\|\theta_m\|_{L^2}^2 + C^*\|\rho_m\|_{L^2}^2$ and get

$$\begin{aligned}&\int_{\mathbb{R}^d} \theta_m^2(t,x) dx + \frac{\sigma^2}{2} \int_0^t \int_{\mathbb{R}^d} \rho_m^2 dx dr \\ &\leq C_{m,d,\sigma} C_d C^* \int_0^t \Big(\|b^{(1)}\|_{L^d}^2 + \|c^{(1)}\|_{L^d}^2 \Big) \|\rho_m\|_{L^2}^2 ds \\ &\quad + C_{m,d} \int_0^t \Big(\|\nabla b^{(2)}(s)\|_\infty + C_\chi \Big\| \frac{b^{(2)}(s,\cdot)}{1+|\cdot|} \Big\|_\infty + \|c^{(2)}(s)\|_\infty + \|\nabla c^{(2)}(s)\|_\infty + 2 \Big) \int_{\mathbb{R}^d} \theta_m^2 dx ds.\end{aligned}$$

If the smallness condition

$$2C_{m,d,\sigma} C_d C^* \big(\sup_{t\in[0,T]} \|b^{(1)}\|_{L^d}^2 + \sup_{t\in[0,T]} \|c^{(1)}\|_{L^d}^2 \big) \leq \sigma^2 \tag{2.18}$$

is satisfied, we may again apply the Gronwall lemma and the other computations above to conclude the proof of Theorem 2.7 also in the remaining case $(p, q) = (d, \infty)$. □

2.9 Existence of global regular solutions for sTE and sCE

In this section we deduce, from the a priori estimates of Theorem 2.7, the existence of global regular solutions for the stochastic generalized equation (sgTE) and consequently also for sTE and sCE. This can be interpreted, at least for the sTE, as a no-blow-up result. Uniqueness will be treated separately in the next section, see also Remark 2.18 below.

In what follows, we assume that the LPS-integrability condition on b, c with exponents $p \in [d, \infty]$ and $q \in [2, \infty]$ as stated in Section 2.1 is satisfied. We further denote by $p' = p/(p-1)$ the conjugate exponent of p (with $p' = 1$ if $p = \infty$). We now start by defining the notion of solutions of class $L^\theta(W^{1,m}_{\mathrm{loc}})$ of equation (sgTE), for some $\theta \geq 2$ and $m \geq p'$. To this end, we require first of all some measurability and continuous semimartingale properties for terms appearing in (sgTE) after testing against smooth functions. We say that a map $u\colon [0,T] \times \mathbb{R}^d \times \Omega \to \mathbb{R}$ is *weakly progressively measurable* with respect to $(\mathcal{G}_t)_t$ if $x \mapsto u(t,x,\omega) \in L^1_{\mathrm{loc}}(\mathbb{R}^d)$ for a.e. (t,ω) and the

process $(t,\omega) \mapsto \int_{\mathbb{R}^d} u(t,x,\omega)\varphi(x)dx$ is progressively measurable with respect to $(\mathcal{G}_t)_t$, for every $\varphi \in C_c^\infty(\mathbb{R}^d)$. Secondly, we need all relevant integrals to be well-defined. Due to the choice $\theta \geq 2$ we have well-defined stochastic integrals; hence we only need to take care that $b(s)\cdot\nabla u(s)$ and $c(s)u(s)$ are in $L^1_{\mathrm{loc}}(\mathbb{R}^d)$ for a.e. $s \in [0,T]$. Keeping in mind the decompositions $b = b^{(1)} + b^{(2)}$ and $c = c^{(1)} + c^{(2)}$ into (roughly) a vector field of LPS class and a Lipschitz function, we first note that with $\theta \geq 2$ we also have $\theta \geq q/(q-1)$ (recalling $q \geq 2$). Therefore, $s \mapsto \langle b(s)\cdot\nabla u(s) - c(s)u(s), \varphi\rangle$ is integrable according to the choice $m \geq p'$ (here, the symbol $\langle\cdot,\cdot\rangle$ stands for the usual inner product in $L^2(\mathbb{R}^d)$).

These introductory comments now motivate the following definition.

Definition 2.13. *Given $\theta \geq 2$, $m \geq p'$,* a solution of equation (sgTE) of class $L^\theta(W^{1,m}_{\mathrm{loc}})$ *is a map $u : [0,T]\times\mathbb{R}^d\times\Omega\to\mathbb{R}$ with the following properties:*

(o) *it is weakly progressively measurable with respect to $(\mathcal{G}_t)_t$;*

(i) *u is in $L^\theta([0,T]\times\Omega; W^{1,m}(B_R))$ for every $R>0$;*

(ii) *$t \mapsto \langle u(t),\varphi\rangle$ has a modification that is a continuous semimartingale, for every φ in $C_c^\infty(\mathbb{R}^d)$;*

(iii) *for every φ in $C_c^\infty(\mathbb{R}^d)$, for this modification (still denoted by $\langle u(t),\varphi\rangle$) we have uniformly in $t\in[0,T]$, with probability one,*

$$\langle u(t),\varphi\rangle = \langle u_0,\varphi\rangle - \int_0^t \langle b(s)\cdot\nabla u(s) + c(s)u(s),\varphi\rangle\, ds + \sigma\sum_{i=1}^d \int_0^t \langle u(s),\partial_i\varphi\rangle \circ dW^i_s. \quad (2.19)$$

As mentioned above we will now prove the existence of such solutions by exploiting the a priori Sobolev-type estimates for solutions to approximate equations with smooth coefficients. The crucial point is that the estimates only depend on the LPS norms of the coefficients b and c, but not on the approximation itself. Hence, from the regular solutions to these approximate equations we may then pass to a limit function which still has the same Sobolev-type regularity, provided that the approximate coefficients remain bounded in these norms. In a second step we then need to verify that the limit function is indeed a solution to the original equation in the sense of Definition 2.13.

Concerning the approximation of the coefficients, we first observe that, since b and c are assumed to belong to the LPS class (satisfying Condition 2.2), we may choose sequences $(b_\varepsilon)_\varepsilon$, $(c_\varepsilon)_\varepsilon$ which verify the following assumptions:

Condition 2.14. *We assume $b_\varepsilon = b^{(1)}_\varepsilon + b^{(2)}_\varepsilon$, $c_\varepsilon = c^{(1)}_\varepsilon + c^{(2)}_\varepsilon$, such that:*

- *$(b^{(1)}_\varepsilon)_\varepsilon$ is a $C_c^\infty([0,T]\times\mathbb{R}^d)$ approximation of $b^{(1)}$ a.e. and in LPS, in the following sense: if $p,q\in(2,\infty)$, then $b^{(1)}_\varepsilon \to b^{(1)}$ a.e. in $[0,T]\times\mathbb{R}^d$ and in $L^q([0,T];L^p(\mathbb{R}^d))$ as $\varepsilon\to 0$; otherwise, if p or q is ∞, then $b^{(1)}_\varepsilon \to b^{(1)}$ a.e. in $[0,T]\times\mathbb{R}^d$ as $\varepsilon\to 0$ and, for every $\varepsilon>0$, $\|b^{(1)}_\varepsilon\|_{L^q([0,T];L^p(\mathbb{R}^d))} \leq 2\|b^{(1)}\|_{L^q([0,T];L^p(\mathbb{R}^d))}$;*
- *in case of Condition 2.2, 1b), the $\|b^{(1)}_\varepsilon\|_{C([0,T];L^d(\mathbb{R}^d))}$ norms are small enough, uniformly in ε (in case of Condition 2.2 1c), this follows from the previous assumption);*
- *$(c^{(1)}_\varepsilon)_\varepsilon$ is as $(b^{(1)}_\varepsilon)_\varepsilon$ (with $c^{(1)}$ in place of $b^{(1)}$);*
- *$(b^{(2)}_\varepsilon)_\varepsilon$ is a $C_c^\infty([0,T]\times\mathbb{R}^d)$ approximation of $b^{(2)}$ a.e. and in $L^2(C^1_{lin})$, in the following sense: $b^{(2)}_\varepsilon \to b^{(2)}$ a.e. in $[0,T]\times\mathbb{R}^d$ and in $L^2([0,T];C^1_{lin}(\mathbb{R}^d))$ as $\varepsilon\to 0$;*
- *$(c^{(2)}_\varepsilon)_\varepsilon$ is a $C_c^\infty([0,T]\times\mathbb{R}^d)$ approximation of $c^{(2)}$ a.e. in $[0,T]\times\mathbb{R}^d$ and in $L^2(C^1_b)$, in the following sense: $c^{(2)}_\varepsilon \to c^{(2)}$ a.e. in $[0,T]\times\mathbb{R}^d$ and in $L^2([0,T];C^1_b(\mathbb{R}^d))$ as $\varepsilon\to 0$.*

Remark 2.15. *Let us briefly explain how Condition 2.14 allows us to treat general coefficients $b^{(1)}$ and $c^{(1)}$ in $C([0,T];L^d(\mathbb{R}^d))$ (see Condition 2.2, 1b)), without imposing a smallness condition of the associated norm as for the case of coefficients in $L^\infty([0,T];L^d(\mathbb{R}^d))$. In fact, we can rewrite any coefficients $b^{(1)}$ in $C([0,T];L^d(\mathbb{R}^d))$ as a sum of a regular, compactly supported term (say f) and the*

remaining, possibly irregular term $b^{(1)} - f$, whose $C([0,T];L^d(\mathbb{R}^d))$ norm can be made arbitrarily small as a consequence of the density of $C([0,T];C_c^\infty(\mathbb{R}^d))$-functions in $C([0,T];L^d(\mathbb{R}^d))$. Thus we can approximate $b^{(1)} - f$ with $(b_\varepsilon^{(1)})_\varepsilon$ and $f + b^{(2)}$ with $(b_\varepsilon^{(2)})_\varepsilon$ (analogously c), which in turn ensures that Condition 2.14 is fulfilled, in particular the smallness of the norm of $b_\varepsilon^{(1)}$.

Remark 2.16. *Notice that, in any case of Condition 2.2 (or also for more general b's), the family $(b_\varepsilon)_\varepsilon$ converges to b in $L^1([0,T];L^{m'}(B_R))$; the same holds for c.*

Theorem 2.17. *Let $m \geq 2$ be an even integer and let $s \in \mathbb{R}$. Assume that b, c satisfy Condition 2.2 and let $u_0 \in W^{1,2m}_{(1+|\cdot|)^{2s+d+1}}(\mathbb{R}^d)$. Then there exists a solution u of equation (sgTE) of class $L^m(W^{1,m}_{\mathrm{loc}})$, which further satisfies $u(t,\cdot) \in W^{1,m}_{(1+|\cdot|)^s}(\mathbb{R}^d)$ for a.e. (t,ω). Moreover, there holds*

$$\operatorname*{ess\,sup}_{t\in[0,T]} E\Big[\|u(t,\cdot)\|^m_{W^{1,m}_{(1+|\cdot|)^s}(\mathbb{R}^d)}\Big] < \infty. \tag{2.20}$$

Proof. Step 1: Compactness argument. Take $(b_\varepsilon)_\varepsilon$ and $(c_\varepsilon)_\varepsilon$ as in Condition 2.14; take $(u_0^\varepsilon)_\varepsilon$ as $C_c^\infty(\mathbb{R}^d)$ approximations of the initial datum u_0, converging to it a.e. in $\mathbb{R}^d$ and in $W^{1,2m}_{(1+|\cdot|)^{2s+d+1}}(\mathbb{R}^d)$. Let u_ε be the regular solution to (sgTE) corresponding to coefficients b_ε, c_ε instead of b, c, and with initial value u_0^ε, given by Lemma 2.6. From Corollary 2.8 (notice that, in the limit case $p = d$, $b_\varepsilon^{(1)}$ is small enough in view of Condition 2.14), we deduce that the family $(u_\varepsilon)_\varepsilon$ is bounded in $L^\infty([0,T];L^m(\Omega;W^{1,m}_{(1+|\cdot|)^s}(\mathbb{R}^d)))$. Hence, by Remark 2.9, we can extract a subsequence (for simplicity the whole sequence), which converges weakly-$*$ to some u in that space; in particular, weak convergence in $L^m([0,T]\times\Omega;W^{1,m}(B_R))$ holds for every $R > 0$, i.e. (i) of Definition 2.13.

Step 2: Weak progressive measurability. Given $\varphi \in C_c^\infty(\mathbb{R}^d)$, the stochastic processes $(t,\omega) \mapsto \langle u_\varepsilon(t),\varphi\rangle$ are progressively measurable, weakly convergent in $L^m([0,T]\times\Omega)$ to $\langle u,\varphi\rangle$ and the space of progressively measurable processes is closed, so weakly closed, in $L^m([0,T]\times\Omega)$. Thus, u is weakly progressively measurable.

Step 3: Passage from Stratonovich to Itô and viceversa. It will be useful to notice that the last two requirements, namely the semimartingale (ii) and the solution property (iii), in Definition 2.13 can be replaced by the following Itô formulation: for every φ in $C_c^\infty(\mathbb{R}^d)$, for a.e. (t,ω), there holds

$$\begin{aligned}\langle u(t),\varphi\rangle = \langle u_0,\varphi\rangle - \int_0^t \langle b(s)\cdot\nabla u(s) + c(s)u(s),\varphi\rangle\, ds \\ + \sigma\sum_{i=1}^d \int_0^t \langle u(s),\partial_i\varphi\rangle\, dW_s^i + \frac{\sigma^2}{2}\int_0^t \langle u(s),\Delta\varphi\rangle ds.\end{aligned} \tag{2.21}$$

Let us prove this fact. Suppose we have the Stratonovich formulation (with (ii) and (iii)). The Stratonovich integral $\sum_{i=1}^d \int_0^t \langle u(s),\partial_i\varphi\rangle \circ dW_s^i$ is well-defined, thanks to (ii) and our integrability assumptions (with $m,\theta \geq 2$), and it is equal to

$$\sum_{i=1}^d \int_0^t \langle u(s),\partial_i\varphi\rangle \circ dW_s^i = \sum_{i=1}^d \int_0^t \langle u(s),\partial_i\varphi\rangle\, dW_s^i + \sum_{i=1}^d \big[\,\langle u(\cdot),\partial_i\varphi\rangle\,, W^i\big]_t$$

where the brackets $[\cdot,\cdot]$ again denote the quadratic covariation. The semimartingale decomposition of $\langle u(t),\partial_i\varphi\rangle$ is taken from the equation for u (just use $\partial_i\varphi$ instead of φ): the martingale part of $\langle u(t),\partial_i\varphi\rangle$ is $\sigma\sum_{j=1}^d \int_0^t \langle u(s),\partial_j\partial_i\varphi\rangle \circ dW_s^j$, so that we have

$$\big[\,\langle u(\cdot),\partial_i\varphi\rangle\,, W^i\big]_t = \sigma\int_0^t \big\langle u(s),\partial_i^2\varphi\big\rangle\, ds. \tag{2.22}$$

Thus, we get precisely formula (2.21) from (iii).

Now suppose we have the Itô formulation (2.21). This implies that $t \mapsto \langle u(t),\varphi\rangle$ has a modification that is a continuous semimartingale (i.e. (ii) in Definition 2.13). The same is true for $t \mapsto \langle u(t),\partial_i\varphi\rangle$ for $i = 1,\dots,d$, and thus the quadratic covariation $[\langle u(\cdot),\partial_i\varphi\rangle, W^i]_t$ and the

Stratonovich integral $\int_0^t \langle u(s), \partial_i \varphi \rangle \circ dW_s^i$ exist; moreover, by the equation itself we again find (2.22). It follows that

$$\sigma \sum_{i=1}^{d} \int_0^t \langle u(s), \partial_i \varphi \rangle \, dW_s^i + \frac{\sigma^2}{2} \int_0^t \langle u(s), \Delta \varphi \rangle \, ds = \sigma \sum_{i=1}^{d} \int_0^t \langle u(s), \partial_i \varphi \rangle \circ dW_s^i$$

and so we deduce (iii) from (2.21).

Step 4: Verification of the equation. We want to show that u satisfies (sgTE), in the sense of distributions. In view of Step 3, we can use the Itô formulation (2.21). Fix φ in $C_c^\infty(\mathbb{R}^d)$ with support in B_R. We already know from Step 2 that $\langle u_\varepsilon(t), \varphi \rangle - \langle u_0^\varepsilon, \varphi \rangle$ converges to $\langle u_t, \varphi \rangle - \langle u_0, \varphi \rangle$ weakly in $L^2([0,T] \times \Omega)$. We will prove that also the other terms in (2.21) converge, weakly in $L^1([0,T] \times \Omega)$. The idea for the convergence is the following: assume we have a linear continuous map $G = G(u)$ between two Banach spaces and a bilinear map $F = F(b,u)$ mapping from two suitable Banach spaces into a third one; then, if b_ε converges to b strongly and u_ε converges to u weakly in the associated topologies, $G(u_\varepsilon)$ and $F(b_\varepsilon, u_\varepsilon)$ converge weakly to $G(u)$ and $F(b,u)$, respectively.

For the term $\int_0^t \langle b(s) \nabla u(s), \varphi \rangle ds$, we take

$$F \colon L^1\big([0,T]; L^{m'}(B_R)\big) \times L^\infty\big([0,T]; L^m(\Omega; W^{1,m}(B_R))\big) \to L^1([0,T] \times \Omega)$$
$$F(b,u)(t,\omega) \coloneqq \int_0^t \langle b(s) \nabla u(s)(\omega), \varphi \rangle ds;$$

Then F is a bilinear continuous map. Fix Z in $L^\infty([0,T] \times \Omega)$; for the weak $L^1([0,T] \times \Omega)$-convergence we now have to prove that, as $\varepsilon \to 0$,

$$\int_0^T E\big[\big(F(b_\varepsilon, u_\varepsilon) - F(b,u)\big) Z\big] dt \to 0.$$

Since b_ε converges strongly to b in $L^1([0,T]; L^{m'}(B_R))$ (see Remark 2.16) and since u_ε has uniformly (in ε) bounded norm in $L^\infty([0,T]; L^m(\Omega; W^{1,m}(B_R)))$ (according to Step 1), the norm $\|F(b_\varepsilon, u_\varepsilon) - F(b, u_\varepsilon)\|_{L^1([0,T] \times \Omega)}$ is small for ε small, and in particular

$$\int_0^T E[(F(b_\varepsilon, u_\varepsilon) - F(b, u_\varepsilon)) Z] dt \to 0$$

as $\varepsilon \to 0$. It remains to prove that $\int_0^T E[(F(b, u_\varepsilon) - F(b,u)) Z] dt \to 0$ as $\varepsilon \to 0$. For this purpose, we notice that, by the Fubini–Tonelli theorem,

$$\int_0^T E\big[(F(b, u_\varepsilon) - F(b,u)) Z\big] dt = \int_0^T \int_{B_R} E\big[Y(\nabla u_\varepsilon - \nabla u)\big] dx ds,$$

where $Y(s,x,\omega) \coloneqq b(s,x) \varphi(x) \int_s^T Z(t,\omega) dt$. The convergence of the right-hand side now follows easily, since Y is in $L^1([0,T]; L^{m'}(B_R \times \Omega))$ and since ∇u_ε converges weakly-$*$ to ∇u in $L^\infty([0,T]; L^m(B_R \times \Omega))$ (by Step 1), as $\varepsilon \to 0$. This finishes the proof of convergence for F, and the convergence of the term $\int_0^t \langle c(s) u(s), \varphi \rangle ds$ is established analogously.

For the term $\int_0^t \langle u(s), \partial_i \varphi \rangle dW_s^i$, we define $G \colon L^2([0,T] \times B_R \times \Omega) \to L^2([0,T] \times \Omega)$ by

$$G(u)(t,\omega) \coloneqq \int_0^t \langle u(s), \partial_i \varphi \rangle dW_s^i(\omega);$$

G is a linear continuous map, hence weakly continuous. Therefore, as a consequence of the weak convergence $\langle u_\varepsilon(s), \partial_i \varphi \rangle$ to $\langle u(s), \partial_i \varphi \rangle$ in $L^2([0,T] \times \Omega)$, we find that $\int_0^t \langle u_\varepsilon(s), \partial_i \varphi \rangle dW_s^i$ converges weakly (to the obvious limit) in $L^2([0,T] \times \Omega)$. The convergence of the last terms in (2.21) is easier. Thus, the limit function u satisfies the identity (2.21), i.e. it is a solution to (sgTE) in the Itô sense, and via Step 3 the proof of Theorem 2.17 is complete. □

Remark 2.18. *Uniqueness of solutions to* (sgTE) *will be treated later in great generality* (*note that uniqueness of weak solutions of class* $L^m(L^m_{\rm loc})$*, defined in Definition* 3.1*, implies uniqueness of solutions of class* $L^m(W^{1,m}_{\rm loc})$*). However, uniqueness of weak solutions requires the formulation itself of weak solutions, in which we have to assume some integrability of* $\operatorname{div} b$ *which plays no role in Definition* 2.13 *and Theorem* 2.17*. One may ask whether it is possible to prove uniqueness of solutions of class* $L^m(W^{1,m}_{\rm loc})$ *directly, without the theory of weak solutions. The answer is affirmative but we do not repeat the proofs, see for instance* [41, Appendix A].

The previous result holds for (sgTE) and therefore, it covers the sTE by taking $c = 0$. The case of the sCE requires $c = \operatorname{div} b$ and therefore it is better to state explicitly the assumptions. The divergence is understood in the sense of distributions.

Corollary 2.19. *Let* $m \geq 2$ *be an even integer and let* $s \in \mathbb{R}$*. Consider the sCE in the form* (2.1) *under the assumptions of Remark* 2.5 *and let* $u_0 \in W^{1,2m}_{(1+|\cdot|)^{2s+d+1}}(\mathbb{R}^d)$*. Then there exists a solution* u *of equation* (sgTE) *of class* $L^m(W^{1,m}_{\rm loc})$*, which further satisfies* $u(t,\cdot) \in W^{1,m}_{(1+|\cdot|)^s}(\mathbb{R}^d)$ *for a.e.* (t,ω) *and the analogous estimate of* (2.20)*.*

2.10 $W^{2,m}$-regularity

In this section we are interested in the existence of solutions to equation (sgTE) of higher regularity, more precisely of local $W^{2,m}$-regularity in space. To this end we shall essentially follow the strategy of the local $W^{1,m}$-regularity in space presented above. First, we consider second order derivatives of equation (sgTE) (instead of first ones) for the smooth solutions of approximate problems with smooth coefficients and derive a parabolic (deterministic) equation for averages of second order derivatives. For this reason we have to assume some LPS condition not only on the coefficients b and c, but also on their first space derivatives. Once the parabolic equation is derived, we may proceed analogously to above, that is, via the parabolic theory we establish a priori regularity estimates involving second order derivatives, and finally we pass to the limit to get the regularity statement.

Let us now start to clarify the assumptions of this section. As motivated above, we roughly assume that in addition to the coefficients b and c also their first order derivatives $\partial_k b$ and $\partial_k c$ (for $k = 1, \dots, d$) satisfy the assumptions of Section 2.1. More precisely, we assume

Condition 2.20. *The coefficients* b *and* c *can be written as* $b = b^{(1)} + b^{(2)}$*,* $c = c^{(1)} + c^{(2)}$ *with weakly differentiable functions* $b^{(1)}, b^{(2)}, c^{(1)}, c^{(2)}$*, and for every* $k \in \{0, 1, \dots, d\}$ *each of the decompositions* $\partial_k b = \partial_k b^{(1)} + \partial_k b^{(2)}$ *and* $\partial_k c = \partial_k c^{(1)} + \partial_k c^{(2)}$ *satisfies Condition* 2.2*. Note that if Condition* 2.2 1b) *or* 1c) *applies, then we need to assume in addition* $d \geq 3$*.*

We start by deriving, in the smooth setting, suitable a priori estimates involving second order derivatives of the regular solution, following the strategy of Theorem 2.7.

Lemma 2.21. *Let* p, q *be in* $(2, \infty)$ *satisfying* $\frac{2}{q} + \frac{d}{p} \leq 1$ *or* $(p, q) = (\infty, 2)$*, let* m *be positive integer, let* $\sigma \neq 0$*, and let* χ *be a function satisfying* (2.8)*. Assume that* b *and* c *are a vector field and a scalar field, respectively, such that* $b = b^{(1)} + b^{(2)}$*,* $c = c^{(1)} + c^{(2)}$*, with* $b^{(i)}$*,* $c^{(i)}$ *in* $C^\infty_c([0,T] \times \mathbb{R}^d)$ *for* $i = 1, 2$*. Then there exists a constant* C *such that, for every* u_0 *in* $C^\infty_c(\mathbb{R}^d)$*, the smooth solution* u *of equation* (sgTE) *starting from* u_0*, given by Lemma* 2.6*, verifies*

$$\sup_{t\in[0,T]} \sum_{i,j=0}^{d} \int_{\mathbb{R}^d} E\big[(\partial_j \partial_i u(t,x))^m\big]^2 \chi(x)dx \leq C\|u_0\|^{2m}_{W^{2,2m}_\chi(\mathbb{R}^d)}.$$

Here, the constant C *can be chosen similarly as in Theorem* 2.7*, now depending also on the* $L^q([0,T]; L^p(\mathbb{R}^d))$ *norms of* $\partial_k b^{(1)}$ *and* $\partial_k c^{(1)}$*, on the* $L^1([0,T]; C^1_{lin}(\mathbb{R}^d))$ *norms of* $\partial_k b^{(2)}$*, and on the* $L^1([0,T]; C^1_b(\mathbb{R}^d))$ *norms of* $\partial_k c^{(2)}$*, for all* $k \in \{0, 1, \dots, d\}$*.*

The result holds also for $(p, q) = (d, \infty)$*, provided that the* $L^\infty([0,T]; L^d(\mathbb{R}^d))$ *norms of* $\partial_k b^{(1)}$ *and* $\partial_k c^{(1)}$ *are sufficiently small* (*depending only on* m, σ *and* d) *for all* $k \in \{0, 1, \dots, d\}$*, see Condition* 2.2*,* 1c).

Sketch of proof. Let us start again from the formal computation: using the abbreviations $v_i \coloneqq \partial_i u$ and $\nu_{ij} \coloneqq \partial_j \partial_i u$ (thus $\nu_{ij} = \nu_{ji}$) for $i, j \in \{0, 1, \dots, d\}$ (and with ∂_0 the identity operator), we have

$$\mathcal{L} v_i = -(\partial_i b \cdot \nabla u + \partial_i c u + c v_i), \quad \text{for } i = 1, \dots, d, \quad \text{and} \quad \mathcal{L} v_0 = -cu.$$

Differentiating once more, we find for $i, j \in \{1, \dots, d\}$ the identity

$$\begin{aligned} \mathcal{L} \partial_j v_i &= \partial_j \mathcal{L} v_i - \partial_j b \cdot \nabla v_i \\ &= -\partial_j \partial_i b \cdot \nabla u - \partial_i b \cdot \partial_j \nabla u - \partial_j \partial_i c u - \partial_i c \partial_j u - \partial_j c v_i - c \partial_j v_i - \partial_j b \cdot \nabla v_i. \end{aligned}$$

Hence, setting again $b_0 \coloneqq c$, we end up with the equations

$$\mathcal{L} \nu_{ij} = \begin{cases} -\sum_{k=0}^{d} \partial_j \partial_i b_k \nu_{k0} - \sum_{k=0}^{d} (\partial_i b_k \nu_{kj} + \partial_j b_k \nu_{ik}) - b_0 \nu_{ij} & \text{for } i, j \neq 0, \\ \qquad\qquad - \sum_{k=0}^{d} \partial_i b_k \nu_{k0} - b_0 \nu_{i0} & \text{for } i > j = 0, \\ \qquad\qquad\qquad\qquad - b_0 \nu_{00} & \text{for } i = j = 0. \end{cases}$$

We next would like to pass to products of the ν_{ij}'s. To this end we consider K to be an element in $\cup_{m \in \mathbb{N}} (\{0, 1, \dots, d\} \times \{0, 1, \dots, d\})^m$ and set $|K| = m$ if $K \in (\{0, 1, \dots, d\} \times \{0, 1, \dots, d\})^m$. Moreover, we may assume $i \geq j$ for every $(i, j) \in K$. As before, $K \setminus (i, j)$ means that we drop one component in K of value (i, j), and similarly $K \setminus (i, j) \cup \{k, \ell\}$ now means that we substitute in K one component of value (i, j) by one of value (k, ℓ) if $k \geq \ell$ or by one of the value (ℓ, k) otherwise. Again, which component is dropped does not matter because in the end we will only consider expressions which depend on the total number of the single components, but not on their numbering. We now set $\nu_K \coloneqq \prod_{(i,j) \in K} \nu_{ij}$, and we then infer from the previous equations satisfied by ν_{ij}, via the Leibniz rule and by distinguishing the cases when $j \neq 0$, $i > j = 0$ and $i = j = 0$, that

$$\begin{aligned} \mathcal{L} \nu_K &= \sum_{(i,j) \in K} \nu_{K \setminus (i,j)} \mathcal{L} \nu_{ij} \\ &= -b_0 |K| \nu_K - \sum_{(i,j) \in K, i > 0} \sum_{k=0}^{d} \nu_{K \setminus (i,j) \cup \{k,j\}} \partial_i b_k \\ &\quad - \sum_{(i,j) \in K, i,j > 0} \sum_{k=0}^{d} \Big(\nu_{K \setminus (i,j) \cup \{k,0\}} \partial_j \partial_i b_k + \nu_{K \setminus (i,j) \cup \{i,k\}} \partial_j b_k \Big). \end{aligned}$$

Rewriting the Stratonovich term in $\mathcal{L} \nu_K$ via $\sigma \nabla \nu_K \circ dW = \sigma \nabla \nu_K dW - \frac{\sigma^2}{2} \Delta \nu_K dt$ and by (formally) taking the expectation, we then obtain that $\omega_K \coloneqq E[\nu_K]$ satisfies the equation

$$\begin{aligned} &\partial_t \omega_K + b \cdot \nabla \omega_K + b_0 |K| \omega_K + \sum_{(i,j) \in K, i > 0} \sum_{k=0}^{d} \omega_{K \setminus (i,j) \cup \{k,j\}} \partial_i b_k \\ &\qquad + \sum_{(i,j) \in K, i,j > 0} \sum_{k=0}^{d} \Big(\omega_{K \setminus (i,j) \cup \{k,0\}} \partial_j \partial_i b_k + \omega_{K \setminus (i,j) \cup \{i,k\}} \partial_j b_k \Big) = \frac{\sigma^2}{2} \Delta \omega_K. \end{aligned} \tag{2.23}$$

This system of equations is of the same structure as the system (2.10) derived for the averages of products of first order space derivatives of the solution u, with the only difference that now also second order derivatives of the coefficients appear. Analogously to Section 2.6, one can make the previous computations rigorous for the regular solution of Lemma 2.6 to (sgTE), i.e. the functions $\omega_K(t, x)$ are continuously differentiable in time and satisfy the pointwise equation (2.23).

From here on, we can proceed completely analogously to the proof of Theorem 2.7, since – even though there are more terms involved – the structure of the system is essentially the same (note that χ was only introduced after having derived the parabolic equation, hence no second order derivatives of χ appear in the computations). This finishes the sketch of the proof. □

With the previous lemma we can then deduce the existence of a global regular solution for the stochastic generalized equation. To this end, we introduce in analogy to Definition 2.13 the notion of a solution u to (sgTE) of class $W^{2,m}_{\mathrm{loc}}$, just with the additional $W^{2,m}_{\mathrm{loc}}$-regularity in space.

Theorem 2.22. *Let $m \geq 2$ be an even integer and let $s \in \mathbb{R}$ be given. Assume that b, c satisfy Condition 2.20 and let $u_0 \in W^{2,2m}_{(1+|\cdot|)^{2s+d+1}}(\mathbb{R}^d)$. Then there exists a solution u of equation (sgTE) of class $L^m(W^{2,m}_{\mathrm{loc}})$, which further satisfies $u(t,\cdot) \in W^{2,m}_{(1+|\cdot|)^s}(\mathbb{R}^d)$ for a.e. (t,ω). Moreover, there holds*

$$\operatorname*{ess\,sup}_{t\in[0,T]} E\Big[\|u(t,\cdot)\|^m_{W^{2,m}_{(1+|\cdot|)^s}(\mathbb{R}^d)}\Big] < \infty.$$

Sketch of proof. Since b and c are assumed by Condition 2.20 to belong to the extended LPS class (extended in the sense that the decomposition into LPS-part and regular part is available for $\partial_k b$ and $\partial_k c$ for each $k \in \{0,1,\dots,d\}$), we find approximations $b^{(1)}_\varepsilon, b^{(2)}_\varepsilon, c^{(1)}_\varepsilon$ and $c^{(2)}_\varepsilon$ of class $C^\infty_c([0,T]\times\mathbb{R}^d)$ such that all assumptions concerning boundedness or convergence in Condition 2.14 are satisfied for $\partial_k b^{(1)}_\varepsilon, \partial_k b^{(2)}_\varepsilon, \partial_k c^{(1)}_\varepsilon$ and $\partial_k c^{(2)}_\varepsilon$, for every $k = 0,1,\dots,d$. We then set $b_\varepsilon = b^{(1)}_\varepsilon + b^{(2)}_\varepsilon$, $c_\varepsilon = c^{(1)}_\varepsilon + c^{(2)}_\varepsilon$. We further choose an $C^\infty_c(\mathbb{R}^d)$-approximation $(u^\varepsilon_0)_\varepsilon$ of the initial values u_0 with respect to $W^{2,2m}_{(1+|\cdot|)^{2s+d+1}}(\mathbb{R}^d)$ and denote by u_ε the regular solution to (sgTE) given by Lemma 2.6, corresponding to coefficients b_ε, c_ε and initial values u^ε_0.

We now take $\chi = (1+|x|)^{2s+d+1}$ in the previous lemma and then deduce from Hölder's inequality, as in Corollary 2.8, the bound

$$\sup_{t\in[0,T]} E\Big[\|u_\varepsilon(t,\cdot)\|^m_{W^{2,m}_{(1+|\cdot|)^s}(\mathbb{R}^d)}\Big] \leq C\|u_0\|^m_{W^{2,2m}_{(1+|\cdot|)^{2s+d+1}}(\mathbb{R}^d)},$$

with a constant C which does not depend on the particular approximation, but only on its norms, and therefore this bound holds uniformly in $\varepsilon \in (0,1)$. From this stage we can follow the strategy of the proof of Theorem 2.17. Indeed, the previous inequality yields that the family $(u_\varepsilon)_\varepsilon$ is bounded in $L^\infty([0,T]; L^m(\Omega; W^{2,m}(B_R)))$ for every $R > 0$. Hence, there exists a subsequence weakly-$*$ convergent to a limit process u in this space. This yields the asserted Sobolev-type regularity involving derivatives up to second order, while the fact that u is indeed a solution to (sgTE) with coefficients b, c was already established in the proof of Theorem 2.17. □

Remark 2.23. *In a similar way one can show higher order Sobolev regularity of type $W^{\ell,m}_{\mathrm{loc}}$, provided that b and c are more regular, in the sense that they can be decomposed into $b^{(1)} + b^{(2)}$ and $c^{(1)} + c^{(2)}$ such that each derivative of these decompositions up to order $\ell - 1$ satisfies Condition 2.2. However, it remains an interesting open question to prove a similar result for fractional Sobolev spaces.*

3 Path-by-path uniqueness for sCE and sTE

The aim of this section is to prove a path-by-path uniqueness result for both sTE and sCE. Since we deal with weak solutions, where an integration by parts is necessary at the level of the definition, the general stochastic equation (sgTE) is not the most convenient one. Let us consider a similar equation in divergence form

$$du + (\operatorname{div}(bu) + cu)dt + \sigma \operatorname{div}(u \circ dW_t) = 0, \qquad u|_{t=0} = u_0 \tag{3.1}$$

for vector fields $b\colon [0,T]\times\mathbb{R}^d \to \mathbb{R}^d$ and $c\colon [0,T]\times\mathbb{R}^d \to \mathbb{R}$. We observe that

(i) for regular coefficients, the equations (3.1) and (sgTE) are equivalent (renaming b and c);

(ii) the sCE is included in (3.1), with u as density of the measure μ_t with respect to the Lebesgue measure;

(iii) the sTE is included in (3.1), by formally setting $c = -\operatorname{div} b$ (which then gives rise to a restriction on $\operatorname{div} b$ for this equation).

We recall from the introduction that all path-by-path uniqueness results rely heavily on the regularity results achieved in the previous section. For this reason we will always assume Condition 2.2 of Section 2.1, which allows us to decompose the vector fields b and c into rough parts $b^{(1)}$ and $c^{(1)}$ under a LPS-condition and more regular parts $b^{(2)}$ and $c^{(2)}$ under an integrability condition in time (only here the L^2-integrability in time is required, cp. Remark 2.3). Concerning the LPS-condition, we still denote the exponents by $p, q \geq 2$ and the conjugate exponent of p by p'. We will consider the purely stochastic case $\sigma \neq 0$ throughout this section.

We can now introduce the concept of a weak solution of the stochastic equation (3.1), in analogy to Definition 2.13 (in particular, it is easily verified that all integrals are well-defined by the integrability assumptions on the vector fields b and c and on the weak solution). We recall that $(\mathcal{G}_t)_{t\in[0,T]}$ is a filtration satisfying the standard assumptions and W denotes a Brownian motion with respect to $(\mathcal{G}_t)_t$.

Definition 3.1. *Let $m \geq 2$ be given. A* weak solution of equation (3.1) of class $L^m(L^m_{\mathrm{loc}})$ *is a random field $u\colon \Omega \times [0,T] \times \mathbb{R}^d \to \mathbb{R}$ with the following properties:*

(o) *it is weakly progressively measurable with respect to $(\mathcal{G}_t)_t$;*

(i) *it is in $L^m([0,T] \times B_R \times \Omega)$ for every $R > 0$;*

(ii) *$t \mapsto \langle u(t), \varphi\rangle$ has a modification which is a continuous semimartingale, for every φ in $C_c^\infty(\mathbb{R}^d)$;*

(iii) *for every φ in $C_c^\infty(\mathbb{R}^d)$, for this continuous modification (still denoted by $\langle u(t), \varphi\rangle$), it holds for all $t \in [0,T]$*

$$\langle u(t), \varphi\rangle = \langle u_0, \varphi\rangle + \int_0^t \langle u(s), b(s)\cdot\nabla\varphi - c(s)\varphi\rangle\, ds + \sigma \int_0^t \langle u(s), \nabla\varphi\rangle \circ dW_s.$$

Remark 3.2. *The previous definition can be given with different degrees of integrability in time and space, namely for solutions of class $L^\theta(L^m_{\mathrm{loc}})$ with $\theta \geq 2$ and $m \geq p'$ (cp. Definition 2.13). We take $\theta = m$ only to minimize the notations.*

Since our aim is to establish the stronger results of path-by-path uniqueness, we first give a path-by-path formulation of (3.1). Let us recall that we started with a probability space $(\Omega, \mathcal{A}, P)$, a filtration $(\mathcal{G}_t)_{t\geq 0}$ (satisfying the standard assumptions), and a Brownian motion $(W_t)_{t\geq 0}$. We now choose, without restriction, a version of W_t which is continuous for every $\omega \in \Omega$. Given $\omega \in \Omega$, considered here as a parameter, we define

$$\widetilde{b}(\omega, t, x) \coloneqq b(t, x + \sigma W_t(\omega))$$
$$\widetilde{c}(\omega, t, x) \coloneqq c(t, x + \sigma W_t(\omega)).$$

We shall write sometimes $\widetilde{b}^\omega$ and $\widetilde{c}^\omega$ for $\widetilde{b}(\omega, \cdot, \cdot)$ and $\widetilde{c}(\omega, \cdot, \cdot)$, respectively, in order to stress the parameter character of ω. With this new notation we now consider the following deterministic PDE, parametrized by $\omega \in \Omega$, in the unknown $\widetilde{u}^\omega\colon [0,T] \times \mathbb{R}^d \to \mathbb{R}$:

$$\partial_t \widetilde{u}^\omega + \operatorname{div}(\widetilde{b}^\omega \widetilde{u}^\omega) + \widetilde{c}^\omega \widetilde{u}^\omega = 0, \qquad \widetilde{u}^\omega|_{t=0} = u_0. \tag{3.2}$$

Definition 3.3. *Let $m \geq 2$. Given $\omega \in \Omega$, we say that $\widetilde{u}^\omega\colon [0,T] \times \mathbb{R}^d \to \mathbb{R}$ is a* weak solution to equation (3.2) of class $L^m(L^m_{\mathrm{loc}})$ *if*

(i) *$\widetilde{u}^\omega \in L^m([0,T] \times B_R)$, for every $R > 0$,*

(ii) *for each φ in $C^1([0,T]; C_c^\infty(\mathbb{R}^d))$, $t \mapsto \langle \widetilde{u}^\omega(t), \varphi(t)\rangle$ is continuous; precisely, this map has a continuous representative, where by representative we mean a function which coincides with $t \mapsto \langle \widetilde{u}^\omega(t), \varphi(t)\rangle$ for $\mathcal{L}^1$-a.e. $t \in [0,T]$,*

(iii) *for all $\varphi \in C^1([0,T]; C_c^\infty(\mathbb{R}^d))$, we have for this continuous representative, for all t in $[0,T]$,*

$$\langle \widetilde{u}^\omega(t), \varphi(t)\rangle = \langle u_0, \varphi(0)\rangle + \int_0^t \Big\langle \widetilde{u}^\omega(s), \partial_t\varphi(s) + \widetilde{b}^\omega(s)\cdot\nabla\varphi(s) - \widetilde{c}^\omega(s)\varphi(s)\Big\rangle\, ds. \tag{3.3}$$

Notice that we have employed here time-dependent test functions. This is only for a technical convenience (we will use such functions in the following), and the definition with autonomous test functions could be shown to be equivalent to Definition 3.3. Equation (3.2) will be considered as the path-by-path formulation of (3.1). The reason is:

Proposition 3.4. *If u is a weak solution of the stochastic equation* (3.1) *of class $L^m(L^m_{\mathrm{loc}})$ in the sense of Definition* 3.1*, then $\widetilde{u}^\omega$ defined as*

$$\widetilde{u}^\omega(t,x) \coloneqq u(t, x + \sigma W_t(\omega))$$

is, for a.e. $\omega \in \Omega$, a weak solution of the deterministic equation (3.2) *of class $L^m(L^m_{\mathrm{loc}})$ in the sense of Definition* 3.3*.*

Remark 3.5. *The following proof, simple in the idea, becomes tedious because of a technical detail which we will encounter also in the following: equation* (3.1) *resp.* (3.2)*, in its weak formulation, is satisfied by $\langle u, \varphi\rangle$ resp. $\langle \tilde{u}, \varphi\rangle$ only for a.e. $t \in [0,T]$, and the exceptional set in $[0,T]$, where this formulation does not hold, could depend on φ, ω and the initial datum. This problem can be overcome essentially in every case, but with some small work* (*see also Lemma* 3.10)*.*

Proof. The idea of the proof is given by the following formal computation, using the Itô formula (in Stratonovich form):

$$\begin{aligned}\partial_t \tilde{u}(t,x) &= \partial_t u(t, x+\sigma W_t) + \sigma \nabla u(t, x+\sigma W_t) \circ \dot{W}_t \\ &= -\operatorname{div}\big(\tilde{b}(t,x)\tilde{u}(t,x)\big) - \tilde{c}(t,x)\tilde{u}(t,x). \end{aligned} \tag{3.4}$$

Since this does not work rigorously when u is not regular, one could try to apply the change of variable formula on the test function rather than on u itself, i.e. taking $\tilde{\varphi}(t,x) = \varphi(t, x - \sigma W_t)$ (which is smooth) as test function in equation (3.1) for u and then use a change of variable to get equation (3.2) for $\tilde{u}$, with φ as test function. The problem is that $\tilde{\varphi}$, besides being time-dependent, is not deterministic (but Definition 3.1 only allows deterministic test functions). Thus, we proceed by approximation. The idea is the following: taking a family $(\rho_\varepsilon)_\varepsilon$ of standard symmetric, compactly supported mollifiers, we first use a shifted version of ρ_ε as test function, to get an equation for the mollification $u^\varepsilon \coloneqq u * \rho_\varepsilon$ for fixed x; having regularity of u^ε, we can derive a formula for $u^\varepsilon(t,x)\varphi(t, x - \sigma W_t)$. After integrating in x, taking the limit $\varepsilon \to 0$ and a change of variable, we finally get an equation for $\tilde{u}$, still in a weak formulation.

For simplicity of notation, we set $c = 0$ and $\sigma = 1$, but all the arguments are valid with immediate extension also in the general case.

Step 1: For fixed $\varphi \in C^1([0,T]; C^\infty_c(\mathbb{R}^d))$, the mollifications u^ε satisfy, for a.e. (t,x,ω),

$$\begin{aligned} &u^\varepsilon(t,x)\varphi(t, x - W_t) \\ &= u^\varepsilon_0(x)\varphi(0,x) - \int_0^t (u(s)b(s)) * \nabla\rho_\varepsilon(x)\varphi(s, x - W_s)ds \\ &\quad - \int_0^t u(s) * \nabla\rho_\varepsilon(x)\varphi(s, x - W_s)\cdot dW_s + \frac{1}{2}\int_0^t u(s) * \Delta\rho_\varepsilon(x)\varphi(s, x - W_s)ds \\ &\quad - \int_0^t u^\varepsilon(s,x)\nabla\varphi(s, x - W_s)\cdot dW_s + \int_0^t u^\varepsilon(s,x)\big(\partial_t + \frac{1}{2}\Delta\big)\varphi(s, x - W_s)ds \\ &\quad + \int_0^t u(s) * \nabla\rho_\varepsilon(x)\cdot\nabla\varphi(s, x - W_s)ds, \end{aligned} \tag{3.5}$$

and all the addends have modification that are measurable in (t,x,ω) (these are the modifications considered in the equality above). We fix a measurable map u (not equivalence class), so that by Fubini's theorem convolutions of u are measurable maps in (t,x,ω). For fixed $x \in \mathbb{R}^d$, we apply Definition 3.1 of a weak solution with test function $\varphi = \rho_\varepsilon(x - \cdot) \in C^\infty_0(\mathbb{R}^d)$, getting the following equation for a modification $u(\rho_\varepsilon(x-\cdot))$ of $u^\varepsilon(x) = u * \rho_\varepsilon(x) = \langle u, \rho_\varepsilon(x-\cdot)\rangle$ (here the notation $\nabla_\cdot\rho_\varepsilon(x-\cdot)$ means the derivative with respect to the $\cdot$ variable, with x fixed):

$$u(\rho_\varepsilon(x-\cdot))(t) \tag{3.6}$$

$$= \langle u_0, \rho_\varepsilon(x-\cdot)\rangle + \int_0^t \langle u(s), b(s)\cdot\nabla_\cdot \rho_\varepsilon(x-\cdot)\rangle ds + \int_0^t \langle u(s), \nabla_\cdot \rho_\varepsilon(x-\cdot)\rangle \circ dW_s$$
$$= u_0^\varepsilon(x) - \int_0^t (u(s)b(s)) * \nabla\rho_\varepsilon(x) ds - \int_0^t u(s) * \nabla\rho_\varepsilon(x)\cdot dW_s + \frac{1}{2}\int_0^t u(s) * \Delta\rho_\varepsilon(x) ds,$$

where we also have passed from Stratonovich to Itô stochastic integral. Applying Itô's product formula to $u(\rho_\varepsilon(x-\cdot))$ and $\varphi(t, x-W_t)$, we find that P-a.s. it hold for every $t\in[0,T]$

$$\begin{aligned} &u(\rho_\varepsilon(x-\cdot))(t)\varphi(t,x-W_t) \qquad (3.7)\\ &= u_0^\varepsilon(x)\varphi(0,x) - \int_0^t (u(s)b(s)) * \nabla\rho_\varepsilon(x)\varphi(s,x-W_s) ds\\ &\quad - \int_0^t u(s) * \nabla\rho_\varepsilon(x)\varphi(s,x-W_s)\cdot dW_s + \frac{1}{2}\int_0^t u(s) * \Delta\rho_\varepsilon(x)\varphi(s,x-W_s) ds\\ &\quad - \int_0^t u(\rho_\varepsilon(x-\cdot))(s)\nabla\varphi(s,x-W_s)\cdot dW_s + \int_0^t u(\rho_\varepsilon(x-\cdot))(s)\big(\partial_t + \frac{1}{2}\Delta\big)\varphi(s,x-W_s) ds\\ &\quad + \int_0^t u(s) * \nabla\rho_\varepsilon(x)\cdot\nabla\varphi(s,x-W_s) ds. \end{aligned}$$

Since, for fixed x, $u(\rho^\varepsilon(x-\cdot)) = u^\varepsilon(x)$ for a.e. (t,ω), we can replace $u(\rho^\varepsilon(x-\cdot))$ with $u^\varepsilon(x)$ inside the integrals, which implies (3.5) for all (t,ω) in a full-measure set A_x, possibly depending on x. Note that, up to this point, we have not used any measurability in x.

Now let us justify that all the addends in (3.5) have modifications which are measurable in (t,x,ω). By classical Fubini theorem, the mollifications of u and thus all the addends but the stochastic integrals are measurable in (t,x,ω). Concerning the stochastic integrals, their integrands are, in view of the weak progressive measurability of u, measurable in (t,x,ω) with respect to $\mathcal{P}\otimes\mathcal{B}(\mathbb{R}^d)$, where $\mathcal{P}$ is the progressive σ-algebra. Thus, stochastic Fubini theorem (see e.g. [77, Theorem 2.2] applies and gives the existence of measurable modifications in (t,x,ω). For such modifications, (3.5) must holds for a.e. (t,x,ω): if this were not the case, then there would exist a positive measure set B in $\mathbb{R}^d$, such that, for every $x\in B$, there would exist a positive measure set C_x in $[0,T]\times\Omega$ where equality 3.5 would not hold. Since the addends of (3.5) are modifications of the addends of those of (3.7), also (3.7) would not hold on this set, which is a contradiction, cf. Remark 4.8.

Step 2: For fixed $\varphi\in C^1([0,T];C_c^\infty(\mathbb{R}^d))$, $\tilde{u}$ has the solution property (3.3) *a.s..* We may now integrate, for a.e. (t,ω), the identity (3.5) with respect to x, obtaining

$$\begin{aligned} &\int_{\mathbb{R}^d} u^\varepsilon(t,x)\varphi(t,x-W_t)dx \qquad (3.8)\\ &= \int_{\mathbb{R}^d} u_0^\varepsilon(x)\varphi(0,x)dx - \int_0^t\int_{\mathbb{R}^d} (u(s)b(s)) * \nabla\rho_\varepsilon(x)\varphi(s,x-W_s)dxds\\ &\quad - \int_0^t\int_{\mathbb{R}^d} u(s) * \nabla\rho_\varepsilon(x)\varphi(s,x-W_s)dx\cdot dW_s + \frac{1}{2}\int_0^t\int_{\mathbb{R}^d} u(s) * \Delta\rho_\varepsilon(x)\varphi(s,x-W_s)dxds\\ &\quad - \int_0^t\int_{\mathbb{R}^d} u^\varepsilon(s,x)\nabla\varphi(s,x-W_s)dx\cdot dW_s + \int_0^t\int_{\mathbb{R}^d} u^\varepsilon(s,x)(\partial_t+\frac{1}{2}\Delta)\varphi(s,x-W_s)dxds\\ &\quad + \int_0^t\int_{\mathbb{R}^d} u(s) * \nabla\rho_\varepsilon(x)\cdot\nabla\varphi(s,x-W_s)dxds, \end{aligned}$$

where we have also used the Fubini as well as the stochastic Fubini to exchange the order of integration. Employing once again Fubini theorem to bring the convolution on $\varphi(t,\cdot-W_t)$, we get, for a.e. (t,ω),

$$\begin{aligned} &\langle u(t), \varphi^\varepsilon(t,\cdot-W_t)\rangle\\ &= \langle u_0, \varphi^\varepsilon(0)\rangle + \int_0^t \langle u(s), b(s)\cdot\nabla\varphi^\varepsilon(s,\cdot-W_s)\rangle ds\\ &\quad + \int_0^t \langle u(s), \nabla\varphi^\varepsilon(s,\cdot-W_s)\rangle\cdot dW_s + \frac{1}{2}\int_0^t \langle u(s), \Delta\varphi^\varepsilon(s,\cdot-W_s)\rangle ds \end{aligned}$$

$$
\begin{aligned}
&- \int_0^t \langle u(s), \nabla\varphi^\varepsilon(s, \cdot - W_s)\rangle \cdot dW_s + \int_0^t \langle u(s), (\partial_t + \frac{1}{2}\Delta)\varphi^\varepsilon(s, \cdot - W_s)\rangle ds \\
&- \int_0^t \langle u(s), \Delta\varphi^\varepsilon(s, \cdot - W_s)\rangle ds \\
&= \langle u_0, \varphi^\varepsilon(0)\rangle + \int_0^t \langle u(s), (\partial_t + b(s)\cdot\nabla)\varphi^\varepsilon(s, \cdot - W_s)\rangle ds.
\end{aligned}
$$

Letting $\varepsilon \to 0$, since u, bu are in $L^1([0,T]; L^1_{\mathrm{loc}}(\mathbb{R}^d))$ for a.e. ω, we have for a.e. (t,ω),

$$
\langle u(t), \varphi(t, \cdot - W_t)\rangle = \langle u_0, \varphi(0)\rangle + \int_0^t \langle u(s), (\partial_t + b(s)\cdot\nabla)\varphi(s, \cdot - W_s)\rangle ds.
$$

By the change of variable $\tilde{x} = x - W_t$, we therefore end up with the claimed solution property

$$
\langle \tilde{u}(t), \varphi(t)\rangle = \langle u_0, \varphi(0)\rangle + \int_0^t \langle \tilde{u}(s), (\partial_t + \tilde{b}(s)\cdot\nabla)\varphi(s)\rangle ds \tag{3.9}
$$

for fixed test function $\varphi \in C^1([0,T]; C^\infty_c(\mathbb{R}^d))$, for every (t,ω) in a full measure set F_φ, which may still depend on φ.

Step 3: Removal of the dependency on the test function φ. In order to conclude the proof of the proposition, we need to make the "good" full measure set, where $\tilde{u}$ satisfies the solution property, independent of φ. For this purpose, we use a density argument. Let D be a countable set in $C^1([0,T]; C^\infty_c(\mathbb{R}^d))$, which is dense in $C^1([0,T]; C^2_b(\mathbb{R}^d))$, and set $F = \cap_{\varphi\in D} F_\varphi$. Then F is a full measure set and (3.9) holds for every $(t,\omega) \in F$ and $\varphi \in D$; after possibly passing to a smaller full-measure set F we can also assume $\tilde{u}^\omega \in L^m([0,T]; L^m_{\mathrm{loc}}(\mathbb{R}^d))$ (thus, fulfilling Definition 3.3 (i)). Now, for a generic test function $\varphi \in C^1([0,T]; C^\infty_c(\mathbb{R}^d))$, we take a sequence $(\varphi^n)_{n\in\mathbb{N}}$ in D, satisfying equation (3.9) and converging to φ in $C^1([0,T]; C^2_b(\mathbb{R}^d))$; by dominated convergence theorem, we can pass to the limit in the equation, for $(t,\omega) \in F$, getting (3.9) for φ. Hence, for a.e. (t,ω), (3.3) holds and the right-hand side defines the continuous representative. □

Since some technical measurability arguments are delicate in the above proof (based mostly on Fubini and stochastic Fubini theorems), we want to give alternative proofs of Step 1 and formula (3.8) at the beginning of Step 2, which rely on a direct exchange of integral formula obtained by continuity of approximations.

Alternative proof of Step 1 and (3.8). *Step 0: Exchange of integrals formula by approximation.* Let $f\colon [0,T]\times\mathbb{R}^d\times\Omega \to \mathbb{R}$ be a function such that:

- f is measurable in (t,x,ω),
- for every x, $(t,\omega) \mapsto f(t,x,\omega)$ is progressively measurable,
- $f \in L^2([0,T]\times\Omega; C^\alpha_{\mathrm{loc}}(\mathbb{R}^d))$ for some $\alpha > 0$.

Then the family of stochastic integrals $\int_0^t f(r,x)dW_r$, parametrized by x, admits a modification which is measurable in (t,x,ω), for every x progressively measurable in (t,ω), and for a.e. ω locally Hölder continuous in (t,x). This can be proven by Kolmogorov's continuity criterion in (t,x) for the stochastic integrals (joint measurability is a consequence of progressive measurability and continuity in (t,x)). Moreover, for such modification, we have for a.e. $\omega\in\Omega$: for every $t\in[0,T]$,

$$
\int_{\mathbb{R}^d}\int_0^t f(r,x)dW_r dx = \int_0^t \int_{\mathbb{R}^d} f(r,x)dx dW_r,
$$

provided the integrals are well-defined (for example, if f is compactly supported). This is a consequence of stochastic Fubini theorem but can be proved without it:

For this purpose, we first observe that by continuity of the stochastic integrals in (t,x), we can approximate, for fixed t, for a.e. $\omega\in\Omega$, the left-hand side $\int_{\mathbb{R}^d}\int_0^t f(r,x)dW_r dx$ with a finite Riemann sum (in x) of stochastic integrals. We then notice that we can approximate the inner integral $\int_{\mathbb{R}^d} f(r,x)dx$ in $L^2([0,T]\times\Omega)$ with a finite Riemann sum (in x), and as a consequence, we

can approximate, for fixed t, the right-hand side $\int_0^t \int_{\mathbb{R}^d} f(r,x)dxdW_r$ in $L^2(\Omega)$ with the stochastic integral of a finite Riemann sum (in x). At the level of these approximations sums we can finally exchanging sum and stochastic integral, and passing to the limit we get the equality above.

Alternative proof of Step 1 above. As before, we fix a measurable map u (not equivalence class), so that, by Fubini theorem, all the convolutions with u are measurable maps in (t,x,ω), regular in x for a.e. (t,ω) fixed. Then, for fixed $x \in \mathbb{R}^d$, our starting point is the modification $u(\rho_\varepsilon(x-\cdot))$ of $u^\varepsilon(x) = u * \rho_\varepsilon(x) = \langle u, \rho_\varepsilon(x-\cdot)\rangle$ satisfying (3.6) and (3.7). Replacing $u(\rho^\varepsilon(x-\cdot))$ with $u^\varepsilon(x)$ inside the integrals of (3.7) (as before), we get for a.e. ω, for every t,

$$\begin{aligned}
&u(\rho_\varepsilon(x-\cdot))(t)\varphi(t,x-W_t)\\
&= u_0^\varepsilon(x)\varphi(0,x) - \int_0^t (u(s)b(s)) * \nabla\rho_\varepsilon(x)\varphi(s,x-W_s)ds\\
&\quad - \int_0^t u(s) * \nabla\rho_\varepsilon(x)\varphi(s,x-W_s)\cdot dW_s + \frac{1}{2}\int_0^t u(s) * \Delta\rho_\varepsilon(x)\varphi(s,x-W_s)ds\\
&\quad - \int_0^t u^\varepsilon(s,x)\nabla\varphi(s,x-W_s)\cdot dW_s + \int_0^t u^\varepsilon(s,x)\big(\partial_t + \frac{1}{2}\Delta\big)\varphi(s,x-W_s)ds\\
&\quad + \int_0^t u(s) * \nabla\rho_\varepsilon(x)\cdot\nabla\varphi(s,x-W_s)ds.
\end{aligned}$$

For the stochastic integrals, the integrands $u(s) * \nabla\rho_\varepsilon(x)\varphi(s,x-W_s)$ and $u^\varepsilon(s,x)\nabla\varphi(s,x-W_s)$ are measurable in (t,x,ω), progressively measurable for every fixed x, and they also belong to $L^2([0,T]\times\Omega; C^1_{\mathrm{loc}}(\mathbb{R}^d))$. Therefore, by Step 0, there exist "nice" modifications of the stochastic integrals. Using these modifications, we get for every x, for a.e. (t,ω) (where the exceptional set possibly depends on x) precisely the formula (3.5). Moreover, since all the addends are measurable in (t,x,ω) by construction, this equality is true for a.e. (t,x,ω) (otherwise we would find positive measure sets A_x in $[0,T]\times\Omega$, for some x, where the equality above would not hold).

Alternative justification of (3.8). As before we again integrate (3.5) in x, for a.e. (t,ω), but at this stage we may then use Fubini theorem to exchange the integrals in ds and dx, while we may use Step 0 to exchange the integral in dW_s and dx. □

Remark 3.6. *One can ask why such a change of variable works and if this is simply a trick. Actually this is not the case: as we will see in Section* 4*, this change of variable corresponds to looking at the random ODE*

$$d\tilde{X}^\omega = \tilde{b}^\omega(t,\tilde{X}^\omega)dt.$$

A similar change of variable can be done also for more general diffusion coefficients, see the discussion in the Introduction, Paragraph 1.9*.*

3.1 The duality approach in the deterministic case

To prove uniqueness for equation (3.2), we shall follow a duality approach. It is convenient to recall the idea in a deterministic case first, especially in view of condition (3.14) further below. For the sake of illustration, we give here an Hilbert space description, even though the duality approach will be developed later in the stochastic case in a more general set-up.

Assume we have a Hilbert space H with inner product $\langle\cdot,\cdot\rangle_H$ and two Hilbert spaces D_A, D_{A^*} which are continuously embedded in H, $D_A \subset H$ and $D_{A^*} \subset H$. Furthermore, let $A(t)\colon D_A \to H$ and $A(t)^*\colon D_{A^*} \to H$ be two families of bounded linear operators such that

$$\langle A(t)x, y\rangle_H = \langle x, A(t)^* y\rangle_H$$

for all $x \in D_A$, $y \in D_{A^*}$. Consider the linear evolution equation in H

$$\partial_t u(t) - A(t)u(t) = 0 \quad \text{for } t \in [0,T], \qquad u|_{t=0} = u_0 \tag{3.10}$$

and suppose that we want to study uniqueness of weak solutions, defined as those functions $u\colon [0,T] \to H$, bounded and weakly continuous, such that

$$\langle u(t), \varphi\rangle_H = \langle u_0, \varphi\rangle_H + \int_0^t \langle u(s), A(s)^*\varphi\rangle_H \, ds$$

for all $\varphi \in D_{A^*}$ and all $t \in [0,T]$. Assume we can prove that this weak formulation implies

$$\langle u(t), \varphi(t) \rangle_H = \langle u_0, \varphi(0) \rangle_H + \int_0^t \Big\langle u(s), A(s)^* \varphi(s) + \partial_t \varphi(s) \Big\rangle_H ds \tag{3.11}$$

for all $\varphi \in C([0,T]; D_{A^*}) \cap C^1([0,T]; H)$ and all $t \in [0,T]$. In order to identify u at any time $t_f \in [0,T]$, we need to consider the dual problem on $[0,t_f]$ with final condition at time t_f. Thus, given any $t_f \in [0,T]$, we consider the equation

$$\partial_t v(t) + A(t)^* v(t) = 0 \quad \text{for } t \in [0,t_f], \qquad v|_{t=t_f} = v_0 \tag{3.12}$$

and assume that, for every v_0 in a dense set $\mathcal{D}$ of H, it has a regular solution $v \in C([0,t_f]; D_{A^*}) \cap C^1([0,t_f]; H)$. Then by the previous assumption (3.11) we obtain with the choice $\varphi = v$ that

$$\langle u(t_f), v_0 \rangle_H = \langle u_0, v(0) \rangle_H .$$

If $u_0 = 0$, then $\langle u(t_f), v_0 \rangle_H = 0$ for every $v_0 \in \mathcal{D}$, hence $u(t_f) = 0$. This implies uniqueness for equation (3.10) by linearity.

Let us repeat this scheme (still considering the case $u_0 = 0$), when a regularized version of the dual equation is used. Assume we have a sequence of (smooth) approximations of equation (3.12)

$$\partial_t v_\varepsilon(t) + A_\varepsilon(t)^* v_\varepsilon(t) = 0 \quad \text{for } t \in [0,t_f], \qquad v_\varepsilon|_{t=t_f} = v_0,$$

where $A_\varepsilon(t)^* \colon D_{A^*} \to H$. If, for every $v_0 \in \mathcal{D}$, we have regular solutions $v_\varepsilon \in C([0,t_f]; D_{A^*}) \cap C^1([0,t_f]; H)$, then, if $u_0 = 0$, we find

$$\langle u(t_f), v_0 \rangle_H = \int_0^{t_f} \langle u(s), (A^*(s) - {A_\varepsilon}^*(s)) v_\varepsilon(s) \rangle_H \, ds \tag{3.13}$$

again from (3.11), and hence

$$| \langle u(t_f), v_0 \rangle_H | \le \|u\|_{L^\infty(0,T;H)} \int_0^{t_f} \|(A^*(s) - A_\varepsilon^*(s)) v_\varepsilon(s)\|_H ds.$$

If, for every $t_f \in [0,T]$ and $v_0 \in \mathcal{D}$, we have

$$\lim_{\varepsilon \to 0} \int_0^{t_f} \|(A^*(s) - A_\varepsilon^*(s)) v_\varepsilon(s)\|_H ds = 0, \tag{3.14}$$

then we again conclude with $u = 0$, which proves uniqueness for equation (3.10). A property of the form (3.14) will be a basic tool in the sequel.

The problem to apply this method rigorously is the regularity of v (or a uniform control of the regularity of v_ε). For deterministic transport and continuity equations with rough drift, one cannot solve the dual equation (3.12) in a sufficiently regular space. Thus the regularity results of Section 2 are the key point of this approach, specific to the stochastic case.

3.2 Dual sPDE and random PDE

Let us recall that we started with a probability space $(\Omega, \mathcal{A}, P)$, a (complete and right-continuous) filtration $(\mathcal{G}_t)_{t \ge 0}$ and a Brownian motion $(W_t)_{t \ge 0}$. Given $t_f \in [0,T]$ (which will be the final time), we consider the process

$$B_t \coloneqq W_t - W_{t_f}, \qquad t \in [0,t_f] \tag{3.15}$$

and the family of σ-fields, for $t \in [0,t_f]$,

$$\mathcal{F}^t = \sigma(\{B_s, s \in [t,t_f]\} \cup \mathcal{N}), \tag{3.16}$$

where $\mathcal{N}$ is the set of P-null sets in $\mathcal{A}$. The family $(\mathcal{F}^t)_{t \in [0,t_f]}$ is a backward filtration, in the sense that $\mathcal{F}^{t_1} \subset \mathcal{F}^{t_2}$ if $t_1 > t_2$. The process B is a "backward Brownian motion", or a "Brownian motion in the reversed direction of time", with respect to the filtration $(\mathcal{F}^t)_t$:

- $B_{t_f} = 0$ a.s., $t \mapsto B_t$ is a.s. continuous (in fact, for all $\omega \in \Omega$ by our choice of W_t),
- $B_{t-h} - B_t$ is $\mathcal{N}(0,h)$ and independent of $\mathcal{F}^t$, for every $t \in [0,t_f]$ and $h \in (0,t]$,
- $(\mathcal{F}^t)_{t\in[0,t_f]}$ is complete and right-continuous (see e.g. [8, Proposition 2.5]).

Stochastic calculus in the backward direction of time can be developed without any difference (except notational) compared to the common forward stochastic calculus, see [56, Chapter 3]. Thus, we may consider the backward version of the sPDE (sgTE) in Stratonovich form

$$dv + (b\cdot\nabla v - cv)dt + \sigma\nabla v \circ dB = 0 \quad \text{for } t\in[0,t_f], \qquad v|_{t=t_f} = v_0,$$

and define weak or $W^{1,r}$ solutions in the same way as in the forward case. In fact, instead of this equation, we shall use its regularized version

$$dv_\varepsilon + (b_\varepsilon\cdot\nabla v_\varepsilon - c_\varepsilon v_\varepsilon)dt + \sigma\nabla v_\varepsilon \circ dB = 0 \quad \text{for } t\in[0,t_f], \qquad v_\varepsilon|_{t=t_f} = v_0, \tag{3.17}$$

where b_ε, c_ε, for $\varepsilon > 0$, and v_0 are C_c^∞ functions. First, for every $\varepsilon > 0$, this equations has a smooth solution with the properties described in Lemma 2.6. Second, we have the analog of Theorem 2.7 and Corollary 2.8:

Corollary 3.7. *Let m be an even integer and let s be in $\mathbb{R}^d$. Assume that b, c satisfy Condition 2.2, and let b_ε, c_ε be $C_c^\infty([0,T]\times\mathbb{R}^d)$ functions satisfying Condition 2.14. Then there exists a constant C independent of ε such that, for every $v_0 \in C_c^\infty(\mathbb{R}^d)$, the smooth solution v_ε of equation (3.17) verifies for all $\varepsilon \in (0,1)$*

$$\sup_{t\in[0,T]} E\Big[\|v_\varepsilon(t,\cdot)\|^m_{W^{1,m}_{(1+|\cdot|)^s}(\mathbb{R}^d)}\Big] \le C\|v_0\|^m_{W^{1,2m}_{(1+|\cdot|)^{2s+d+1}}(\mathbb{R}^d)}.$$

Moreover, the analog of Proposition 3.4 holds. But, about this, let us pay attention to the notations. The result here is:

Corollary 3.8. *With the notations $\widetilde{v}^B_\varepsilon(t,x) \coloneqq v_\varepsilon(t,x+\sigma B_t)$, $\widetilde{b}^B_\varepsilon(t,x) \coloneqq b_\varepsilon(t,x+\sigma B_t)$, $\widetilde{c}^B_\varepsilon(t,x) \coloneqq c_\varepsilon(t,x+\sigma B_t)$ we have for a.e. $\omega\in\Omega$ that $\widetilde{v}^B_\varepsilon$ has paths in $C^1([0,t_f];C_c^\infty(\mathbb{R}^d))$ and that there holds*

$$\partial_t\widetilde{v}^B_\varepsilon + \widetilde{b}^B_\varepsilon\cdot\nabla\widetilde{v}^B_\varepsilon - \widetilde{c}^B_\varepsilon\widetilde{v}^B_\varepsilon = 0, \qquad \widetilde{v}^B_\varepsilon|_{t=t_f} = v_0. \tag{3.18}$$

To check that the substitution $x+\sigma B_t$ is correct, we should repeat step by step the proof of Proposition 3.4 in the backward case, but, since this is lengthy, let us only convince ourselves with a formal computation, similar to (3.4), which would be rigorous if W (hence B) and v were smooth:

$$\Big(\partial_t\widetilde{v}^B_\varepsilon + \widetilde{b}^B_\varepsilon\cdot\nabla\widetilde{v}^B_\varepsilon - \widetilde{c}^B_\varepsilon\widetilde{v}^B_\varepsilon\Big)(t,x)$$
$$= \Big(\partial_t v_\varepsilon + b_\varepsilon\cdot\nabla v_\varepsilon - c_\varepsilon v_\varepsilon\Big)(t,x+\sigma B_t) + \sigma\nabla v_\varepsilon(t,x+\sigma B_t)\circ\frac{dB}{dt} = 0.$$

Unfortunately, equation (3.18) is not dual to equation (3.2) (up to the fact that coefficients are smoothed) because, in the coefficients, x is translated by W in (3.2) and by B in (3.18). If we introduce $\widetilde{v}_\varepsilon(t,x) \coloneqq v_\varepsilon(t,x+\sigma W_t)$ we have $\widetilde{v}_\varepsilon(t,x) = \widetilde{v}^B_\varepsilon(t,x+\sigma W_{t_f})$ and therefore:

Corollary 3.9. *With the notations $\widetilde{v}_\varepsilon(t,x) \coloneqq v_\varepsilon(t,x+\sigma W_t)$, $\widetilde{b}_\varepsilon(t,x) \coloneqq b_\varepsilon(t,x+\sigma W_t)$, $\widetilde{c}_\varepsilon(t,x) \coloneqq c_\varepsilon(t,x+\sigma W_t)$, we have for a.e. $\omega\in\Omega$*

$$\partial_t\widetilde{v}_\varepsilon + \widetilde{b}_\varepsilon\cdot\nabla\widetilde{v}_\varepsilon - \widetilde{c}_\varepsilon\widetilde{v}_\varepsilon = 0, \qquad \widetilde{v}_\varepsilon(t_f,x) = v_0(x+\sigma W_{t_f}). \tag{3.19}$$

So $\widetilde{v}_\varepsilon$ solves the dual equation to (3.2) (more precisely, the regularized version of the dual equation), but with a randomized final condition at time t_f. Having in mind the scheme of the previous section, we have found the operator $A^*_\varepsilon(t)$.

3.3 Duality formula

The aim of this section is to prove the duality formula (3.20), in order to repeat the ideas described in Section 3.1. Notice that, by the explicit formula (2.5), smooth solutions of equation (sgTE) with smooth, compactly supported initial data, and therefore also the smooth solution $v_\varepsilon(\omega,t,x)$ of the backward stochastic equation (3.17) with smooth, compactly supported final data, are compactly supported in space, with support depending on (t,ω). The same is true for the function $\widetilde{v}_\varepsilon(\omega,t,x) \coloneqq v_\varepsilon(\omega,t,x+\sigma W_t(\omega))$ (since we have assumed that W_t is continuous for every $\omega\in\Omega$). We shall write v_ε^ω and $\widetilde{v}_\varepsilon^\omega$ for these functions, respectively, for a given $\omega\in\Omega$.

Before going on, we need to give a meaning to equation (3.3) for every t, for a certain fixed (i.e. independent of φ) modification of $\tilde{u}^\omega$ (see Remark 3.5). To this end, we establish the next lemma, in which we denote by $\mathcal{B}_b$ the set of bounded Borel functions and $H^{-1}(B_R) \coloneqq (W_0^{1,2}(B_R))^*$.

Lemma 3.10. *Suppose that $\tilde{u}^\omega$ is a weak solution to equation (3.2) of class $L^m(L^m_{\mathrm{loc}})$ according to Definition 3.3, for some fixed $\omega\in\Omega$. Then there exists a representative of $\tilde{u}^\omega$ (that is, a measurable map $[0,T]\to\mathcal{D}'(\mathbb{R}^d)$ which coincides with $\tilde{u}^\omega$ up to negligible sets in $[0,T]$), still denoted by $\tilde{u}^\omega$, which belongs to $\cap_{R>0}\mathcal{B}_b([0,T];H^{-1}(B_R))$ and satisfies formula (3.3) for every $t\in[0,T]$.*

Unless otherwise stated, we will use this representative $\tilde{u}^\omega$ (and the validity of formula (3.3) for every $t\in[0,T]$).

Proof. We will omit the superscript ω in the following. In order to construct the representative, we fix $t\in[0,T]$ and define $F_t\colon C_c^\infty(\mathbb{R}^d)\to\mathbb{R}$ via

$$\langle F_t,\varphi\rangle = \langle u_0,\varphi\rangle + \int_0^t \langle \tilde{u}(s), \tilde{b}(s)\cdot\nabla\varphi - \tilde{c}(s)\varphi\rangle ds.$$

By our integrability assumption on b and c, if φ has support in B_R, then $|\langle F_t,\varphi\rangle|$ is bounded by $C_R\|\varphi\|_{W^{1,2}}$, with a constant C_R which is independent of t. So, for any $R>0$, F_t can be extended to a linear continuous functional on $W_0^{1,2}(B_R)$, with norms uniformly bounded in t.

Let us verify that F is a representative of $\tilde{u}$. By equation (3.3), for every time-independent test function φ in $C_c^\infty(\mathbb{R}^d)$, there exists a full $\mathcal{L}^1$-measure set A_φ in $[0,T]$ such that, for all t in A_φ, $\langle F_t,\varphi\rangle$ coincides with $\langle\tilde{u}_t,\varphi\rangle$. Hence, for a countable dense set $\mathcal{D}$ in $C_c^\infty(\mathbb{R}^d)$, F_t and $\tilde{u}_t$ must coincide for all t in $\cap_{\varphi\in\mathcal{D}}A_\varphi$, which is still a full measure set in $[0,T]$.

It remains to prove that F satisfies the identity (3.3) for time-dependent test functions φ in $C^1([0,T];C_c^\infty(\mathbb{R}^d))$. To this end we notice that, since F is a representative of $\tilde{u}$, it must verify (3.3) for a.e. $t\in[0,T]$. Moreover, $t\mapsto\langle F_t,\varphi(t)\rangle$ is continuous, which follows from the uniform (in time) bound of the H^{-1} norm of F: indeed $|\langle F_t,\varphi(t)\rangle - \langle F_s,\varphi(s)\rangle| \le |\langle F_t - F_s,\varphi(t)\rangle| + |\langle F_s,\varphi(t)-\varphi(s)\rangle|$; so when $s\to t$, then $|\langle F_t-F_s,\varphi(t)\rangle|\to 0$, as a consequence of the definition of F, and also $|\langle F_s,\varphi(t)-\varphi(s)\rangle|\to 0$, since $\varphi(s)\to\varphi(t)$ and $\sup_{s\in[0,T]}\|F_s\|_{H^{-1}(B_R)}\le C_R$ for every $R>0$. Since the right-hand side of (3.3) is continuous in time as well, we conclude that (3.3) holds in fact for every $t\in[0,T]$, and the proof of the lemma is complete. □

Remark 3.11. *The map $(t,\omega)\mapsto\tilde{u}^\omega(t)$ is actually H^{-1}-weakly-$*$ progressively measurable: indeed, for every test function φ, the weak-$*$ continuity of $\tilde{u}(t)$ implies that the map $(t,\omega)\mapsto\langle\tilde{u}^\omega(t),\varphi\rangle$ can be approximated by simple progressively measurable functions: we can take for instance*

$$\tilde{u}^n(t) = \sum_{j=1}^{\lfloor 2^n T\rfloor} 2^n \int_{t_{j-1}}^{t_j} \tilde{u}(s)ds\, \mathbb{1}_{[t_j,t_{j+1})}(t),$$

where $t_j = 2^{-n}j$, $j\in\mathbb{N}$, is a dyadic partition of $[0,T]$. Analogously, the H^{-1}-weakly-$*$ continuous version of u, defined from the weakly-$*$ continuous representative of $\tilde{u}$ via $u(t)=\tilde{u}(t,\cdot-\sigma W)$, is H^{-1}-weakly-$*$ progressively measurable. Mind that these continuous, distribution-valued versions, do not need to be functions in (t,x,ω).

With this "weakly continuous" representative, we can now state the duality formula for approximations.

Proposition 3.12. *Given $t_f \in [0,T]$, $v_0 \in C_c^\infty(\mathbb{R}^d)$ and $\varepsilon > 0$, let v_ε be the smooth solution of the backward stochastic equation* (3.17)*. For some $\omega \in \Omega$ assume that $v_\varepsilon^\omega \in C^1([0,T]; C_c^\infty(\mathbb{R}^d))$ and that identity* (3.19) *holds for $\widetilde{v}_\varepsilon^\omega$. If $\widetilde{u}^\omega$ is any weak solution of equation* (3.2) *of class $L^m(L^m_{\mathrm{loc}})$ corresponding to that ω, then we have*

$$\begin{aligned}&\big\langle \widetilde{u}^\omega(t_f), v_0(\cdot + \sigma W_{t_f}(\omega))\big\rangle \\ &\qquad = \langle u_0, \widetilde{v}_\varepsilon^\omega(0)\rangle + \int_0^{t_f} \Big\langle \widetilde{u}^\omega(s), (\widetilde{b}^\omega(s) - \widetilde{b}_\varepsilon^\omega(s)) \cdot \nabla \widetilde{v}_\varepsilon^\omega(s) - (\widetilde{c}^\omega(s) - \widetilde{c}_\varepsilon^\omega(s))\widetilde{v}_\varepsilon^\omega(s)\Big\rangle\, ds. \quad (3.20)\end{aligned}$$

Proof. This follows directly from (3.3) (which can be stated for $t = t_f$ fixed, by the previous Lemma 3.10) for $\varphi = \widetilde{v}_\varepsilon^\omega$ and identity (3.19). □

3.4 Path-by-path uniqueness

By linearity of equation (sgTE), in order to prove uniqueness it is sufficient to prove that $u_0 = 0$ implies $u = 0$. To this aim, we will combine identity (3.20) and Corollary 3.7, similarly to the idea explained in Section 3.1 for the deterministic case. The problem in the stochastic case, however, is that we have regularity control in average and we want to deduce path-by-path uniqueness. Let us first state the analog of (3.14). Here, we need to impose $m > 2$ (while the Definition 3.3 of weak L^m-solutions requires only $m \geq 2$).

Lemma 3.13. *Take $m > 2$, $\beta > 0$ and assume Condition* 2.2 *on b and c and Condition* 2.14 *on the families $(b_\varepsilon)_\varepsilon$ and $(c_\varepsilon)_\varepsilon$. Given $t_f \in [0,T]$ and $v_0 \in C_c^\infty(\mathbb{R}^d)$, let $(v_\varepsilon)_\varepsilon$ be the family of smooth solutions of the backward stochastic equation* (3.17)*. Then*

$$\lim_{\varepsilon \to 0} E\bigg[\int_0^{t_f} \int_{\mathbb{R}^d} (1+|x|)^\beta \big(|(b - b_\varepsilon) \cdot \nabla v_\varepsilon| + |(c - c_\varepsilon) v_\varepsilon| \big)^{m'} dx ds \bigg] = 0.$$

Proof. We only prove the convergence in the case $c = 0$, since the terms with c are similar to or easier than those with b. We will prove the assertion for every b and family $(b_\varepsilon)_\varepsilon$ which satisfy

$$\int_0^T \Big(\int_{\mathbb{R}^d} (1+|x|)^{-\alpha} |b - b_\varepsilon|^{\tilde{p}} dx \Big)^{m'/\tilde{p}} ds \to 0 \quad (3.21)$$

for some $\alpha \geq 0$ and $\tilde{p} > m'$. This condition is more general than the LPS condition and includes the cases of

- $b^{(1)}$ as in Condition 2.2 1a), for $p < \infty$, or $b^{(1)}$ as in 1b) or in 1c), with $\alpha = 0$, $\tilde{p} = p$;
- $b^{(1)}$ as in Condition 2.2 1a) for $(p,q) = (\infty, 2)$, with $\alpha > d$, any $\tilde{p}$ with $m' < \tilde{p} < 2$: indeed $|b - b_\varepsilon|$ converges a.e. to 0 and $\int_0^T \int_{\mathbb{R}^d} (1+|x|)^{-\alpha} |b - b_\varepsilon|^2 dx ds$ is uniformly (in ε) bounded (by Hölder's inequality), thus (3.21) follows from Vitali's theorem, in this form: if ν is a finite measure (here $(1+|x|)^{-\alpha} dx ds$), f_n converges to 0 ν-a.e. and $\int |f_n|^a d\nu$ is uniformly bounded for some $a > 1$ (here $a = 2/\tilde{p}$), then f_n tends to 0 in L^1;
- $b^{(2)}$ as in Condition 2.2 2), with $\alpha > d + 2$, $\tilde{p} = 2$ (here we need L^2 integrability in time instead of L^1 for $b^{(2)}$, see Remark 2.3).

Assuming (3.21), we write $\beta = (\beta + \alpha m'/\tilde{p}) - \alpha m'/\tilde{p}$ and apply Hölder's inequality, first in x and ω with exponent $\tilde{p}/m' > 1$, then in time with exponent 1. In this way, we find

$$\begin{aligned}&E\bigg[\int_0^{t_f} \int_{\mathbb{R}^d} (1+|x|)^\beta |(b - b_\varepsilon) \cdot \nabla v_\varepsilon|^{m'} dx ds \bigg] \\ &\leq \bigg(\int_0^{t_f} \Big(\int_{\mathbb{R}^d} (1+|x|)^{-\alpha} |b - b_\varepsilon|^{\tilde{p}} dx \Big)^{m'/\tilde{p}} ds \bigg) \\ &\quad \times \sup_{t \in [0,t_f]} E\bigg[\Big(\int_{\mathbb{R}^d} (1+|x|)^{(\beta\tilde{p} + \alpha m')/(\tilde{p} - m')} |\nabla v_\varepsilon|^{m'\tilde{p}/(\tilde{p}-m')} dx \Big)^{1 - m'/\tilde{p}} \bigg].\end{aligned}$$

Now Corollary 3.7 gives that the second term is uniformly bounded and we get the claim. □

For the following uniqueness statement we have to restrict the behavior at infinity of weak L^m-solutions (in the definition, they are just $L^m_{\mathrm{loc}}(\mathbb{R}^d)$). The restriction is not severe: we just need at most polynomial growth at infinity. To be precise, we need that for some $\alpha > 0$ we have

$$\int_0^T \int_{\mathbb{R}^d} \frac{1}{1+|x|^\alpha} |\widetilde{u}^\omega(t,x)|^m dx dt < \infty. \tag{3.22}$$

Theorem 3.14. *Assume $m > 2$. There exists a full measure set $\Omega_0 \subset \Omega$ such that for all $\omega \in \Omega_0$ the following property holds: for every $u_0 \colon \mathbb{R}^d \to \mathbb{R}$ such that $\int_{\mathbb{R}^d} (1+|x|^\alpha)^{-1} |u_0(x)|^m dx < \infty$ for some $\alpha > 0$, equation (3.2) has at most one weak solution $\widetilde{u}^\omega \colon [0,T] \times \mathbb{R}^d \to \mathbb{R}$ of class $L^m(L^m_{\mathrm{loc}})$ which satisfies the additional condition (3.22).*

Proof. Step 1: Identification of Ω_0. From Lemma 3.13, given $t_f \in [0,T]$ and $v_0 \in C_c^\infty(\mathbb{R}^d)$, there exist a full measure set $\Omega_{t_f,v_0} \subset \Omega$ and a sequence $(\varepsilon_n)_{n\in\mathbb{N}}$ with $\varepsilon_n \to 0$ as $n \to \infty$ such that

$$\begin{gathered} \widetilde{v}^\omega_{\varepsilon_n} \text{ belongs to } C^1([0,T]; C_c^\infty(\mathbb{R}^d)) \text{ and satisfies (3.19), for all } n \in \mathbb{N}, \\ \lim_{n\to\infty} \big\| (1+|\cdot|^{\alpha/m})(b^\omega - b^\omega_{\varepsilon_n}) \cdot \nabla v^\omega_{\varepsilon_n} \big\|_{L^{m'}([0,t_f]\times\mathbb{R}^d)} = 0 \end{gathered} \tag{3.23}$$

for all $\omega \in \Omega_{t_f,v_0}$. Hence, we also have

$$\lim_{n\to\infty} \big\| (1+|\cdot + \sigma W(\omega)|^{\alpha/m})(\widetilde{b}^\omega - \widetilde{b}^\omega_{\varepsilon_n}) \cdot \nabla \widetilde{v}^\omega_{\varepsilon_n} \big\|_{L^{m'}([0,t_f]\times\mathbb{R}^d)} = 0 \tag{3.24}$$

for all $\omega \in \Omega_{t_f,v_0}$. By a diagonal procedure, given a countable set $\mathcal{D}$ in $C_c^\infty(\mathbb{R}^d)$ which is dense in $L^2(\mathbb{R}^d)$, there exist a full measure set $\Omega_b \subset \Omega$ and a sequence $(\varepsilon_n)_{n\in\mathbb{N}}$ with $\varepsilon_n \to 0$ as $n \to \infty$ such that properties (3.23) and (3.24) hold for all $t_f \in [0,T] \cap \mathbb{Q}$, $v_0 \in \mathcal{D}$ and $\omega \in \Omega_b$. Since, for a given ω, there exists a constant $C_\omega > 0$ such that

$$(1+|x|^{\alpha/m}) \le C_\omega (1+|x+\sigma W_t(\omega)|^{\alpha/m})$$

for all $x \in \mathbb{R}^d$ and $t \in [0,T]$, we may replace (3.24) by

$$\lim_{n\to\infty} \big\| (1+|\cdot|^{\alpha/m})(\widetilde{b}^\omega - \widetilde{b}^\omega_{\varepsilon_n}) \cdot \nabla \widetilde{v}^\omega_{\varepsilon_n} \big\|_{L^{m'}([0,t_f]\times\mathbb{R}^d)} = 0. \tag{3.25}$$

An analogous selection is possible for $(1+|\cdot|^{\alpha/m})(\widetilde{c}^\omega - \widetilde{c}^\omega_{\varepsilon_n})\widetilde{v}^\omega_{\varepsilon_n}$, which provides another full measure set Ω_c, and Ω_0 is then defined as the intersection $\Omega_b \cap \Omega_c$.

Step 2: Path-by-path uniqueness for equation (3.2) on Ω_0. Given $\omega \in \Omega_0$ and a weak solution $\widetilde{u}^\omega$ to (3.2) of class $L^m(L^m_{\mathrm{loc}})$ with $u_0 = 0$, by identity (3.20) and property (3.23) we have

$$\big\langle \widetilde{u}^\omega(t_f), v_0(\cdot + \sigma W_{t_f}(\omega)) \big\rangle = \int_0^{t_f} \Big\langle \widetilde{u}^\omega, (\widetilde{b}^\omega - \widetilde{b}^\omega_{\varepsilon_n}) \cdot \nabla \widetilde{v}^\omega_{\varepsilon_n} - (\widetilde{c}^\omega - \widetilde{c}^\omega_{\varepsilon_n}) \widetilde{v}^\omega_{\varepsilon_n} \Big\rangle ds$$

for all $t_f \in [0,T] \cap \mathbb{Q}$, $v_0 \in \mathcal{D}$ and $n \in \mathbb{N}$. Now we pass to the limit. By Hölder's inequality we get

$$\begin{aligned} &\big| \big\langle \widetilde{u}^\omega(t_f), v_0(\cdot + \sigma W_{t_f}(\omega)) \big\rangle \big| \\ &\le \Big(\int_0^{t_f} \int_{\mathbb{R}^d} \frac{1}{(1+|x|^{\alpha/m})^m} |\widetilde{u}^\omega(s,x)|^m dx ds \Big)^{1/m} \\ &\quad \times \Big(\int_0^{t_f} \int_{\mathbb{R}^d} (1+|x|^{\alpha/m})^{m'} \big| (\widetilde{b}^\omega - \widetilde{b}^\omega_{\varepsilon_n}) \cdot \nabla \widetilde{v}^\omega_{\varepsilon_n} - (\widetilde{c}^\omega - \widetilde{c}^\omega_{\varepsilon_n}) \widetilde{v}^\omega_{\varepsilon_n} \big|^{m'} dx ds \Big)^{1/m'} \end{aligned}$$

and thus $\big\langle \widetilde{u}^\omega(t_f), v_0(\cdot + \sigma W_{t_f}(\omega)) \big\rangle = 0$ by (3.25) (and the analogue for $(1+|\cdot|^{\alpha/m})(\widetilde{c}^\omega - \widetilde{c}^\omega_{\varepsilon_n})\widetilde{v}^\omega_{\varepsilon_n}$). This is equivalent to $\big\langle \widetilde{u}^\omega(t_f, \cdot - \sigma W_{t_f}(\omega)), v_0 \big\rangle = 0$, which implies $\widetilde{u}^\omega(t_f, \cdot - \sigma W_{t_f}(\omega)) = 0$ by the density of $\mathcal{D}$ in $L^2(\mathbb{R}^d)$ and thus $\widetilde{u}^\omega(t_f, \cdot) = 0$. This holds true for every $t_f \in [0,T] \cap \mathbb{Q}$; since $t \mapsto \widetilde{u}^\omega(t)$ is continuous in the sense of distributions, we get $\widetilde{u}^\omega(t,\cdot) = 0$ for every $t \in [0,T]$. The proof of the theorem is complete. □

3.5 Existence for (3.1)

So far we have proved that path-by-path uniqueness holds for the stochastic equation (3.1). It remains to prove existence of a (distributional) solution. The proof is based on a priori estimates and is somehow similar to that of Theorem 2.7 and Theorem 2.17, without the difficulty of taking derivatives. Thus, we will state the result and only sketch the proof.

Proposition 3.15. *Let p, q be in $(2,\infty)$ satisfying $\frac{2}{q}+\frac{d}{p}\leq 1$ or $(p,q)=(\infty,2)$. Assume that b and c are a vector field and a scalar field, respectively, such that $b=b^{(1)}+b^{(2)}$, $c=c^{(1)}+c^{(2)}$, with $b^{(i)}$, $c^{(i)}$ in $C_c^\infty([0,T]\times\mathbb{R}^d)$ for $i=1,2$. Let χ be a function satisfying (2.8). Then there exists a constant C such that, for every u_0 in $C_c^\infty(\mathbb{R}^d)$, the smooth solution u of equation (3.1) starting from u_0, given by Lemma 2.6, verifies*

$$\sup_{t\in[0,T]}\int_{\mathbb{R}^d} E\left[(u(t,x))^m\right]^2\chi(x)dx\leq C\left\|u_0\right\|^{2m}_{L^{2m}_\chi(\mathbb{R}^d)}.$$

Moreover, the constant C can be chosen to have continuous dependence on m,d,σ,χ,p,q and on the $L^q([0,T];L^p(\mathbb{R}^d))$ norms of $b^{(1)}$ and $c^{(1)}$, on the $L^1([0,T];C^1_{lin}(\mathbb{R}^d))$ norm of $b^{(2)}$, and on the $L^1([0,T];C^1_b(\mathbb{R}^d))$ norm of $c^{(2)}$.
The result holds also for $(p,q)=(d,\infty)$ with the additional hypothesis that the $L^\infty([0,T];L^d(\mathbb{R}^d))$ norms of $b^{(1)}$ and $c^{(1)}$ are smaller than δ, see Condition 2.2, 1c) (in this case the continuous dependence of C on these norms is up to δ).

Proof. We proceed similarly as in the the proof of Theorem 2.7, but aiming for a priori estimates for u and not for its derivatives. To this end, we consider the equation for $E[u^m]$, which is a parabolic closed equation. The same method of proof as in Theorem 2.7 can then be applied (without the difficulty of having a system with many indices), which then shows the claim.. □

Theorem 3.16. *Let $m\geq 4$ be an even integer and let s be a real number. Assume that b, c satisfy Condition 2.2 and let $u_0\in L^{2m}_{(1+|\cdot|)^{2s+d+1}}(\mathbb{R}^d)$. There exists a weak solution u to equation (3.1) of class $L^m(L^m_{\mathrm{loc}})$. Moreover,*

$$\operatorname*{ess\,sup}_{t\in[0,T]} E\left[\left\|u(t,\cdot)\right\|^m_{L^m_{(1+|\cdot|)^s}(\mathbb{R}^d)}\right]<\infty.$$

Finally, pathwise uniqueness holds among such solutions and actually among all solutions such that $\tilde{u}$ satisfies (3.22) for a.s..

Proof. The existence of a weak solution to equation (3.1) of class $L^m(L^m_{\mathrm{loc}})$ follows by the same arguments as in the proof of Theorem 2.17. The main differences are that weak-$*$ convergence holds in $L^\infty([0,T];L^m(\Omega;L^m(B_R)))$ instead of in $L^\infty([0,T];L^m(\Omega;W^{1,m}(B_R)))$ and that all the derivatives must be carried over to the test function φ.

Pathwise uniqueness follows from Theorem 3.14: Let u, u_1 be two solutions to (3.1) satisfying (3.22) on the same filtered probability space $(\Omega,(\mathcal{G}_t)_t,P)$, such that W is a Brownian motion with respect to $(\mathcal{G}_t)_t$. Then, according to Lemma 3.4, the function $\tilde{u}_1$ given by $\tilde{u}_1(t,x)=u_1(t,x+\sigma W_t)$ solves the deterministic PDE (3.2) for a.e. $\omega\in\Omega$, so $\tilde{u}_1$ must coincide with $\tilde{u}$ for a.e. (t,x,ω), which implies the claim $u_1=u$. □

Remark 3.17. *For solution to equation (3.1), non-negativity of initial values is preserved, i.e. if $u_0\geq 0$, then $u(t,x,\omega)\geq 0$ for a.e. (t,x,ω): this is true in the regular case (for u, b and c smooth and compactly supported), thanks to the representation formula (2.5) (where, for the application to equation (3.1), c is replaced by $c+\operatorname{div} b$). This carries over to the general case, since u is constructed as weak-$*$ limit in $L^\infty([0,T];L^m(\Omega;W^{1,m}(B_R)))$ of solutions with regularized coefficients and initial condition.*

4 Path-by-path uniqueness and regularity of the flow solving the sDE

In this section we want to apply the previous results to study equation (sDE). We will get existence, strong (even path-by-path) uniqueness and regularity for the stochastic flow solving the sDE,

where b is in the LPS class and $\sigma \neq 0$. Once again, we recall that no such result holds in the deterministic case ($\sigma = 0$), which means that for the stochastic case ($\sigma \neq 0$) the evolution of the finite-dimensional system gets better due to the additional stochastic term.

In order to state the result, we need to make the formal links between (sDE) and (sCE) precise. This will be done for the deterministic case, in the first subsection, using Ambrosio's theory of Lagrangian flows. Then we will use this link (read in a proper way in the stochastic case) combined with uniqueness and regularity for the stochastic equations to arrive at our result.

4.1 The deterministic case

Consider the ODE

$$\frac{d}{dt}X = f(t, X) \tag{4.1}$$

on $\mathbb{R}^d$. If f is a regular field (e.g. $C_c^1([0,T] \times \mathbb{R}^d)$), there exists a unique flow $\Phi\colon [0,T] \to \mathrm{Diff}(\mathbb{R}^d)$ of diffeomorphisms on $\mathbb{R}^d$ solving the ODE, i.e. for every x in $\mathbb{R}^d$, $t \mapsto \Phi(t,x)$ is of class C^1 and solves the ODE starting from $\Phi(0,x) = x$. If φ is a test function in $C_c^\infty(\mathbb{R}^d)$, then the chain rule gives the following equation for $\varphi(\Phi)$:

$$\frac{d}{dt}\varphi(\Phi_t) = \nabla\varphi(\Phi_t) \cdot f(t, \Phi_t).$$

If we integrate this equation with respect to a finite signed measure μ_0 on $\mathbb{R}^d$, we get

$$\langle \mu_t, \varphi \rangle = \langle \mu_0, \varphi \rangle + \int_0^t \langle \mu_s, f(s,\cdot) \cdot \nabla\varphi \rangle ds, \tag{4.2}$$

where $\mu_t = (\Phi_t)_\# \mu_0$ is the image measure on $\mathbb{R}^d$ of μ_0 under Φ_t, i.e. $\int g d\mu_t = \int g(\Phi_t) d\mu_0$ for every measurable bounded function g on $\mathbb{R}^d$. Equation (4.2) is the continuity equation (CE) for μ (starting from μ_0), which we have written so far in compact form as

$$\partial_t \mu + \mathrm{div}(f\mu) = 0. \tag{CE}$$

It is easy to see that the previous passages still hold when f is not regular. Starting from this remark, DiPerna–Lions' and Ambrosio's theory extends the above link between the ODE and the CE (in some generalized sense) to the irregular case, so that one can study the CE in order to study the ODE. We will follow Ambrosio's theory of Lagrangian flows, which allows to transfer a well-posedness result for the CE to a well-posedness result for the ODE.

In the general theory, one considers a convex set $\mathcal{L}_f$ of solutions $\mu = (\mu_t)_t$ to the equation (CE), with values in the set $\mathcal{M}_+(\mathbb{R}^d)$ of finite positive measures on $\mathbb{R}^d$, which satisfies $\int_0^T \int_{\mathbb{R}^d} (1+|x|)^{-1} |f(t,x)| \mu_t(dx)dt < \infty$ for every μ in $\mathcal{L}_f$ and

$$0 \leq \mu_t' \leq \mu_t,\ \mu \in \mathcal{L}_f,\ \mu' \text{ solves (CE)} \ \Rightarrow \mu' \in \mathcal{L}_f \tag{4.3}$$

("solution of (CE)" is here intended in the sense of distributions). For our purposes, $\mathcal{L}_f$ will be, for some m fixed a priori, the set

$$\mathcal{L}_f = \Big\{ \mu = (\mu_t)_t \colon \mu \text{ solves (CE)},\ \mu_t = u_t \mathcal{L}^d \text{ for some non-negative } u \in L^1 \cap L^m([0,T] \times \mathbb{R}^d), \int_0^T \int_{\mathbb{R}^d} \frac{|f(t,x)|}{1+|x|} \mu_t(dx)dt < \infty \Big\}.$$

Sometimes, we will use u to indicate also μ and viceversa.

Definition 4.1. *A $\mathcal{L}_f$ (local) Lagrangian flow, starting from some fixed (non-negligible) Borel set S in $\mathbb{R}^d$, is a Borel map $\Phi\colon [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ such that*

- *for $\mathcal{L}^d$-a.e. x in S, for every t, $\Phi_t(x) = x + \int_0^t f(s, \Phi_s(x))ds$;*
- *$\mu_t \coloneqq (\Phi_t)_\#(\mathbb{1}_S \mathcal{L}^d)$ is in $\mathcal{L}_f$.*

A $\mathcal{L}_f$ global Lagrangian flow *is a Borel map $[0,T]\times\mathbb{R}^d\to\mathbb{R}^d$ which is a local Lagrangian flow from every (non-negligible) Borel set S.*

Theorem 4.2. *Suppose that uniqueness holds for the CE, starting from every $\mathbb{1}_S\mathcal{L}^d$, in the class $\mathcal{L}_f$. Then local uniqueness (i.e. uniqueness from every S) holds among $\mathcal{L}_f$ Lagrangian flows (that is, if Φ^1 and Φ^2 are two such flows, then, for a.e. x in S: for every t, $\Phi^1_t(x)=\Phi^2_t(x)$). If in addition existence holds for the CE in the class $\mathcal{L}_f$ starting from $\mathbb{1}_{B_N}\mathcal{L}^d$, for every positive integer N, then there exists a global Lagrangian flow.*

This theorem is stated and proved in [3, Theorem 18]. We give here a concise proof (similar to the one in [3]) only of uniqueness, since the existence part is more technical and long (though not difficult). The idea for uniqueness is again to use the link between the ODE and the CE: whenever one has two flows Φ^1 and Φ^2, then $(\Phi^1_t)_\#\mathbb{1}_S\mathcal{L}^d$ and $(\Phi^2_t)_\#\mathbb{1}_S\mathcal{L}^d$ are solutions to (CE) in the class $\mathcal{L}_f$, so, by uniqueness, they must coincide, so that Φ^1 and Φ^2 coincide on S.

Proof. By continuity in time of the Lagrangian flow (for a.e. fixed x), it is enough to show that, given two $\mathcal{L}_f$ Lagrangian flows Φ^1, Φ^2 starting from the same S, then, for every t, we have $\Phi^1_t(x)=\Phi^2_t(x)$ for a.e. $x\in S$. Suppose by contradiction this is not the case. Then there exist a time $t\in[0,T]$, two disjoint Borel sets E^1, E^2 in $\mathbb{R}^d$ and a Borel set S' in S with $0<\mathcal{L}^d(S')<\infty$ such that $\Phi^i_t(x)$ is in E^i for every x in S', $i=1,2$. Define μ^i_t as $(\Phi^i_t)_\#\mathbb{1}_{S'}\mathcal{L}^d$ for $i=1,2$. Then μ^1 and μ^2 are maps from $[0,T]$ to $\mathcal{M}_+(\mathbb{R}^d)$, which are weakly continuous solutions to (CE), still in the class $\mathcal{L}_f$ (as they are restrictions of $(\Phi^1_t)_\#\mathbb{1}_S\mathcal{L}^d$ and $(\Phi^2_t)_\#\ \mathbb{1}_S\mathcal{L}^d$), and differ at least in one point t. This is a contradiction and uniqueness is proved. □

Remark 4.3. *If existence and uniqueness (starting from every $\mathbb{1}_S\mathcal{L}^d$ in the class $\mathcal{L}_f$) hold for* (CE) *and if f is in $L^2_{\mathrm{loc}}([0,T]\times\mathbb{R}^d)$, then for every $t\in[0,T]$ there holds $\langle u_0,\varphi(\Phi_t)\rangle=\langle u_t,\varphi\rangle$, where we have denoted by Φ the Lagrangian flow and by u the H^{-1}-weakly-$*$ continuous version of the solution to* (CE)*. Indeed, $t\mapsto(\Phi_t)_\#u_0$ is also a H^{-1}-weakly-$*$ continuous solution too* (CE)*, hence it must coincide with u at all times.*

Having Theorem 4.2 at our disposal, we can employ existence and uniqueness for the CE in the class $\mathcal{L}_f$ in order to prove well-posedness for the Lagrangian flow associated with the ODE. This is what DiPerna–Lions and Ambrosio have done (using mainly the transport equation instead of the CE and with a different $\mathcal{L}_f$) for weakly differentiable functions f ([29, 2]). We will follow this strategy, but for f in LPS class and with noise, using our well-posedness result for (sCE).

4.2 Stochastic Lagrangian flow: existence, uniqueness and regularity

We consider the equation (sDE) on $\mathbb{R}^d$. Since we use also here the results of the previous sections, we again assume the same LPS Condition 2.2 on b. As before, we consider the purely stochastic case $\sigma\neq0$, and W is a standard d-dimensional Brownian motion, endowed with its natural completed filtration $(\mathcal{F}_t)_t$ (the smallest among all the possible filtrations), which is also right-continuous (see [8, Proposition 2.5]).

With the change of variable $\tilde{X}_t=X_t-\sigma W_t$, this sDE becomes a family of (random) ODEs, parametrized by $\omega\in\Omega$:

$$\frac{d}{dt}\tilde{X}=\tilde{b}^\omega(t,\tilde{X}),\tag{4.4}$$

where, as usual, $\tilde{b}^\omega(t,x)=b(t,x+\sigma W_t(\omega))$. More precisely, if X is a progressively measurable process, then X solves (sDE) if and only if $\tilde{X}$ solves the ODE (4.4) for a.e. ω. For this family of ODEs, the concepts of Lagrangian flow and CE (at ω fixed) make sense and the CE associated with this ODE is precisely the random PDE (3.2) with $c=0$, that is the random CE

$$\partial_t\tilde{u}+\mathrm{div}(\tilde{b}\tilde{u})=0.\tag{4.5}$$

Thus we can hope to apply our existence and uniqueness result for (sCE) (remembering that, by Lemma 3.4, a solution to the random CE (4.5) is given by $\tilde{u}(t,x)=u(t,x+\sigma W_t)$, when u solves (sCE)).

Definition 4.4. *A* stochastic (global) Lagrangian flow solving the equation (sDE) *is a measurable map* $\Phi\colon [0,T]\times\mathbb{R}^d\times\Omega\to\mathbb{R}^d$ *with the following properties:*

- *for a.e.* $\omega\in\Omega$, $(t,x)\mapsto\tilde{\Phi}^\omega_t(x):=\Phi^\omega_t(x)-\sigma W_t(\omega):=\Phi(t,x,\omega)-\sigma W_t(\omega)$ *is a* $\mathcal{L}_{\tilde{b}^\omega}$ *(global) Lagrangian flow (solving the ODE* (4.4) *with that* ω *fixed);*
- Φ *is progressively measurable, i.e. it is* $\mathcal{P}\otimes\mathcal{B}(\mathbb{R}^d)$*-measurable, where* $\mathcal{P}$ *is the progressive* σ*-algebra.*

Given a certain class A *of functions from* $\mathbb{R}^d$ *to* $\mathbb{R}^d$*, e.g.* $W^{1,m}_{\mathrm{loc}}(\mathbb{R}^d)$*, the flow is said to be of class* A *if, for every* $t\in[0,T]$, Φ_t *is in class* A *with probability* 1.

Let us now state the main result of this section:

Theorem 4.5. *Let* $m\geq 4$ *be an even integer and assume that* b *verifies Condition* 2.2*. Then*

1. *local path-by-path uniqueness holds among Lagrangian flows solving* (sDE)*, i.e., for a.e.* $\omega\in\Omega$*, local uniqueness holds among* $\mathcal{L}_{\tilde{b}^\omega}$ *Lagrangian flows solving the ODE* (4.4) *with that* ω *fixed;*
2. *there exists a global stochastic Lagrangian flow solving* (sDE)*;*
3. *this flow is of class* $W^{1,m}_{\mathrm{loc}}(\mathbb{R}^d)$*.*

Remark 4.6. *If* $m>d$*, we deduce by Sobolev embedding that, for every* $t\in[0,T]$*, there exists a representative of* Φ_t *which is of class* $C^{0,\alpha}_{\mathrm{loc}}(\mathbb{R}^d)$ *for* $\alpha=1-d/m$*. However, we are not able to show that this representative is jointly continuous in* (t,x) *(we actually do not even show joint measurability), though such joint continuity is known to be true in the subcritical case (see* [35]*).*

Before proceeding to the proof, which is essentially an application of our well-posedness result for the deterministic PDE (3.2), we make some comments on this result.

Remark 4.7. *The existence part gives essentially a family a flows* Φ^ω*, parametrized by* ω *in* Ω*, such that* $\Phi(x)$ *solves* (sDE) *for a.e.* x*. The regularity part gives local weak differentiability of the flow and, if* $m>d$*, even Hölder continuity for a.e.* ω *and every* t *(unfortunately, the full* P*-measure set* Ω_t *where this property holds can depend on* t*). The uniqueness part implies pathwise uniqueness among stochastic Lagrangian flows: given stochastic Lagrangian flows* Φ^1 *and* Φ^2 *solving* (sDE) *(even adapted to some filtration larger than* $(\mathcal{F}_t)_t$*) and starting from the same initial datum of the form* $\mathbb{1}_S\mathcal{L}^d$*, they necessarily coincide. Indeed, for a.e.* ω*,* $\tilde{\Phi}^1(\omega)$ *and* $\tilde{\Phi}^2(\omega)$ *are Lagrangian flows solving the random ODE* (4.4) *(with that* ω*), so* $\Phi^1=\Phi^2$ *a.e..*

Let us emphasize that path-by-path uniqueness is stronger than pathwise uniqueness: it says that, for each fixed ω *in a full* P*-measure set, any two Lagrangian flows, solving* (sDE) *(interpreted as the random ODE* (4.4)*) with* ω *fixed, must coincide, without any need to have adapted flows. On the other hands, while we can manage flows, we are not able to compare two solutions to the ODE* (4.4)*, at* ω *fixed, starting from a fixed* x*, so we have no uniqueness result for* (sDE) *with* x *as initial datum. Let us remind however that pathwise uniqueness holds for* (sDE) *(with* x *fixed) under Krylov-Röckner conditions, see* [54].

Remark 4.8. *We wish to recall a basic argument in measure theory, that we will use quite often: Let* $(E,\mathcal{E},\mu)$*,* $(F,\mathcal{F},\nu)$ *be two* σ*-finite measure spaces and let* $f\colon E\times F\to\mathbb{R}$ *be a map such that, for* ν*-a.e.* z*, the map* $y\mapsto f(y,z)$ *is* $\mathcal{E}$*-measurable. Assume that* f *has a* $\mathcal{E}\otimes\mathcal{F}$*-measurable version* $g\colon E\times F\to\mathbb{R}$*, i.e. there exist a full measure set* F_0 *and, for every* $z\in F_0$*, a full measure set* E^z_0*, such that* $f(y,z)=g(y,z)$ *for all* $z\in F_0$ *and* $y\in E^z_0$*. Let* $BP=BP(a)$ *be a Borel property defined for* $a\in\mathbb{R}$ *(in the sense that the subset where* BP *is true is a Borel set), for example* $\varphi(a)=0$ *for some Borel function* φ*. Assume that, for* ν*-a.e.* z*, it holds:* $BP(f(y,z))$ *for* μ*-a.e.* y*. Then* $BP(g(y,z))$ *holds for* $(\mu\times\nu)$*-a.e.* (y,z)*. A similar property also holds for more than two variables.*

Proof. If this were not true, then the set $A=\{(y,z)\in E\times F\colon \neg BP(g(y,z))\}$ is $\mathcal{E}\otimes\mathcal{F}$-measurable (by measurability of BP and g) and of positive measure. Therefore, by Fubini theorem, there exists a positive measure set $F_\neg\subset F$ such that, for every $z\in F_\neg$, the set $E^z_\neg:=\{y\in E\colon \neg BP(g(y,z))\}$ is $\mathcal{E}$-measurable and of positive measure. But g is a version of f. Therefore, for every z in the positive measure set $F_\neg\cap F_0$, the set $\{y\in E\colon \neg BP(f(y,z))\}$ contains the positive measure set $E^z_\neg\cap E^z_0$, which is in contradiction with the assumption on $BP(f)$. □

Proof of Theorem 4.5. *Part 1: Uniqueness of Lagrangian flows solving the ODE* (4.4). Theorem 3.14, applied to the random CE (4.5), gives a full P-measure set Ω_0 in Ω such that, for every $\omega \in \Omega_0$, for every Borel set S, there exists at most one solution $\tilde{u}^\omega$ to the CE in the class $\mathcal{L}_{\tilde{b}^\omega}$, which starts from $\mathbb{1}_S$ (note that Ω_0 is independent of the initial datum). Thus, for every $\omega \in \Omega_0$, the first part of Theorem 4.2 gives local uniqueness among Lagrangian flows solving (4.4) at ω fixed.

Part 2: Existence of a global stochastic Lagrangian flow. The idea is to proceed in three steps and use Ambrosio's theory for the random ODE (4.4) to get the existence, at ω fixed, of a Lagrangian flow, then to use the progressive measurability of the solution to the sCE (sCE) to show progressive measurability of (a version of) the Lagrangian flow and to conclude. As we are going to take various modifications of the same function, we keep the following convention: we use the notation $\bar{\Phi}$ for a solution of the sDE which is continuous in time (at x fixed), but not necessarily measurable in ω, the notation $\tilde{\Phi}$ for a solution of the random ODE and the notation $\bar{\tilde{\Phi}}$ for a solution of the random ODE which is also continuous in time (again at x fixed). For the solution to the (s)CE, we do not use the "bar" as we consider, unless otherwise stated, versions that are both weakly-$*$ measurable and weakly-$*$ continuous.

In the *first step*, we get the existence, at ω fixed, of a Lagrangian flow $\bar{\tilde{\Phi}}^\omega$ solving the ODE (4.4). We take $S = B_N$ for an arbitrary positive integer N. By Theorem 3.16 (applied with $c = 0$), Remark 3.17 and Proposition 3.4, we find a full P-measure set Ω_0 in Ω, independently of N (by a diagonal procedure), such that, for every $\omega \in \Omega_0$ and N, there exists a (unique) solution $\tilde{u}^{\omega,N}$ to the CE (4.5) in the class $\mathcal{L}_{\tilde{b}^\omega}$, starting from $\mathbb{1}_{B_N}$. Thus, the second part of Theorem 4.2 gives the claimed existence of a global Lagrangian flow $\bar{\tilde{\Phi}}^\omega$ solving the ODE (4.4).

Now we define $\bar{\Phi} = \bar{\tilde{\Phi}} + \sigma W$, which seems at first the natural candidate for the stochastic Lagrangian flow solution to (sDE). The main problem is that $\bar{\Phi}$ does not have any measurability property in ω. Therefore, in the *second step*, we find a progressively measurable map $\Phi\colon [0,T] \times \mathbb{R}^d \times \Omega \to \mathbb{R}^d$ version on $\bar{\Phi}$, that is $P(\bar{\Phi}(t,x,\omega) = \Phi(t,x,\omega)$ for a.e. $(t,x)) = 1$ (keep in mind that this set is not a priori measurable in ω). To this end, we shall use the link between ODE and CE (at the deterministic level) and the progressive measurability of the solution to (sCE).

In what follows, we denote by φ_n functions in $C_c^\infty(\mathbb{R}^d)$ with $\varphi_n(x) = x$ for $|x| \le n$. By the deterministic theory (Theorem 4.2 and Remark 4.3), we know that, for every n and $u_0 \in C_c^\infty(\mathbb{R}^d)$, we have for every ω in a full measure set $\Omega_{u_0,n}$: for every $t \in [0,T]$ there holds $\langle u_0, \varphi_n(\bar{\tilde{\Phi}}_t^\omega)\rangle = \langle \tilde{u}^\omega(t), \varphi_n\rangle$ and so $\langle u_0, \varphi_n(\bar{\Phi}_t^\omega)\rangle = \langle u^\omega(t), \varphi_n\rangle$, where $u^\omega(t)$ is the H^{-1}-weakly-$*$ continuous version, as in Remark 3.11, of the solution to (sCE) starting from u_0. In particular, the map $(t,\omega) \mapsto \langle u_0, \varphi_n(\bar{\Phi}_t^\omega)\rangle$ coincides with a progressively measurable map for every $t \in [0,T]$, for a.e. ω (with the exceptional set independent of t) for every $u_0 \in C_c^\infty(\mathbb{R}^d)$. Hence, up to redefining $\bar{\Phi}_t^\omega$ on a P-null set independent of t, $\langle u_0, \varphi_n(\bar{\Phi}_t^\omega)\rangle$ is progressively measurable for every φ in $C_c^\infty(\mathbb{R}^d)$ and thus by density also for every $u_0 \in L^2(B_R)$. Therefore, $\varphi_n(\bar{\Phi}_t^\omega)$ is weakly-$*$ progressively measurable in $L^2(B_R)$. Since $L^2(B_R)$ is a separable reflexive space, Pettis measurability theorem applies and gives that $\varphi_n(\bar{\Phi}_t^\omega)$ is strongly progressively measurable with values in $L^2(B_R)$, in particular there exists $\Phi_n\colon [0,T] \times B_R \times \Omega \to \mathbb{R}^d$, $\mathcal{P} \otimes \mathcal{B}(\mathbb{R}^d)$-measurable, version of $\varphi_n(\bar{\Phi})$, that is, for a.e. (t,ω) there holds $\Phi_n(t,x,\omega) = \varphi_n(\bar{\Phi}_t^\omega(x))$ for a.e. $x \in B_R$ (cf. [62, Proposition A.6]). Using Remark 4.8 and the analogous properties for $\bar{\Phi}$, one can check that Φ_n does not depend on R and is definitively constant in n (for a.e. (t,x,ω)), so we get Φ, which is $\mathcal{P} \times \mathcal{B}(\mathbb{R}^d)$-measurable and a version of $\bar{\Phi}$. The second step is complete.

To conclude the proof of existence, we have to prove that $\Phi - \sigma W$ is a (global) Lagrangian flow solving (4.4). However, since Φ coincides with $\bar{\Phi}$ only for a.e. (t,x) (for fixed ω), $\Phi(\cdot,x,\omega) - \sigma W(\omega)$ does not need to be continuous in time and satisfies (4.4) only for a.e. $t \in [0,T]$. In the *third step*, we prove that there exists a measurable version of Φ, and so of $\bar{\Phi}$, which is continuous in t for a.e. (x,ω), and use this version to conclude. The conceptual idea is that, given a path γ which has a continuous version, its continuous version can be constructed from γ in a measurable way, so that this version is measurable (with respect to some other variable) if γ is measurable.

For any $N \in \mathbb{N}$, we choose a dyadic partition $t_j^N = 2^{-N}j$, we set $I_j^N := [t_j^N, t_{j+2}^N)$ and, for $t \in [0,T]$, we define $I^N(t)$ as I_j^N for the minimal j with $t \in I_j^N$. Note that, for a.e. ω, it holds: for a.e. (t,x), $\Phi(t,x,\omega) = \bar{\Phi}(t,x,\omega)$ (as Φ is a modification of $\bar{\Phi}$ and both are measurable in (t,x)

for ω fixed). In particular, for a.e. ω, it holds: for a.e. x,

$$\max_j \big[\operatorname*{ess\,sup}_{t\in I_j^N} \Phi(t,x,\omega) - \operatorname*{ess\,inf}_{t\in I_j^N} \Phi(t,x,\omega)\big] = \max_j \big[\operatorname*{ess\,sup}_{t\in I_j^N} \bar\Phi(t,x,\omega) - \operatorname*{ess\,inf}_{t\in I_j^N} \bar\Phi(t,x,\omega)\big].$$

The continuity property of $\bar\Phi$ implies that for a.e. ω the following is true: for a.e. x and every $m\in\mathbb{N}$, there exists $N\in\mathbb{N}$ such that $\max_j[\operatorname{ess\,sup}_{t\in I_j^N}\bar\Phi(t,x,\omega) - \operatorname{ess\,inf}_{t\in I_j^N}\bar\Phi(t,x,\omega)] < 1/m$. Therefore, by Remark 4.8, the set

$$\bigcap_{m\in\mathbb{N}}\bigcup_{N\in\mathbb{N}} \big\{(x,\omega)\colon \max_j \big[\operatorname*{ess\,sup}_{t\in I_j^N} \Phi(t,x,\omega) - \operatorname*{ess\,inf}_{t\in I_j^N} \Phi(t,x,\omega)\big] < 1/m\big\}$$

has full measure. Then, for a.e. (x,ω), the limit

$$A(t,x,\omega) = \lim_{N\to\infty} \operatorname*{ess\,sup}_{s\in I^N(t)} \Phi(s,x,\omega)$$

is well-defined and finite for every t. Moreover, the map A (defined zero on the exceptional set where the above limit does not exist) is measurable in (t,x,ω), and continuous in t for a.e. (x,ω). For a.e. ω it holds: $A(t,x,\omega)=\bar\Phi(t,x,\omega)$ for a.e. (t,x) (since $\bar\Phi(t,x,\omega) = \lim_{N\to\infty}\operatorname{ess\,sup}_{s\in I^N(t)}\bar\Phi(s,x,\omega)$ for a.e. (t,x)). So, again by Remark 4.8, $A=\Phi$ for a.e. (t,x,ω). With a little abuse of notation, we will use now Φ also for its modification which is continuous in t.

It remains to show that $\tilde\Phi = \Phi - \sigma W$ is a Lagrangian flow solving (4.4). The integrand $b(\tilde\Phi)$ is in $L^1(0,T)$ for a.e. (x,ω) and the ODE (4.4) is satisfied for a.e. (t,x,ω): otherwise, since Φ is a version of $\bar\Phi$, reasoning as in Remark 4.8, for some ω in a positive measure set, the ODE would not be satisfied even by $\bar{\tilde\Phi}$. The continuity in time implies that, for a.e. (x,ω), the ODE (4.4) is satisfied for every t. Therefore, this Φ is the desired stochastic Lagrangian flow.

Part 3: $W^{1,m}_{\rm loc}(\mathbb{R}^d)$-regularity of Φ. We prove a stability result, which is interesting in itself.

Lemma 4.9. *Let $m\geq 4$ be an even integer and assume that $(b_\varepsilon)_\varepsilon$ verifies Condition 2.14. If Φ^ε are the associated regular stochastic flows, then, for every $t\in[0,T]$, $(\Phi^\varepsilon_t)_\varepsilon$ converges to Φ_t weakly in $L^m(\Omega; W^{1,m}_{(1+|\cdot|)^{-d-1-m}}(\mathbb{R}^d))$.*

This weak convergence result yields in particular that, for every $t\in[0,T]$, Φ_t belongs to $L^m(\Omega;W^{1,m}_{\rm loc}(\mathbb{R}^d))$ and, if $m>d$, to $L^m(\Omega;C^{0,\alpha}_{\rm loc}(\mathbb{R}^d))$ (for $\alpha=1-d/m$) by Sobolev immersion. The proof is complete. □

Proof of Lemma 4.9. Step 1: Representation formula for fixed time. For every $u_0\in C^\infty_c(\mathbb{R}^d)$, for every $\varphi\in C^\infty_c(\mathbb{R}^d)$, for a.e. (t,x,ω), we have

$$\langle u(t),\varphi\rangle = \langle u_0, \varphi(\Phi_t)\rangle, \tag{4.6}$$

as a consequence of the analogous property for $\bar\Phi$ and of Remark 4.8. In particular, taking the H^{-1}-valued weakly-$*$ time continuous version for u (Remark 3.11), we get the above formula for every t, for every ω (in a full measure set independent of t). Moreover, by Theorem 3.16 extended to every time by weak-$*$ continuity, $\sup_t E[\|u(t,\cdot)\|^m_{L^m_{(1+|\cdot|)^\alpha}(\mathbb{R}^d)}]$ is finite for every real α (since u_0 is bounded compactly supported); therefore, calling again φ_n functions in $C^\infty_c(\mathbb{R}^d)$ with $\varphi_n(x)=x$ for $|x|\leq n$, we have that, for every t fixed: $\langle u_0,\Phi_t\rangle$ is in $L^m(\Omega)$ and

$$E\big[\,|\langle u(t), id-\varphi_n\rangle|^m\,\big] = E\big[\,|\langle u_0, \Phi_t - \varphi_n(\Phi_t)\rangle|^m\,\big] \to 0 \quad \text{as } n\to\infty.$$

Step 2: Approximation and conclusion. Fix $t\in[0,T]$ and $u_0\in C^\infty_c(\mathbb{R}^d)$. Note that, for every $\varphi\in C^\infty_c(\mathbb{R}^d)$, $\varphi(\Phi^\varepsilon_t)$ is the solution v^ε, at time 0, to the backward approximated stochastic transport equation, with final time t and final datum φ. The approximated duality formula (3.20) (for $c=0$ and with a change of variable to avoid the "tilde"), the approximation Condition 2.14 on $(b_\varepsilon)_\varepsilon$ and equation (4.6) then give

$$E\left[|\langle u_0, \varphi(\Phi_t) - \varphi(\Phi^\varepsilon_t)\rangle|^m\right] = E\left[|\langle u(t),\varphi\rangle - \langle u_0, v^\varepsilon\rangle|^m\right] \to 0 \quad \text{as } \varepsilon\to 0. \tag{4.7}$$

On the other hand, Corollary 3.7 gives that $v^\varepsilon(0) = \varphi(\Phi_t^\varepsilon)$ is bounded in $L^m(\Omega; W^{1,m}_{(1+|\cdot|)^{-(d+1+m)}}(\mathbb{R}^d))$, uniformly in ε. Since this space is reflexive (see Remark 2.9), $\varphi(\Phi_t^\varepsilon)$ converges weakly, as $\varepsilon \to 0$ and up to the choice of a subsequence, to an element Ψ_t^φ with

$$\|\Psi_t^\varphi\|_{L^m(\Omega;W^{1,m}_{(1+|\cdot|)^{-(d+1+m)}}(\mathbb{R}^d))} \le C\|\varphi\|_{W^{1,2m}_{(1+|\cdot|)^{-(d+1+2m)}}}.$$

for a constant C which is independent of t. In particular, for every $u_0 \in C_c^\infty(\mathbb{R}^d)$ and $F \in L^\infty(\Omega)$, we get

$$E\big[\langle u_0, \Psi_t^\varphi - \varphi(\Phi_t^\varepsilon)\rangle F\big] \to 0 \quad \text{as } \varepsilon \to 0.$$

Therefore, by (4.7) we find $\varphi(\Phi_t) = \Psi_t^\varphi$ for a.e. (x,ω). Now, taking $\varphi = \varphi_n$ (with bounded $W^{1,2m}_{(1+|\cdot|)^{-(d+1+2m)}}$ norm), we have

$$\|\varphi_n(\Phi_t)\|_{L^m(\Omega;W^{1,m}_{(1+|\cdot|)^{-(d+1+m)}}(\mathbb{R}^d))} \le C,$$

where the constant C is independent of n and t. As a consequence, $\varphi_n(\Phi_t)$ converges weakly in $L^m(\Omega; W^{1,m}_{(1+|\cdot|)^{-(d+1+m)}}(\mathbb{R}^d))$, as $n \to \infty$ and up to the choice of a subsequence. On the other hand, by Step 1 we know that $\langle u_0, \Phi_t - \varphi_n(\Phi_t)\rangle \to 0$ in $L^m(\Omega)$, as $n \to \infty$. So, by a similar argument to the one for $\varepsilon \to 0$, any weak limit of $\varphi_n(\Phi_t)$ has to be Φ_t, and hence

$$\|\Phi_t\|_{L^m(\Omega;W^{1,m}_{(1+|\cdot|)^{-(d+1+m)}}(\mathbb{R}^d))} \le C. \qquad \square$$

5 Towards classical pathwise uniqueness

So far we have investigated the problem of path-by-path uniqueness for the equations (sCE) and (sDE). In some sense, this is the strongest type of uniqueness we know. Indeed, we can come back heuristically to pathwise uniqueness for (sDE) in this way: given two processes X and Y which are solutions to (sDE) with the same initial datum, then, for a.e. ω, $X(\omega)$ and $Y(\omega)$ solve the sDE at fixed ω (more precisely, $\tilde{X}^\omega$ and $\tilde{Y}^\omega$ solve the random ODE (4.4)), so that, by path-by-path uniqueness, they must coincide. However, since we only deal with flows, we are not able to give a "classical" pathwise uniqueness result (among processes instead of flows), as a direct consequence of Theorem 4.5. Thus, we will now see how to modify the duality argument to get a more classical pathwise uniqueness, though still the initial datum cannot be a single point $x \in \mathbb{R}^d$, but it has to be a suitable diffused random variable.

5.1 The first result

The easiest consequence of Theorem 3.14 (applied to the continuity equation) is pathwise uniqueness among solutions with conditional laws (given the Brownian motion) in $L^m([0,T], L^m_{(1+|\cdot|)^{-\alpha}}(\mathbb{R}^d))$.

The relevant concept of solution and the result are shown below, but let us explain the idea. As already mentioned, we need the initial datum X_0 to be diffuse. We could take e.g. the probability space $(C([0,T];\mathbb{R}^d) \times B_R(y_0), Q \otimes \mathcal{L}^d)$ (with the suitable σ-algebra), with Q as Wiener measure, for some $R > 0$, $y_0 \in \mathbb{R}^d$, and $X_0(\gamma, x) = x$, $W_t(\gamma, x) = \gamma_t$; the filtration must be any filtration $(\mathcal{G}_t)_t$ (satisfying the standard assumptions) such that $\mathcal{G}_t$ contains $\sigma\{X_0, W_s | s \le t\}$. The solution $X = X(\gamma, x)$ should be thought of as a flow, for fixed $\gamma \in C([0,T];\mathbb{R}^d)$, solving the sDE at this fixed γ. Now we ask: among which class of processes path-by-path uniqueness applies, implying pathwise uniqueness? We have to require (again heuristically) that, for Q-a.e. Brownian trajectory $W = \gamma$, $(X_t(\gamma,\cdot))_\# \mathcal{L}^d$ is a diffuse measure. This is true in the case above, while for the general case (of a general probability space and general initial datum X_0), we must require that X_0 has a diffuse law and that "the law of X_t for fixed γ is diffuse" too. This law of X_t for fixed γ is the conditional law of X_t given the Brownian motion W; see e.g. [76, Chapter 1] for a reference on conditional law.

Definition 5.1. *Let $m \ge 1$, $\alpha \in \mathbb{R}$; let W be a Brownian motion (on a probability space $(\Omega, \mathcal{A}, P)$), let $(\mathcal{F}_t)_t$ be its natural completed filtration. An $\mathbb{R}^d$-valued process X on Ω is said to have* conditional (marginal) laws *(given the Brownian motion W) in $L^m([0,T], L^m_{(1+|\cdot|)^\alpha}(\mathbb{R}^d))$ if, for a.e. $t \in [0,T]$, the conditional law of X_t given $\mathcal{F}_t$ has a density (with respect to Lebesgue measure) $\rho(t,x,\omega)$ and, for a.e. $\omega \in \Omega$, $\rho(\cdot,\cdot,\omega)$ belongs to $L^m([0,T], L^m_{(1+|\cdot|)^\alpha}(\mathbb{R}^d))$.*

Theorem 5.2. *Let $m \geq 4$, $s \in \mathbb{R}$. Let W, $(\mathcal{F}_t)_t$ be as above and let X_0 be a random variable on Ω, independent of W, such that the law of X_0 has a density (with respect to the Lebesgue measure) in $L^m_{(1+|\cdot|)^{2s+d+1}}(\mathbb{R}^d)$. Assume Condition 2.2. Then, for every $\alpha \leq s$, strong existence and pathwise uniqueness hold for* (sDE) *with initial datum X_0, among solutions with conditional laws in $L^m([0,T], L^m_{(1+|\cdot|)^\alpha}(\mathbb{R}^d))$. More precisely, if $(\mathcal{G}_t)_t$ is an admissible filtration (satisfying the standard assumptions) on Ω (i.e. X_0 is $\mathcal{G}_0$-measurable and W is a Brownian motion with respect to $(\mathcal{G}_t)_t$), then there exists a unique $\mathcal{G}$-adapted process solving* (sDE)*, starting from X_0 and with conditional laws in $L^m([0,T], L^m_{(1+|\cdot|)^\alpha}(\mathbb{R}^d))$.*

We will not give all the details of the proof, also because the proof is similar to the one of the next Theorem 5.4.

Proof. The proof of uniqueness is similar to the one of the first part of Theorem 4.2. Suppose by contradiction that there exist two different solutions X and Y with the properties above. Then it is possible to find a time t_0, two disjoint Borel sets E and F in $\mathbb{R}^d$ and a measurable set Ω' in Ω with $P(\Omega') > 0$ such that $X_{t_0}(\omega)$ belongs to E and $Y_{t_0}(\omega)$ belongs to F for every ω in Ω'.

On $C([0,T];\mathbb{R}^d)$ we denote by Q the Wiener measure and by Γ the essential image of Ω' under the map W, i.e. $\Gamma := \{\gamma \in C([0,T];\mathbb{R}^d) \colon P(\Omega' | W = \gamma) > 0\}$ (this definition makes sense up to Q-negligible sets). Since Q is the image measure of P under W, we have $Q(\Gamma) > 0$. For every t, we define $\tilde{\mu}_t$ as the conditional law on $\mathbb{R}^d$ of $\tilde{X}_t$, restricted to Ω', given W, i.e., for every φ in $C_b(\mathbb{R}^d)$,

$$\langle \tilde{\mu}^\gamma_t, \varphi \rangle = E[\varphi(\tilde{X}_t) \mathbb{1}_{\Omega'} | W = \gamma], \quad \text{for } Q\text{-a.e. } \gamma \in \Gamma.$$

We analogously define $\tilde{\nu}_t$ for $\tilde{Y}$ instead of $\tilde{X}$. Then one can show that:

- for Q-a.e. $\gamma \in \Gamma$, $\tilde{\mu}^\gamma$ and $\tilde{\nu}^\gamma$ are weakly continuous (in time) solutions to the random CE (4.5) at γ fixed;
- $\tilde{\mu}$ and $\tilde{\nu}$ belong to the $\mathcal{L}_{\tilde{b}}$ class;
- $\tilde{\mu}$ and $\tilde{\nu}$ differ at time t_0.

So we have found two different $\mathcal{L}_{\tilde{b}^\gamma}$ solutions to the random CE at γ fixed, for a non-negligible set of γ. This is a contradiction, and thus, the proof of uniqueness is complete.

Strong existence is a consequence of the existence of a stochastic Lagrangian flows Φ solving (sDE). Indeed, defining $X_t(\omega) := \Phi^\omega_t(X_0(\omega))$, we observe the following facts:

- Since $\Phi(x)$ solves the sDE with initial datum x, for a.e. x, and X_0 is absolutely continuous (with respect to the Lebesgue measure), X verifies for a.e. ω

$$X_t = X_0 + \int_0^t b(s, X_s) ds + W_t.$$

- X is obviously $\mathcal{H}$-adapted, where $\mathcal{H}_t = \sigma(\{X_0, W_s | s \leq t\} \cup \mathcal{N})$ is the minimal admissible filtration ($\mathcal{N}$ are the P-null sets).
- Let u_0 be the density of the law of X_0 and u the solution in $L^\infty([0,T]; L^m(\Omega; L^m_{(1+|\cdot|)^s}(\mathbb{R}^d)))$ to the sCE as in Theorem 3.16, with initial datum u_0. Then the law of X_t has $u(t)$ as conditional density, given W. To prove this, notice that W and Φ are adapted to the Brownian (completed) filtration $\mathcal{F}$ and that X_0 is independent of $\mathcal{F}_T$, so, for any test function $\varphi \in C^\infty_c(\mathbb{R}^d)$ and any $\psi \in C_b(C([0,T];\mathbb{R}^d))$, we have

$$\begin{aligned} E[\varphi(X_t)\psi(W)] &= E[\varphi(\Phi_t(X_0))\psi(W)] \\ &= E\Big[\int_{\mathbb{R}^d} \varphi(\Phi_t(x)) u_0(x) \psi(W) dx\Big] = E\Big[\int_{\mathbb{R}^d} \varphi(x) u(t,x) \psi(W) dx\Big], \end{aligned}$$

  where in the second passage we used independence (precisely, in the form of Lemma 5.5 in the following paragraph) and the last passage is a consequence of $u(t) = (\Phi_t)_\# u_0$.

Thus, X is the desired solution and also existence is proved. □

5.2 The second result

The previous result was somehow limited, at least for uniqueness, by our hypothesis on conditional laws. In this paragraph we prove that actually pathwise uniqueness holds among processes whose marginal laws are diffuse (the precise hypothesis is stated below), with no need to control conditional laws.

To understand the relation with the previous Theorem 5.2, consider again the case discussed at the beginning of the previous paragraph and notice that, given a process X on $(C([0,T];\mathbb{R}^d) \times B_R(y_0), Q \otimes \mathcal{L}^d)$, the law ρ_t of X_t is the Q-average, on $C([0,T];\mathbb{R}^d)$, of the conditional laws ρ_t^γ of $X_t(\gamma, x)$, given the Brownian trajectory γ. So the fact that the law (that is, the mean of the conditional laws) is diffuse is a weaker condition than the hypothesis on Q-a.e. conditional law. Hence the class of processes whose marginal laws are diffuse is larger that the class used in Theorem 5.2, in particular the uniqueness result in the following Theorem 5.4 is morally stronger. Actually no implication holds between the two uniqueness results (for a technicality on the bounds on the densities, see the next definition), but still the idea is that uniqueness is stronger in Theorem 5.4.

Definition 5.3. *Let $m \geq 1$, $\alpha \in \mathbb{R}$. An $\mathbb{R}^d$-valued process X is said to have* (marginal) laws in $L^\infty([0,T], L^m_{(1+|\cdot|)^\alpha}(\mathbb{R}^d))$ *if, for a.e. $t \in [0,T]$, the law of X_t has a density* (*with respect to the Lebesgue measure*) *$\rho(t,x)$, which belongs to $L^\infty([0,T], L^m_{(1+|\cdot|)^\alpha}(\mathbb{R}^d))$.*

As previously mentioned, this class seems to be larger than that of Definition 5.1. Rigorously speaking, it is not: to deduce $\mu_t \in L^m([0,T] \times \mathbb{R}^d)$ from $\mu_t^\omega \in L^m([0,T] \times \mathbb{R}^d)$ for a.e. ω, we need the additional condition that $\int \|\mu_t^\omega\|_{L^m([0,T]\times\mathbb{R}^d)} P(d\omega)$ is finite.

Here is the main pathwise uniqueness result:

Theorem 5.4. *Let $m \geq 4$, $s \in \mathbb{R}$. Let W be a Brownian motion* (*on a probability space $(\Omega, \mathcal{A}, P)$*). *Let X_0 be a random variable on Ω, independent of W, such that the law of X_0 has a density* (*with respect to the Lebesgue measure*) *in $L^m_{(1+|\cdot|)^{2s+d+1}}(\mathbb{R}^d)$. Assume Condition 2.2. Then, for every $\alpha \leq s$, strong existence and pathwise uniqueness hold for* (sDE) *with initial datum X_0, among solutions with laws in $L^\infty([0,T], L^m_{(1+|\cdot|)^\alpha}(\mathbb{R}^d))$. More precisely, if $(\mathcal{G}_t)_t$ is an admissible filtration* (*satisfying the standard assumptions*) *on Ω* (*i.e. X_0 is $\mathcal{G}_0$-measurable and W is a Brownian motion with respect to $(\mathcal{G}_t)_t$*), *then there exists a unique $\mathcal{G}$-adapted process solving* (sDE), *starting from X_0 and with laws in $L^\infty([0,T], L^m_{(1+|\cdot|)^\alpha}(\mathbb{R}^d))$.*

Proof of uniqueness. First we give the idea of the proof. Let X, Y be two solutions to (sDE) which are adapted to an admissible filtration $(\mathcal{G}_t)_t$. Set $\mu_t := \delta_{X_t} - \delta_{Y_t}$; then μ is a random distribution which solves the sCE

$$\partial_t \mu + \operatorname{div}(b\mu) + \sum_{k=1}^d \partial_k \mu \circ \dot{W} = 0$$

in the sense of distributions, with $\mu_0 = 0$. We have to prove that $\mu \equiv 0$. We again want to use duality: if v solves the backward sTE

$$\partial_t v + b \cdot \nabla v + \sum_{k=1}^d \partial_k v \circ \dot{W} = 0$$

with final time t_f and final condition $v(t_f) = \varphi$ fixed, then formally it holds $\langle \mu_t, \varphi \rangle = \langle \mu_0, v_0 \rangle = 0$. But now we must be careful: expressions like

$$\langle \mu_s, b(s) \cdot \nabla v(s) \rangle, \tag{5.1}$$

which appear naturally in the rigorous proof of the duality formula, are no more under control: μ_s is only a measure, while $b(s)$ and $\nabla v(s)$ are not continuous (not even bounded). There are two key facts. The first one is where the integrability hypothesis plays a role: if we replace μ_s by its average $\rho_s = E[\mu_s] = (X_s)_\# P - (Y_s)_\# P$, we can estimate (5.1) since the density of ρ is in the correct integrability class for Hölder's inequality. However, taking the expectation, we have to deal with $E[\mu_s \nabla v(s)]$. Here enters the second key fact, namely that μ_s and $\nabla v(s)$ are independent, since μ_s is $\mathcal{G}_s$-measurable, while $v(t)$ (as backward solution) is adapted to the Brownian backward

(completed) filtration $\mathcal{F}_{s,t}$, which is independent of $\mathcal{G}_s$. Having this in mind, we come to the rigorous proof of the result.

Take $t_f \in [0,T]$ and $\varphi \in C_c^\infty(\mathbb{R}^d)$. Let b_ε be as in Condition 2.14, let v_ε be the solution to the approximated backward transport equation

$$\partial_t v_\varepsilon + b_\varepsilon \cdot \nabla v_\varepsilon + \sum_{k=1}^d \partial_k v_\varepsilon \circ \dot{W} = 0$$

with final time t_f and final datum $v_\varepsilon(t_f) = \varphi$. With the usual notation with tilde ($\tilde{v}_\varepsilon(s,x) = v_\varepsilon(s, x+\sigma W_s)$, $\tilde{X}_s = X_s - \sigma W_s$), the chain rule gives

$$\begin{aligned} dv_\varepsilon(t, X_t) = d\tilde{v}_\varepsilon(t, \tilde{X}_t) &= \tilde{b}(t, \tilde{X}_t) \cdot \nabla \tilde{v}_\varepsilon(t, \tilde{X}_t) dt - \tilde{b}_\varepsilon(t, \tilde{X}_t) \cdot \nabla \tilde{v}_\varepsilon(t, \tilde{X}_t) dt \\ &= [(b - b_\varepsilon) \cdot \nabla v_\varepsilon](t, X_t) dt \end{aligned}$$

and similarly for Y. Subtracting the expression for Y from that for X, we get

$$\varphi(X_{t_f}) - \varphi(Y_{t_f}) = \int_0^{t_f} [(b - b_\varepsilon) \cdot \nabla v_\varepsilon](s, X_s) ds - \int_0^{t_f} [(b - b_\varepsilon) \cdot \nabla v_\varepsilon](s, Y_s) ds.$$

We now claim that

$$\lim_{\varepsilon \to 0} \int_0^{t_f} E\big[|(b - b_\varepsilon) \cdot \nabla v_\varepsilon|(s, X_s)\big] ds = 0 \tag{5.2}$$

and similarly for Y. Assuming this, we obtain $\varphi(X_{t_f}) = \varphi(Y_{t_f})$ and then, by the arbitrariness of φ and t_f, also $X \equiv Y$.

For proving (5.2), we want to exploit the independence of $\nabla v_\varepsilon(s)$ and X_s, for fixed $s \in [0,T]$. To this end, we need the following elementary lemma:

Lemma 5.5. *Consider two measurable spaces $(F_1, \mathcal{F}_1)$, $(F_2, \mathcal{F}_2)$ and a probability measure P on $(F_2, \mathcal{F}_2)$. Let $f \colon F_1 \times F_2 \to \mathbb{R}$, $Z \colon F_2 \to F_1$ be two measurable functions and denote by ρ the law of Z on F_1. Suppose that there exists a σ-algebra $\mathcal{A} \subset \mathcal{F}_2$ such that f is $\mathcal{F}_1 \otimes \mathcal{A}$-measurable and Z is independent of $\mathcal{A}$. Assume also $\int_{F_1} \int_{F_2} |f(y,\omega)| P(d\omega) \rho(dy) < \infty$. Then it holds*

$$\int_{F_2} f(Z(\omega), \omega) P(d\omega) = \int_{F_1} \int_{F_2} f(y, \omega) P(d\omega) \rho(dy).$$

Proof. The lemma is clear for $f(y,\omega) = g(y)h(\omega)$, when g is $\mathcal{F}_1$-measurable and integrable (with respect to η), and h is $\mathcal{A}$-measurable and integrable (with respect to P). The general case is obtained by approximating f with sums of functions as above. □

Applying this lemma with $F_1 = \mathbb{R}^d$, $F_2 = \Omega$, $f = |(b - b_\varepsilon) \cdot \nabla v_\varepsilon|$ and $Z = X$ with law ρ, for fixed $s \in [0, t_f]$, and then integrating over $s \in [0, t_f]$, we obtain

$$\int_0^{t_f} E\big[|(b - b_\varepsilon) \cdot \nabla v_\varepsilon|(s, X_s)\big] ds = \int_0^{t_f} \langle E[|(b(s) - b_\varepsilon(s)) \cdot \nabla v_\varepsilon(s)|], \rho_s \rangle ds.$$

We would like to use Hölder's inequality to conclude with (5.2). Since the density of ρ belongs to $L^\infty([0,T]; L^m_{(1+|\cdot|)^{-\alpha}}(\mathbb{R}^d))$ by assumption, it is enough to prove that

$$\int_0^{t_f} \Big(\int_{\mathbb{R}^d} |b - b_\varepsilon|^{m'} |E[|\nabla v_\varepsilon|^{m'}](1 + |x|)^{\alpha m'/m} dx \Big)^{1/m'} ds < \infty. \tag{5.3}$$

The proof of (5.3) is almost the same of that of Lemma 3.13. The only change is the exponent $1/m'$ in the time integral. For this reason we need even less, namely it suffices that

$$\int_0^T \Big(\int_{\mathbb{R}^d} (1 + |x|)^{-\beta} |b - b_\varepsilon|^{\tilde{p}} dx \Big)^{1/\tilde{p}} ds < \infty$$

for some $\beta \geq 0$ and $\tilde{p} > m'$, which holds under Condition 2.2. The rest of the proof requires only obvious changes compared to the proof of Lemma 3.13. □

Proof of existence. As for Theorem 5.2, strong existence is an easy consequence of the existence of a stochastic Lagrangian flows Φ solving (sDE). Indeed, defining $X_t(\omega) \coloneqq \Phi_t^\omega(X_0(\omega))$, we observe the following facts.

- Since $\Phi(x)$ solves the sDE with initial datum x, for a.e. x, and X_0 is absolutely continuous (with respect to the Lebesgue measure), X verifies for a.e. ω

$$X_t = X_0 + \int_0^t b(s, X_s)ds + W_t.$$

- X is obviously $\mathcal{H}$-adapted, where $\mathcal{H}_t = \sigma(\{X_0, W_s | s \leq t\} \cup \mathcal{N})$ is the minimal admissible filtration ($\mathcal{N}$ are the P-null sets).

- Let u_0 be the density of the law of X_0 and u the solution in $L^\infty([0,T]; L^m(\Omega; L^m_{(1+|\cdot|)^s}(\mathbb{R}^d)))$ to the sCE as in Theorem 3.16, with initial datum u_0. Then the law of X_t has density given by $\mu_t = E[u(t)]$. Indeed Φ and X_0 are independent (which allows to use Lemma 5.5), so, for any test function $\varphi \in C_c^\infty(\mathbb{R}^d)$, we have

$$E[\varphi(X_t)] = E[\varphi(\Phi_t(X_0))] = E\Big[\int_{\mathbb{R}^d} \varphi(\Phi_t(x))u_0(x)dx\Big] = E\Big[\int_{\mathbb{R}^d} \varphi(x)u(t,x)dx\Big],$$

  where the last passage is a consequence of $u(t) = (\Phi_t)_\# u_0$. We further have for μ

$$\sup_{t\in[0,T]} \int_{\mathbb{R}^d} |\mu_t|^m (1+|x|)^{-\alpha} dx \leq \sup_{t\in[0,T]} \int_{\mathbb{R}^d} E[|u_t|^m](1+|x|)^{-\alpha} dx < \infty,$$

  so X has law in $L^\infty([0,T], L^m_{(1+|\cdot|)^s}(\mathbb{R}^d))$.

Thus, X is the desired solution and also existence is proved. □

6 Path-by-path results for sDE

6.1 Path-by-path uniqueness of individual trajectories

We next consider equation (sDE). Its integral formulation is

$$X_t(\omega) = x + \int_0^t b(s, X_s(\omega))ds + \sigma W_t(\omega)$$

and therefore we may give a path-by-path meaning to it. Assume for some constant $C > 0$ that $b\colon [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ is a measurable locally bounded function (defined for all (t,x), not only a.e.). As before, let us assume that W has continuous trajectories (everywhere). Given $\omega \in \Omega$, hence given the continuous function $t \mapsto W_t(\omega)$, consider all continuous functions $y\colon [0,T] \to \mathbb{R}^d$ which satisfy the identity

$$y(t) = x + \int_0^t b(s, y(s))ds + \sigma W_t(\omega)$$

and call $C(\omega, x)$ the set of all such functions. Denote by $Card(C(\omega, x))$ the cardinality of $C(\omega, x)$.

Remark 6.1. *If b is continuous with $|b(t,x)| \leq C(1+|x|)$ for all $(t,x) \in [0,T] \times \mathbb{R}^d$, then by classical deterministic arguments $C(\omega, x)$ is non empty.*

Definition 6.2. *We say that the sDE satisfies* path-by-path uniqueness *if*

$$P\big(Card(C(\omega, x)) \leq 1 \text{ for all } x \in \mathbb{R}^d\big) = 1,$$

namely if for a.e. $\omega \in \Omega$, $C(\omega, x)$ is at most a singleton for every x in $\mathbb{R}^d$.

To our knowledge, the only two results on path-by-path uniqueness are [25] and [15]. We present here a new strategy for this kind of results.

Let $y \in C(\omega, x)$ be a solution. Set

$$z^\omega(t) \coloneqq y(t) - \sigma W_t(\omega)$$

solution of

$$z^\omega(t) = x + \int_0^t b(s, z^\omega(s) + \sigma W_s(\omega))ds = x + \int_0^t \widetilde{b}^\omega(s, z^\omega(s))ds,$$

where, as usual, $\widetilde{b}^\omega(t,x) = b(t, x + \sigma W_t(\omega))$. Consider the time-dependent Dirac measure $\widetilde{\mu}_\omega(t) = \delta_{z^\omega(t)}$ on $\mathbb{R}^d$. For every $\varphi \in C_c^\infty(\mathbb{R}^d)$, we write $\langle \widetilde{\mu}^\omega(t), \varphi \rangle$ for $\int_{\mathbb{R}^d} \varphi d\widetilde{\mu}^\omega(t)$, which, in this particular case, is simply $\varphi(z^\omega(t))$.

Lemma 6.3. *For all $t \in [0,T]$ and $\varphi \in C^1([0,T]; C_c^\infty(\mathbb{R}^d))$, we have*

$$\langle \widetilde{\mu}^\omega(t), \varphi(t) \rangle = \langle \widetilde{\mu}^\omega(0), \varphi(0) \rangle + \int_0^t \left\langle \widetilde{\mu}^\omega(s), \widetilde{b}^\omega(s) \cdot \nabla\varphi(s) + \partial_t \varphi(s) \right\rangle ds.$$

Proof. We have to prove that

$$\varphi(t, z^\omega(t)) = \varphi(0, z^\omega(0)) + \int_0^t \big(\widetilde{b}^\omega(s, z^\omega(s)) \cdot \nabla\varphi(s, z^\omega(s)) + \partial_t\varphi(s, z^\omega(s))\big)ds,$$

which is true by ordinary calculus. □

We can now prove a central fact. For a bounded function f and a Borel set E, denote with $\|f\|_{0,E}$ the supremum of f over E; in general, this is not the essential supremum, unless f is continuous.

Theorem 6.4. *Let $(b_\varepsilon)_{\varepsilon \in (0,1)}$ be a family in $C_c^\infty([0,T] \times \mathbb{R}^d)$. Assume that, for every $t_f \in [0,T]$ and $v_0 \in C_c^\infty(\mathbb{R}^d)$, we have*

$$P - \lim_{\varepsilon \to 0} \int_0^{t_f} \|(b - b_\varepsilon) \cdot \nabla v_\varepsilon^\omega\|_{0,B_R} ds = 0 \tag{6.1}$$

for every positive R, and where, for every $\varepsilon \in (0,1)$, v_ε is the smooth solution of the backward sPDEs (3.17) *corresponding to b_ε and v_0, with $c_\varepsilon = 0$ (v_ε^ω denotes $v_\varepsilon(\cdot,\cdot,\omega)$ as before). Then path-by-path uniqueness holds for* (sDE).

Proof. Step 1: Identification of Ω_0, independently of x. By assumption (6.1), given $t_f \in [0,T]$ and $v_0 \in C_c^\infty(\mathbb{R}^d)$, there exist a full measure set $\Omega_{t_f, v_0} \subset \Omega$ and a sequence $\varepsilon_n \to 0$ such that

$$\widetilde{v}_{\varepsilon_n}^\omega \in C^1([0,T]; C_c^\infty(\mathbb{R}^d)), \text{ identity (3.19) holds for } \widetilde{v}_{\varepsilon_n}(\omega) \text{ (with } c = 0), \text{ for all } n \in \mathbb{N} \tag{6.2}$$

$$\lim_{n \to \infty} \int_0^{t_f} \big\|(b^\omega - b_{\varepsilon_n}^\omega) \cdot \nabla v_{\varepsilon_n}^\omega\big\|_{0,B_R} ds = 0.$$

for all $\omega \in \Omega_{t_f, v_0}$. Hence also

$$\lim_{n \to \infty} \int_0^{t_f} \big\|(\widetilde{b}^\omega - \widetilde{b}_{\varepsilon_n}^\omega) \cdot \nabla \widetilde{v}_{\varepsilon_n}^\omega\big\|_{0, B_R - \sigma W_t(\omega)} ds = 0 \tag{6.3}$$

for all $\omega \in \Omega_{t_f, v_0}$. Let $\mathcal{D} \subset C_c^\infty(\mathbb{R}^d)$ be a countable set which separates points, i.e. for all $a \neq b \in \mathbb{R}^d$, there is $v_0 \in \mathcal{D}$ with $v_0(a) \neq v_0(b)$. By a diagonal procedure, there exist a full measure set $\Omega_0 \subset \Omega$ and a sequence $\varepsilon_n \to 0$ such that properties (6.2) and (6.3) hold for all $t_f \in [0,T] \cap \mathbb{Q}$, $v_0 \in \mathcal{D}$, $n, R \in \mathbb{N}$ and $\omega \in \Omega_0$. Since, given $\omega \in \Omega_0$ and $R \in \mathbb{N}$, there exists $R'_\omega \in \mathbb{N}$ such that $B_R \subset B_{R'_\omega} - \sigma W_t(\omega)$ for all $t \in [0,T]$, we may replace (6.3) by

$$\lim_{n \to \infty} \int_0^{t_f} \big\|(\widetilde{b}^\omega - \widetilde{b}_{\varepsilon_n}^\omega) \cdot \nabla \widetilde{v}_{\varepsilon_n}^\omega\big\|_{0,B_R} ds = 0 \tag{6.4}$$

for all $t_f \in [0,T] \cap \mathbb{Q}$, $v_0 \in \mathcal{D}$, $R \in \mathbb{N}$ and $\omega \in \Omega_0$.

Step 2: $C(\omega, x)$ is a singleton for every $x \in \mathbb{R}^d$ and $\omega \in \Omega_0$, i.e. path-by-path uniqueness. Given $\omega \in \Omega_0$ and $y^{(i)} \in C(\omega, x)$, $i = 1, 2$, we define the (signed) measure

$$\widetilde{\rho}^{\omega}(t) \coloneqq \delta_{y^{(1)}(t) - \sigma W_t(\omega)} - \delta_{y^{(2)}(t) - \sigma W_t(\omega)},$$

which satisfies

$$\langle \widetilde{\rho}^{\omega}(t_f), \varphi(t_f) \rangle = \int_0^{t_f} \Big\langle \widetilde{\rho}^{\omega}(s), \widetilde{b}^{\omega}(s) \cdot \nabla \varphi(s) + \partial_t \varphi(s) \Big\rangle \, ds$$

for all $t_f \in [0,T]$ and $\varphi \in C^1([0,T]; C_c^{\infty}(\mathbb{R}^d))$, due to Lemma 6.3. In particular, this holds for $\varphi = \widetilde{v}^{\omega}_{\varepsilon_n}$ and thus, by (3.19), we get

$$\big\langle \widetilde{\rho}^{\omega}(t_f), v_0(\cdot + \sigma W_{t_f}(\omega)) \big\rangle = \int_0^{t_f} \Big\langle \widetilde{\rho}^{\omega}(s), (\widetilde{b}^{\omega}(s) - \widetilde{b}^{\omega}_{\varepsilon_n}(s)) \cdot \nabla \widetilde{v}^{\omega}_{\varepsilon_n}(s) \Big\rangle \, ds.$$

Then, if $R > 0$ is such that $|y^{(i)}(t)| \leq R$ for $t \in [0,T]$ and $i = 1, 2$, we find

$$\big| \big\langle \widetilde{\rho}^{\omega}(t_f), v_0(\cdot + \sigma W_{t_f}(\omega)) \big\rangle \big| \leq 2 \int_0^{t_f} \big\| (\widetilde{b}^{\omega}(s) - \widetilde{b}^{\omega}_{\varepsilon_n}(s)) \cdot \nabla \widetilde{v}^{\omega}_{\varepsilon_n}(s) \big\|_{0, B_R} ds,$$

and thus $\big\langle \widetilde{\rho}^{\omega}(t_f), v_0(\cdot + \sigma W_{t_f}(\omega)) \big\rangle = 0$ by (6.4). This is equivalent to $\big\langle \widetilde{\rho}^{\omega}(t_f, \cdot - \sigma W_{t_f}(\omega)), v_0 \big\rangle = 0$, which implies $\widetilde{\rho}^{\omega}(t_f, \cdot - \sigma W_{t_f}(\omega)) = 0$ since $v_0 \in \mathcal{D}$ was arbitrary and $\mathcal{D}$ separates points. Consequently, $y^{(1)}(t_f) = y^{(2)}(t_f)$ follows. This holds true for every $t_f \in [0,T] \cap \mathbb{Q}$, and since $t \mapsto y^{(1)}(t)$ is continuous, we get $y^{(1)}(t) = y^{(2)}(t)$ for every $t \in [0,T]$. This finishes the proof of the Theorem. □

Theorem 6.4 is, in a sense, our main result on path-by-path uniqueness, although assumption (6.1) is not explicit in terms of b. Roughly speaking, this condition is true when we have a uniform bound (in some probabilistic sense) for $\|\nabla \widetilde{v}^{\omega}_{\varepsilon_n}\|_{0, B_R}$. It introduces a new approach to the very difficult question of path-by-path uniqueness, which may be easily generalized to sDEs in infinite dimensions, for instance (this will be treated in separate works). A simple consequence is:

Corollary 6.5. *Let $(b_\varepsilon)_{\varepsilon \in (0,1)}$ be a family in $C_c^{\infty}([0,T] \times \mathbb{R}^d)$ which converges uniformly to b on compact sets $[0,T] \times B_R$, for every $R > 0$. Assume that, for every $t_f \in [0,T]$, $R > 0$ and $v_0 \in C_c^{\infty}(\mathbb{R}^d)$, we have*

$$\sup_{\varepsilon \in (0,1)} E \int_0^{t_f} \|\nabla v_\varepsilon\|_{0, B_R} \, ds < \infty \tag{6.5}$$

where v_ε is the smooth solution of the backward sPDEs (3.17) *corresponding to b_ε and v_0, with $c_\varepsilon = 0$. Then path-by-path uniqueness holds for* (sDE).

In Section 2.10 we have proved (reformulated for the backward sPDE) that, for every $t_f \in [0,T]$, $R > 0$, m an even positive integer and $v_0 \in C_c^{\infty}(\mathbb{R}^d)$,

$$\sup_{\varepsilon \in (0,1)} \sup_{[0,T]} E \Big[\, \|v_\varepsilon\|^m_{W^{2,m}(B_R)} \, \Big] < \infty.$$

Therefore, by Sobolev embedding, we obtain for $m > d$

$$\sup_{\varepsilon \in (0,1)} \sup_{[0,T]} E \Big[\, \|\nabla v_\varepsilon\|^m_{0, B_R} \, \Big] < \infty, \tag{6.6}$$

which implies condition (6.5) of Corollary 6.5. Hence, we have:

Corollary 6.6. *Under the conditions of Section* 2.10 (*Db of class LPS*) *we have path-by-path uniqueness for* (sDE).

Notice that the conditions of Section 2.10 with $m > d$ imply $b \in C^{\varepsilon}_{\mathrm{loc}}(\mathbb{R}^d, \mathbb{R}^d)$ for some $\varepsilon > 0$. Thus, in the case $m > d$, this result is included in Corollary 6.8 below and already in [25]. However, also the limit case $m = d \geq 3$ is included in our statement. Otherwise, we may take estimate (6.6) from [41] in the case

$$b \in L^\infty(0, T; C^\alpha_b(\mathbb{R}^d)); \tag{6.7}$$

(essential boundedness in time, with values in $C^\alpha_b(\mathbb{R}^d)$, is actually enough, since the measure solutions to the continuity equation are only space-valued). Precisely, the following result is proved in [41]. We give here an independent proof for the sake of completeness.

Lemma 6.7. *Let b satisfy* (6.7) *and take a family $b_\varepsilon \in C^\infty_c([0,T] \times \mathbb{R}^d)$ which converges uniformly to b on compact sets. Then the flows Φ^ε_t associated to* (sDE) *with coefficients b_ε satisfy for every $m \geq 1$*

$$\sup_{\varepsilon \in (0,1)} \sup_{[0,T]} E\Big[\|D\Phi^\varepsilon_t\|^m_{\infty, B_R} \Big] < \infty. \tag{6.8}$$

Proof. Step 1: Formula for $D\Phi^\varepsilon_t(x)$ via Itô–Tanaka trick. Let us introduce the vector field $U_\varepsilon(t,x)$, for $t \in [0,T]$, $x \in \mathbb{R}^d$, and with components $U^i_\varepsilon(t,x)$, for $i = 1, \dots, d$, satisfying the backward parabolic equation (where b^i_ε is the i-component of b_ε)

$$\partial_t U^i_\varepsilon + b_\varepsilon \cdot DU^i_\varepsilon + \frac{\sigma^2}{2} \Delta U^i_\varepsilon = -b^i_\varepsilon + \lambda U^i_\varepsilon, \qquad U^i_\varepsilon(T,x) = 0 \tag{6.9}$$

for some $\lambda > 0$. As explained in [41, Section 2] based on classical results of [53] (see also a partial probabilistic proof in [38]), this equation has a unique solution of class $U^i_\varepsilon \in C^1([0,T]; C^\alpha_b(\mathbb{R}^d)) \cap C([0,T]; C^{2,\alpha}_b(\mathbb{R}^d))$, and there is a uniform constant $C > 0$ such that

$$\sup_{\varepsilon \in (0,1)} \sup_{[0,T]} \|U_\varepsilon\|_{C^{2,\alpha}_b(\mathbb{R}^d)} \leq C. \tag{6.10}$$

Moreover, given any $\delta > 0$, there exists $\lambda > 0$ large enough such that

$$\|DU_\varepsilon\|_\infty \leq \delta. \tag{6.11}$$

Here and below we denote by $\|\cdot\|_\infty$ the L^∞ norm both in time and space. We may apply the Itô formula to $U^i_\varepsilon(t, \Phi^\varepsilon_t(x))$ and use (6.9) to get

$$U^i_\varepsilon(t, \Phi^\varepsilon_t(x)) = U^i_\varepsilon(0,x) + \int_0^t (-b^i + \lambda U^i_\varepsilon)(s, \Phi^\varepsilon_s(x)) ds + \sigma \int_0^t \nabla U^i_\varepsilon(s, \Phi^\varepsilon_s(x)) \cdot dW_s.$$

This allows us to rewrite the equation

$$\Phi^{\varepsilon,i}_t(x) = x^i + \int_0^t b^i_\varepsilon(s, \Phi^\varepsilon_s(x)) ds + \sigma W^i_t$$

in the form

$$\Phi^{\varepsilon,i}_t(x) = x^i + U^i_\varepsilon(0,x) - U^i_\varepsilon(t, \Phi^\varepsilon_t(x)) + \int_0^t \lambda U^i_\varepsilon(s, \Phi^\varepsilon_s(x)) ds + \sigma \int_0^t \nabla U^i_\varepsilon(s, \Phi^\varepsilon_s(x)) \cdot dW_s + \sigma W^i_t.$$

Since b_ε is smooth and compactly supported, we a priori know from [56] that Φ^ε_t is differentiable; hence we may use the differentiability properties of U_ε and the result of differentiation under stochastic integral of [56] to have

$$\begin{aligned} \partial_k \Phi^{\varepsilon,i}_t(x) &= \delta_{ik} + \partial_k U^i_\varepsilon(0,x) - \sum_{j=1}^d \partial_j U^i_\varepsilon(t, \Phi^\varepsilon_t(x)) \partial_k \Phi^{\varepsilon,j}_t(x) \\ &\quad + \int_0^t \lambda \sum_{j=1}^d \partial_j U^i_\varepsilon(s, \Phi^\varepsilon_s(x)) \partial_k \Phi^{\varepsilon,j}_s(x) ds + \sigma \int_0^t \sum_{j,l=1}^d \partial_l \partial_j U^i_\varepsilon(s, \Phi^\varepsilon_s(x)) \partial_k \Phi^{\varepsilon,j}_s(x) dW^l_s. \end{aligned} \tag{6.12}$$

Step 2: Uniform pointwise estimate for $D\Phi_t^\varepsilon(x)$. We first use the previous identity to estimate $E[|\partial_k \Phi_t^{\varepsilon,i}(x)|^r]$ uniformly in (t,x) and ε, for each $r>1$. Denoting by $C_r>0$ a generic constant depending only on r, we have

$$E[|\partial_k \Phi_t^\varepsilon(x)|^r] \le C_r + C_r\|DU_\varepsilon\|_\infty^r + C_r\|DU_\varepsilon\|_\infty^r E[|\partial_k\Phi_t^\varepsilon(x)|^r] \\ + \lambda^r C_r\|DU_\varepsilon\|_\infty^r \int_0^t E[|\partial_k\Phi_s^\varepsilon(x)|^r]ds + \sigma^r C_r\|D^2U_\varepsilon\|_\infty^r \int_0^t E[|\partial_k\Phi_s^\varepsilon(x)|^r]ds,$$

where we have used the Burkholder–Davis–Gundy inequality in the last term. By choosing λ so large that $C_r\,\|DU_\varepsilon\|_\infty^r \le 1/2$ (possible by (6.11), we find

$$\frac12 E[|\partial_k \Phi_t^\varepsilon(x)|^r] \le C_r + \lambda^r \int_0^t E[|\partial_k\Phi_s^\varepsilon(x)|^r]ds + \sigma^r C_r\|D^2U_\varepsilon\|_\infty^r \int_0^t E[|\partial_k\Phi_s^\varepsilon(x)|^r]ds.$$

Now it is sufficient to apply Gronwall's lemma and the uniform estimate (6.10) to conclude that

$$\sup_{\varepsilon\in(0,1)}\sup_{t\in[0,T]}\sup_{x\in\mathbb{R}^d} E[|D\Phi_t^\varepsilon(x)|^r] < \infty. \tag{6.13}$$

Step 3: Conclusion via Kolmogorov's regularity criterion. To get the supremum in x inside the expectation, let us try to apply the Kolmogorov regularity criterion. Given $x,y\in\mathbb{R}^d$, $r>1$, we derive from (6.12) (using suitable vector notations)

$$E[|\partial_k\Phi_t^\varepsilon(x) - \partial_k\Phi_t^\varepsilon(y)|^r] \le C_r(I_1 + I_{21} + I_{22} + \lambda^r I_{31} + \lambda^r I_{32} + \sigma^r I_{41} + \sigma^r I_{42})$$

with the following abbreviations

$$\begin{aligned}
I_1 &= |\partial_k U_\varepsilon(0,x) - \partial_k U_\varepsilon(0,y)|^r \\
I_{21} &= E[|DU_\varepsilon(t,\Phi_{\varepsilon,t}(x))|^r |\partial_k\Phi_t^\varepsilon(x) - \partial_k\Phi_t^\varepsilon(y)|^r] \\
I_{22} &= E[|DU_\varepsilon(t,\Phi_{\varepsilon,t}(x)) - DU_\varepsilon(t,\Phi_{\varepsilon,t}(y))|^r |\partial_k\Phi_t^\varepsilon(y)|^r] \\
I_{31} &= \int_0^t E[|DU_\varepsilon(s,\Phi_t^\varepsilon(x))|^r |\partial_k\Phi_s^\varepsilon(x) - \partial_k\Phi_s^\varepsilon(y)|^r]ds \\
I_{32} &= \int_0^t E[|DU_\varepsilon(s,\Phi_s^\varepsilon(x)) - DU_\varepsilon(s,\Phi_s^\varepsilon(y))|^r |\partial_k\Phi_s^\varepsilon(y)|^r]ds \\
I_{41} &= \int_0^t E[|D^2U_\varepsilon(s,\Phi_t^\varepsilon(x))|^r |\partial_k\Phi_s^\varepsilon(x) - \partial_k\Phi_s^\varepsilon(y)|^r]ds \\
I_{42} &= \int_0^t E\big[|D^2U_\varepsilon(s,\Phi_s^\varepsilon(x)) - D^2U_\varepsilon(s,\Phi_s^\varepsilon(y))|^r |\partial_k\Phi_s^\varepsilon(y)|^r\big]ds.
\end{aligned}$$

For the last term we have used again the Burkholder–Davis–Gundy inequality. Let us denote by $C_U>0$ (resp. $C_{r,\Phi}>0$) a constant independent of $\varepsilon\in(0,1)$, based on the uniform estimate (6.10) (resp. on (6.13)) and let us write $\delta>0$ for the constant in (6.11). We have

$$\begin{aligned}
I_1 &\le \|DU_\varepsilon\|_\infty^r |x-y|^r \le C_U^r |x-y|^r \\
I_{21} &\le \|DU_\varepsilon\|_\infty^r E[|\partial_k\Phi_t^\varepsilon(x) - \partial_k\Phi_t^\varepsilon(y)|^r] \le \delta^r E[|\partial_k\Phi_t^\varepsilon(x) - \partial_k\Phi_t^\varepsilon(y)|^r] \\
I_{22} &\le \|D^2U_\varepsilon\|_\infty^r E\Big[\int_0^1 |D\Phi_t^\varepsilon(\theta x + (1-\theta)y)|^r d\theta |\partial_k\Phi_t^\varepsilon(y)|^r\Big]|x-y|^r \\
&\le \|D^2U_\varepsilon\|_\infty^r E\Big[|\partial_k\Phi_t^\varepsilon(y)|^{2r}\Big]^{1/2}\Big(\int_0^1 E\big[|D\Phi_t^\varepsilon(ux+(1-u)y)|^{2r}\big]du\Big)^{1/2}|x-y|^r \\
&\le C_U^r C_{2r,\Phi}^{1/2} C_{2r,\Phi}^{1/2}|x-y|^r = C_U^r C_{2r,\Phi}|x-y|^r.
\end{aligned}$$

Similarly, we get

$$I_{31} \le C_U^r \int_0^t E[|\partial_k\Phi_s^\varepsilon(x) - \partial_k\Phi_s^\varepsilon(y)|^r]ds$$

$$I_{32} \leq TC_U^r C_{2r,\Phi}|x-y|^r$$

and finally

$$\begin{aligned} I_{41} &\leq C_U^r \int_0^t E\big[|\partial_k \Phi_s^\varepsilon(x) - \partial_k \Phi_s^\varepsilon(y)|^r\big] ds \\ I_{42} &\leq \sup_{[0,T]} \|D^2 U_\varepsilon\|_{C^\alpha}^r \int_0^t E\Big[\int_0^1 |D\Phi_s^\varepsilon(ux+(1-u)y)|^{\alpha r} du |\partial_k \Phi_s^\varepsilon(y)|^r\Big] ds |x-y|^{\alpha r} \\ &\leq TC_U^r C_{2\alpha r,\Phi}^{1/2} C_{2r,\Phi}^{1/2} |x-y|^{\alpha r}. \end{aligned}$$

Taking δ sufficiently small (and thus λ large enough), from Gronwall's lemma we deduce

$$E\big[|\partial_k \Phi_t^\varepsilon(x) - \partial_k \Phi_t^\varepsilon(y)|^r\big] \leq C_r |x-y|^{\alpha r}.$$

Since $r > 1$ is arbitrary, we may apply Kolmogorov's regularity criterion (see for instance the quantitative version of [56] for the bound on the moments of supremum norm in x) and entail (6.8), which finishes the proof of the lemma. □

It follows:

Corollary 6.8. *Under condition* (6.7) *we have path-by-path uniqueness for* (sDE).

Proof. In view of the formula $v_\varepsilon(t,x) = v_0(\Phi_t^\varepsilon(x))$ and (6.8), the estimate (6.6) holds, which in turn implies (6.1). Hence, the path-by-path uniqueness follows immediately from Theorem 6.4. □

7 Examples and counterexamples

In this section we present some examples of drifts under LPS conditions, which exhibit regularization by noise phenomena (i.e. the ODE is ill-posed, while the sDE is well-posed), and an example outside of the LPS conditions where our results do not hold.

7.1 Examples of regularization by noise

Example 7.1. *Given a real number α, we consider on $\mathbb{R}^d$ the autonomous vector field*

$$b(x) := \mathbb{1}_{0<|x|\leq 1}|x|^\alpha \hat{x} + \mathbb{1}_{|x|>1} x,$$

where $\hat{x} = x/|x|$ for $x \neq 0$ (and $\hat{0} = 0$). The vector field b is Lipschitz continuous if and only if $\alpha \geq 1$ and it satisfies the LPS conditions if and only if $\alpha > -1$. In the deterministic case, if $-1 < \alpha < 1$, we see that:

- *if $x_0 \neq 0$, then there exists a unique solution Y to the ODE $dX_t/dt = b(t, X_t)$ (that is* (sDE) *with $\sigma = 0$) starting from x_0, namely*

  $$Y(t) = (|x_0|^{1-\alpha} + (1-\alpha)t)^{1/(1-\alpha)} \hat{x}_0 \mathbb{1}_{t\leq t_1} + e^{t-t_1}\hat{x}_0 \mathbb{1}_{t>t_1},$$

  where t_1 is the first time that $|Y| = 1$ ($t_1 = 0$ if $|x_0| > 1$);

- *if $x_0 = 0$, then there is an infinite number of solutions to the ODE starting from 0, namely any*

  $$Y(t) = \mathbb{1}_{t>t_a}((1-\alpha)(t-t_0))^{1/(1-\alpha)} \hat{x}_a \mathbb{1}_{t\leq t_1} + e^{t-t_1}\hat{x}_a \mathbb{1}_{t>t_1}$$

  for some t_a in $[0,\infty]$ (for $t_a = \infty$, we find the null solution) and some x_a in the unit sphere $\mathbb{S}^{d-1}$ (and with t_1 as before).

Consequently, if $-1 < \alpha < 1$, there cannot exist a continuous flow solving the ODE: continuity fails in $x_0 = 0$. This also implies that non-uniqueness appears for the transport equation with discontinuous (at 0) initial datum.

On the contrary, for $\alpha > -1$ the drift is in the LPS class, hence all our results apply to this example and show regularization by noise in various ways. In particular, the discontinuity of the flow in the origin is removed ω-wise; moreover, for $\alpha > 0$ (when b is Hölder continuous), we have proved path-by-path uniqueness starting from $x_0 = 0$. Let us also remind that pathwise uniqueness holds, starting from 0, by [54].

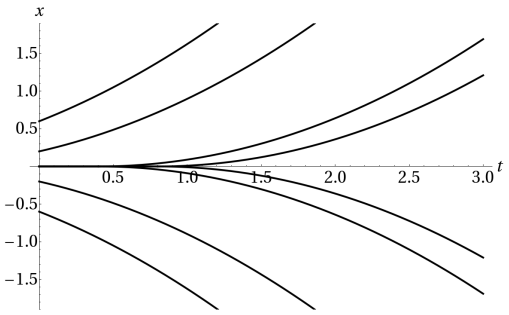

Figure 1: Non-uniqueness of trajectories

In this particular example it is possible to get an intuitive idea, "by hands", of what happens. If one consider the ODE without noise starting from 0, any solution Y grows near 0 no faster than $t^{1/(1-\alpha)}$; on the contrary, the Brownian motion W near 0 grows as $t^{1/2}$ (up to a logarithmic correction, which does not affect the intuition). Heuristically, we could say that the "speed" of Y near 0 caused by the drift is like $t^{\alpha/(1-\alpha)}$, while the one caused by W is like $t^{-1/2}$. So what we expect to happen is that the Brownian motion moves the particle immediately away from 0, faster than the action of the drift, and this prevents the formation of non-uniqueness or singularities. At least in the one-dimensional case, this can be seen also through speed measure and scale function, see [13].

Notice that if $\alpha < -1$ the opposite phenomenon appears (Y is faster than W), so that we expect ill-posedness. This is also an argument for the optimality of the LPS conditions (even if, in this case, we do not look at critical cases in LPS hypotheses), see also Example 7.4.

Example 7.2. *We consider a similar autonomous vector field as in the previous example, but change sign:*

$$b(x) := -\mathbb{1}_{0<|x|\leq 1}|x|^{\alpha}\hat{x} - \mathbb{1}_{|x|>1}x.$$

In this case, we see that, for every initial x_0, there exists a unique solution Y to the ODE, which reaches 0 in finite time and then stays in 0. Thus, concentration happens in 0, so that there does not exist a Lagrangian flow (the image measure of the flow at time t can have a Dirac delta in 0). Moreover the solution to the continuity equation concentrates in 0. Again our results apply when $\alpha > -1$, so these concentration phenomena disappear in the stochastic case.

Example 7.3. *Take $d = 2$ for simplicity. The following field is a combination of the previous two examples:*

$$b(x) = \mathbb{1}_{x\in A}\big[\mathbb{1}_{0<|x|\leq 1}|x|^{\alpha}\hat{x} + \mathbb{1}_{|x|>1}x\big] + \mathbb{1}_{x\in A^c}\big[-\mathbb{1}_{0<|x|\leq 1}|x|^{\alpha}\hat{x} - \mathbb{1}_{|x|>1}x\big],$$

where $A = \{x \in \mathbb{R}^2 \colon x_1 > 0 \text{ or } (x_1 = 0, x_2 > 0)\}$. It is easy to see that, for $\alpha < 1$, in the deterministic case one can construct flows with discontinuity, concentration of the mass in 0 or both; in particular, non-uniqueness holds. Again, for $\alpha > -1$, well-posedness (as in Theorem 4.5) *is restored.*

7.2 A counterexample in the supercritical case

Example 7.4. *Finally let us show that outside the LPS class there are equations and diffuse initial conditions without any solution; in particular the statement of Theorem* 5.4 *does not hold in this case. We now consider equation* (sDE) *on $\mathbb{R}^d$, with $\sigma = 1$ and drift b defined as*

$$b(x) := -\beta|x|^{-2}x\mathbb{1}_{x\neq 0},$$

with $\beta > d/2$. Notice that, for $d \geq 2$, this drift is just outside the LPS class (in the sense that $|x|^{\alpha-1}x$ belongs to that class for any $\alpha > -1$). For this particular sDE, we have: for some $T > 0$ and $M > 0$, if X_0 is a random variable, independent of W and uniformly distributed on $[-M, M]$, then there does not exist a weak solution, starting from X_0.

Proof. Step 1: (sDE) *does not have a weak solution for* $X_0 = 0$ (*for any* $T > 0$). Notice that this does not prevent from extending Theorem 5.4 to this case (because the initial datum is concentrated on 0), but it is a first step. The method is taken from [17].

Assume, by contradiction, that a weak solution on $[0, T]$ exists, i.e. there is a filtered probability space $(\Omega, \mathcal{A}, \mathcal{G}_t, P)$ (satisfying standard assumptions), a Brownian motion W in $\mathbb{R}^d$ with respect to $(\mathcal{G}_t)_t$, an $(\mathcal{G}_t)_t$-adapted continuous process $(X_t)_{t\geq 0}$ in $\mathbb{R}^d$, such that $\int_0^T |b(X_t)|dt < \infty$ and, a.s.,

$$X_t = \int_0^t b(X_s)ds + W_t.$$

Hence, X is a continuous semimartingale, with quadratic covariation $\langle X^i, X^j \rangle_t = \delta_{ij} t$. By the Itô formula, we have

$$\begin{aligned} d\,|X_t|^2 &= -2\beta \mathbb{1}_{X_t \neq 0} dt + 2X_t \cdot dW_t + d \cdot dt \\ &= (d\mathbb{1}_{X_t=0} - (2\beta - d)\mathbb{1}_{X_t \neq 0})dt + 2X_t \cdot dW_t. \end{aligned}$$

We now claim that

$$\int_0^T \mathbb{1}_{X_s=0} ds = 0 \tag{7.1}$$

holds with probability one. This implies

$$|X_t|^2 = -(2\beta - d)\int_0^t \mathbb{1}_{X_s \neq 0} ds + \int_0^t 2X_s \cdot dW_s.$$

Therefore, $|X_t|^2$ is a positive local supermartingale, vanishing at $t = 0$. This implies $|X_t|^2 \equiv 0$, hence $X_t \equiv 0$, which contradicts the fact that $\langle X^i, X^j \rangle_t = \delta_{ij} t$.

It remains to prove the claim (7.1). Consider the random set $\{t \in [0, T] \colon X_t = 0\}$. Since it is a subset of $A_1 = \{t \in [0, T] \colon X_t^1 = 0\}$, it is sufficient to prove that A_1 is of Lebesgue measure zero, P-a.s. and this is equivalent to $P\big(\int_0^T \mathbb{1}_{X_s^i=0} ds = 0\big) = 1$. Since X is a continuous semimartingale with quadratic covariation $\langle X^i, X^j \rangle_t = \delta_{ij} t$, also X^1 is a continuous semimartingale, with quadratic covariation $\langle X^1, X^1 \rangle_t = t$. Hence, by the occupation times formula (see [72, Chapter VI, Corollary 1.6])

$$\int_0^T \mathbb{1}_{X_s^1=0} ds = \int_{\mathbb{R}} \mathbb{1}_{a=0} L_T^a(X^1) da$$

where $L_T^a(X^1)$ is the local time at a on $[0, T]$ of the process X^1. Hence, a.s., $\int_0^T \mathbb{1}_{X_s^1=0} ds = 0$.

Step 2: (sDE) *does not have a weak solution starting from* X_0 *uniformly distributed on* $[-M, M]$ (*for some* $T > 0$ *and* $M > 0$). Again, we suppose by contradiction that there exists such a solution X (associated with some filtration $(\mathcal{G}_t)_t$), on a probability space $(\Omega, \mathcal{A}, P)$. Let τ be the first time when X hits 0 (it is a stopping time with respect to $(\mathcal{G}_t)_t$), with $\tau = \infty$ when X does not hit 0. We now claim that

$$P(\tau < \infty) > 0. \tag{7.2}$$

Assuming this, we can construct a new process Y, which is a weak solution to (sDE), starting from $Y_0 = 0$. This is in contradiction with Step 1. The process Y is built as follows. Take $\tilde{\Omega} = \{\tau < \infty\}$, $\tilde{\mathcal{A}} = \{A \cap \tilde{\Omega} \colon A \in \mathcal{A}\}$, $Q = P(\tilde{\Omega})^{-1} P|_{\tilde{\mathcal{A}}}$, then define $Y_t := X_{t+\tau}$, $\tilde{W}_t := W_{t+\tau} - W_\tau$ on $\tilde{\Omega}$ and $\mathcal{H}_t = \sigma(\{\tilde{W}_s, Y_s | s \leq t\} \cup \tilde{\mathcal{N}})$ (σ-algebra on $\tilde{\Omega}$), where $\tilde{\mathcal{N}}$ is the set of Q-null sets of $\tilde{\Omega}$. We observe the following facts:

- $\tilde{W}$ is a natural Brownian motion on the space $(\tilde{\Omega}, \mathcal{A}, Q)$, i.e., for every positive integer n, for every $0 < t_1 < \ldots < t_n$ and for every $f_1, \ldots, f_n$ in $C_b(\mathbb{R}^d)$, there holds

$$E\Big[\mathbb{1}_{\tilde{\Omega}} \prod_{j=1}^n f_j(W(t_j + \tau) - W(t_{j-1} + \tau))\Big] = P(\tilde{\Omega}) \prod_{j=1}^n \int_{\mathbb{R}^d} f_j d\mathcal{N}(0, (t_j - t_{j-1})I), \tag{7.3}$$

  where $\mathcal{N}(m, A)$ is the Gaussian law of mean m and covariance matrix A. This can be verified, for a general $\mathcal{G}$-stopping time, with a standard argument: first one proves (7.3)

when τ is a stopping time with discrete range in $[0,\infty]$, then, for the general case, one uses an approximation of τ with stopping times τ_k with discrete range such that $\tau_k \downarrow \tau$ (as $k \to \infty$) and $\{\tau = \infty\} = \{\tau_k = \infty\}$ for every k.

- $\tilde{W}$ is a Brownian motion with respect to the filtration $\mathcal{H}$, i.e., for every $0 = t_0 < t_1 < \ldots < t_n \le s < t$ and for every $f, g_1, \ldots, g_n$ in $C_b(\mathbb{R}^d)$, there holds

$$E\Big[\mathbb{1}_{\tilde{\Omega}} f(W(t+\tau) - W(s+\tau)) \prod_{j=0}^{n} g_j(X(t_j+\tau))\Big] = \int_{\mathbb{R}^d} f d\mathcal{N}(0,(t-s)I) E\Big[\prod_{j=0}^{n} g_j(X(t_j+\tau))\Big].$$

  Again this can be shown by approximation (with stopping times with discrete range).

- Y is a weak solution to (sDE), starting from $Y_0 = 0$. This follows immediately from

$$X_{s'} = X_s + \int_s^{s'} b(X_r)dr + W_{s'} - W_s,$$

  setting $s' = t + \tau$ and $s = \tau$.

It remains to prove claim (7.2). We suppose by contradiction that $\tau = \infty$ a.s.; this implies that, for every $t \in [0,T]$, we have $P(X_t \neq 0) = 1$. Then, computing $E[|X|^2]$ by the Itô formula, we get

$$\frac{d}{dt} E[|X_t|^2] = -2\beta P(X_t \neq 0) + d = -2\beta + d < 0,$$

hence, there exists a time $t_0 > 0$ such that $E[|X_{t_0}|^2] < 0$, which is a contradiction. This completes the proof. □

Remark 7.5. *The restriction on β is due to the first step. The claim (7.2) holds in fact for the more general case $\beta > (d-2)/2$, which can be achieved by an alternative approach.*

Sketch of proof. The idea is that the symmetry properties of the drift allow to reduce the solution X to the sDE to a one-dimensional Bessel process, for which the probability of hitting 0 is known. Fix $R > 0$ and, for any $\varepsilon > 0$, $x \in B_R \setminus \bar{B}_\varepsilon$, denote by $\tau_\varepsilon(x)$ the exit time from the annulus $B_R \setminus \bar{B}_\varepsilon$ of the solution $Z(x)$ to (sDE), starting from x (Z exists up to $\tau_\varepsilon(x)$ since the drift is regular in the annulus). If we prove that, for every $x \neq 0$, there exist $T > 0$, $\delta > 0$ and R large enough such that, for every $\varepsilon > 0$,

$$P\big(\tau_\varepsilon(x) < T, Z(x,\tau_\varepsilon(x)) = \varepsilon\big) > \delta, \tag{7.4}$$

we have shown the claim. In order to prove (7.4) we notice that $P(\tau_\varepsilon(x) < T, Z(x,\tau_\varepsilon(x)) = \varepsilon) = u(0,x)$, where u solves the backward parabolic PDE, on $B_R \setminus \bar{B}_\varepsilon$,

$$\partial_t u + b \cdot \nabla u + \frac{1}{2}\Delta u = 0,$$

with final and boundary conditions

$$u(T,\cdot) \equiv 0 \text{ in } B_R \setminus \bar{B}_\varepsilon, \quad u(t,\cdot) \equiv 1 \text{ on } \partial B_\varepsilon \text{ and } u(t,\cdot) \equiv 1 \text{ on } \partial B_R, \text{ for all } t \in [0,T].$$

By the symmetry properties of the drift b, the solution u is given by $u(t,x) = v(t,|x|)$, where $v\colon [0,T] \times (\varepsilon, R) \to \mathbb{R}$ solves the PDE

$$\partial_t v + \bar{b} \cdot \nabla v + \frac{1}{2}\Delta v = 0$$

for $\bar{b}(r) = (-\beta + (d-1)/2)|r|^{-1}\mathbb{1}_{r\neq 0}$, with final and boundary conditions

$$v(T,\cdot) \equiv 0, \quad v(\cdot,\varepsilon) \equiv 1, \quad v(\cdot,R) \equiv 0.$$

Then

$$P\big(\tau_\varepsilon(x) < T, Z(x,\tau_\varepsilon(x)) = \varepsilon\big) = v(0,|x|) = P\big(\sigma_\varepsilon(|x|) < T, \xi(|x|,\sigma_\varepsilon(|x|)) = \varepsilon\big), \tag{7.5}$$

where $\xi = \xi(r)$ is the one-dimensional process solution to the sDE

$$d\xi = \Big(-\beta + \frac{d-1}{2}\Big)|\xi|^{-1}\mathbb{1}_{\xi\neq 0}dt + dB_t$$

(B is a one-dimensional Brownian motion), with initial condition $\xi_0(r) = r$, and $\sigma_\varepsilon(r)$ is the exit time of $\xi(r)$ from the interval (ε, R). Now standard tools of one-dimensional diffusion processes (speed measure and scale function, see [13, Chapter 16]) allow to deduce that, since $\beta > (d-2)/2$, for every $r > 0$, $\xi(r)$ hits 0 in finite time with positive probability. This implies that, for every $r > 0$, there exist $T > 0$, $\delta > 0$ and R large enough such that, for every $\varepsilon > 0$, $P(\sigma_\varepsilon(r) < T, \xi(r, \sigma_\varepsilon(r)) = \varepsilon) > \delta$. In view of (7.5) we have established (7.4) and the sketch of proof of the final Remark is complete. □

References

[1] M. Aizenman, *On vector fields as generators of flows: a counterexample to Nelson's conjecture*, Ann. Math. **107** (1978), no. 2, 287–296.

[2] L. Ambrosio, *Transport equation and Cauchy problem for BV vector fields*, Invent. Math. **158** (2004), no. 2, 227–260.

[3] L. Ambrosio and G. Crippa, *Existence, uniqueness, stability and differentiability properties of the flow associated to weakly differentiable vector fields, Transport equations and multi-D hyperbolic conservation laws*, Lect. Notes Unione Mat. Ital., vol. **5**, Springer, Berlin, 2008, pp. 3–57.

[4] S. Attanasio and F. Flandoli, *Renormalized solutions for stochastic transport equations and the regularization by bilinear multiplication noise*, Comm. Partial Differential Equations **36** (2011), no. 8, 1455–1474.

[5] H. Bahouri, J.-Y. Chemin, and R. Danchin, *Fourier Analysis and Nonlinear Partial Differential Equations*, Springer, 2011.

[6] V. Bally, I. Gyöngy, and É. Pardoux, *White noise driven parabolic SPDEs with measurable drift*, J. Funct. Anal. **120** (1994), no. 2, 484–510.

[7] V. Barbu, M. Röckner, and D. Zhang, *Stochastic nonlinear Schrödinger equations: no blow-up in the non-conservative case*, J. Differential Equations **263** (2017), no. 11, 7919–7940.

[8] R.F. Bass, *Stochastic processes*, Cambridge Series in Statistical and Probabilistic Mathematics, vol. **33**, Cambridge University Press, Cambridge, 2011.

[9] R.F. Bass and Z.-Q. Chen, *Brownian motion with singular drift*, The Annals of Probab. **31** (2003), n. 2, 791–817.

[10] H. Beirão Da Veiga, *Remarks on the smoothness of the $L^\infty(0,T;L^3)$ solutions of the 3-D Navier–Stokes equations*, Portugal. Math. **54** (1997), no. 4, 381–391.

[11] L.A. Bianchi, *Uniqueness for an inviscid stochastic dyadic model on a tree*, Electron. Commun. Probab. **18** (2013), no. 8, 1–12.

[12] V. Bogachev, G. Da Prato, and M. Röckner, *Uniqueness for solutions of Fokker-Planck equations on infinite dimensional spaces*, Comm. Partial Differential Equations **36** (2011), no. 6, 925–939.

[13] L. Breiman, *Probability*, Corrected reprint of the 1968 original, Classics in Applied Mathematics, vol. **7**, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1992.

[14] G. Cannizzaro and K. Chouk, *Multidimensional SDEs with singular drift and universal construction of the polymer measure with white noise potential*, Ann. Probab. **46** (2018), no. 3, 1710–1763.

[15] R. Catellier and M. Gubinelli, *Averaging along irregular curves and regularisation of ODEs*, Stochastic Process. Appl. **126** (2016), no. 8, 2323–2366.

[16] R. Catellier, *Rough linear transport equation with an irregular drift*, Stoch. Partial Differ. Equ. Anal. Comput. **4** (2016), no. 3, 477–534.

[17] A.S. Cherny and H.-J. Engelbert, *Singular stochastic differential equations*, LNM 1858. Springer-Verlag, Berlin, 2005.

[18] M. Chaves, K. Gawędzki, P. Horvai, A. Kupiainen, and M. Vergassola, *Lagrangian dispersion in Gaussian self-similar velocity ensembles*, J. Statist. Phys. **113** (2003), no. 5-6, 643–692.

[19] Z.-Q. Chen and Z. Zhao, *Diffusion processes and second order elliptic operators with singular coefficients for lower order terms*, Math. Ann. **302** (1995), 323–357.

[20] K. Chouk and B. Gess, *Path-by-path regularization by noise for scalar conservation laws*, arXiv:1708.00823 (2017).

[21] K. Chouk and M. Gubinelli, *Nonlinear PDEs with modulated dispersion I: Nonlinear Schrödinger equations*, Comm. Partial Differential Equations **40** (2015), no. 11, 2047–2081.

[22] K. Chouk and M. Gubinelli, *Nonlinear PDEs with modulated dispersion II: Korteweg–de Vries equation*, arXiv:1406.7675 (2014).

[23] G. Crippa and C. De Lellis, *Estimates and regularity results for the DiPerna–Lions flow*, J. Reine Angew. Math. **616** (2008), 15–46.

[24] R. Courant and D. Hilbert, *Methods of Mathematical Physics*, Interscience Publishers, Inc., New York, N.Y., 1953.

[25] A. M. Davie, *Uniqueness of solutions of stochastic differential equations*, Int. Math. Res. Not. IMRN (2007), no. 24, Art. ID rnm124, 26.

[26] A. Debussche and Y. Tsutsumi, *1D quintic nonlinear Schrödinger equation with white noise dispersion*, J. Math. Pures Appl. (9) **96** (2011), no. 4, 363–376.

[27] F. Delarue and R. Diel, *Rough paths and 1d SDE with a time dependent distributional drift: application to polymers*, Probab. Theory Related Fields **165** (2016), no. 1-2, 1-63.

[28] F. Delarue, F. Flandoli, and D. Vincenzi, *Noise prevents collaps of Vlasov-Poisson point charges*, Comm. Pure Appl. Math. **67** (2014), no. 10, 1700–1736.

[29] R.J. DiPerna and P.-L. Lions, *Ordinary differential equations, transport theory and Sobolev spaces*, Invent. Math. **98** (1989), no. 3, 511–547.

[30] R. Duboscq and A. Réveillac, *Stochastic regularization effects of semi-martingales on random functions*, J. Math. Pures Appl. (9) **106** (2016), no. 6, 1141–1173.

[31] R. Duboscq and A. Réveillac, *On a stochastic Hardy–Littlewood–Sobolev inequality with application to Strichartz estimates for the white noise dispersion*, arXiv:1711.07188 (2017)

[32] L. Escauriaza, G. Seregin, and V. Šverák, *Backward uniqueness for parabolic equations*, Arch. Ration. Mech. Anal. **169** (2003), no. 2, 147–157.

[33] E. Fedrizzi and F. Flandoli, *Noise prevents singularities in linear transport equations*, J. Funct. Anal. **264** (2013), no. 6, 1329–1354.

[34] E. Fedrizzi and F. Flandoli, *Hölder flow and differentiability for SDEs with nonregular drift.* Stoch. Anal. Appl. **31** (2013), no. 4, 708–736.

[35] E. Fedrizzi and F. Flandoli, *Pathwise uniqueness and continuous dependence of SDEs with non-regular drift*, Stochastics **83** (2011), no. 3, 241–257.

[36] E. Fedrizzi, W. Neves, and C. Olivera, *On a class of stochastic transport equations for L^2_{loc} vector fields*, Annali della Scuola Normale Superiore di Pisa. Classe di scienze **18** (2018), no. 2, 397–419.

[37] A. Figalli, *Existence and uniqueness of martingale solutions for SDEs with rough or degenerate coefficients*, J. Funct. Anal. **254** (2008), no. 1, 109–153.

[38] F. Flandoli, *Random perturbation of PDEs and fluid dynamic models*, Saint Flour summer school lectures 2010, Lecture Notes in Math., vol. **2015**, Springer, Berlin, 2011.

[39] F. Flandoli, M. Gubinelli, and E. Priola, *Remarks on the stochastic transport equation with Hölder drift*, Rend. Sem. Mat. Univ. Pol. Torino, **70**, (2012) 53–73.

[40] F. Flandoli, M. Gubinelli, and E. Priola, *Does noise improve well-posedness of fluid dynamic equations?*, Stochastic partial differential equations and applications, Quad. Mat., vol. **25**, Dept. Math., Seconda Univ. Napoli, Caserta, 2010, pp. 139–155.

[41] F. Flandoli, M. Gubinelli, and E. Priola, *Well-posedness of the transport equation by stochastic perturbation*, Invent. Math. **180** (2010), no. 1, 1–53.

[42] F. Flandoli, M. Gubinelli, and E. Priola, *Full well-posedness of point vortex dynamics corresponding to stochastic 2D Euler equations*, Stoch. Proc. Appl. **121** (2011), no. 7, 1445–1463.

[43] F. Flandoli, E. Issoglio, and F. Russo, *Multidimensional stochastic differential equations with distributional drift*, Trans. Amer. Math. Soc. **369** (2017), no. 3, 1665–1688.

[44] F. Flandoli, M. Maurelli, and M. Neklyudov, *Regularization by noise of a stochastic vector advection equation*, J. Math. Fluid Mech. **16** (2014), no. 4, 805–822.

[45] P. Friz and A. Shekhar, *Doob–Meyer for rough paths*, Bull. Inst. Math. Acad. Sin. (N.S.) **8** (2013), no. 1, 73–84.

[46] G.P. Galdi, *An introduction to the Navier–Stokes initial-boundary value problem*, in: Fundamental directions in mathematical fluid mechanics, Adv. Math. Fluid Mech., Birkhäuser, Basel, 2000, pp. 1–70.

[47] P. Gassiat and B. Gess, *Regularization by noise for stochastic Hamilton–Jacobi equations*, Probab. Theory Related Fields **173** (2019), no. 3-4, 1063–1098.

[48] B. Gess and M. Maurelli, *Well-posedness by noise for scalar conservation laws*, Comm. Partial Differential Equations **43** (2018), no. 12, 1702–1736.

[49] B. Gess and S. Smith, *Stochastic continuity equations with conservative noise*, J. Math. Pures Appl. (2019), to appear.

[50] B. Gess and P.E. Souganidis, *Long-time behavior, invariant measures, and regularizing effects for stochastic scalar conservation laws*, Comm. Pure Appl. Math. **70** (2017), no. 8, 1562–1597.

[51] I. Gyöngy and T. Martínez, *On stochastic differential equations with locally unbounded drift*, Czechoslovak Math. J. **51**(126) (2001), no. 4, 763–783.

[52] A.A. Kiselev and O.A. Ladyženskaya, *On the existence and uniqueness of the solution of the nonstationary problem for a viscous, incompressible fluid*, Izv. Akad. Nauk SSSR. Ser. Mat. **21** (1957), 655–680.

[53] N.V. Krylov, *Lectures on elliptic and parabolic equations in Hölder spaces*, Graduate Studies in Mathematics, **12**, AMS, Providence, RI, 1996.

[54] N.V. Krylov and M. Röckner, *Strong solutions of stochastic equations with singular time dependent drift*, Probab. Theory Related Fields **131** (2005), no. 2, 154–196.

[55] H. Kunita, *First order stochastic partial differential equations*, Stochastic analysis (Katata/Kyoto, 1982), North-Holland Math. Library, vol. **32**, North-Holland, 1984, pp. 249–269.

[56] H. Kunita, *Stochastic differential equations and stochastic flows of diffeomorphisms*, École d'été de probabilités de Saint-Flour, XII—1982, Lecture Notes in Math., vol. **1097**, Springer, Berlin, 1984, pp. 143–303.

[57] H. Kunita, *Stochastic flows and stochastic differential equations*, Cambridge Studies in Advanced Mathematics, vol. **24**, Cambridge University Press, Cambridge, 1990.

[58] O.A. Ladyženskaja, *Uniqueness and smoothness of generalized solutions of Navier–Stokes equations*, Zap. Naučn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI) **5** (1967), 169–185.

[59] J.-L. Lions and G. Prodi, *Un théorème d'existence et unicité dans les équations de Navier–Stokes en dimension 2*, C. R. Acad. Sci. Paris **248** (1959), 3519–3521.

[60] P.-L. Lions, *Mathematical topics in fluid mechanics. Vol. 1*, Oxford Lecture Series in Mathematics and its Applications, vol. **3**, The Clarendon Press Oxford University Press, New York, 1996.

[61] M. Maurelli, *Wiener chaos and uniqueness for stochastic transport equation*, C. R. Math. Acad. Sci. Paris **349** (2011), no. 11-12, 669–672.

[62] M. Maurelli, *Regularization by noise in finite dimension*, Ph.D. Thesis, Scuola Normale Superiore di Pisa, 2016.

[63] S.-E.A. Mohammed, T. Nilssen, and F. Proske, *Sobolev differentiable stochastic flows for SDEs with singular coefficients: applications to the transport equation*, Ann. Probab. **43** (2015), no. 3, 1535–1576.

[64] K. Nam, *Stochastic Differential Equations with Critical Drifts*, arXiv:1802.00074 (2018).

[65] W. Neves and C. Olivera, *Wellposedness for stochastic continuity equations with Ladyzhenskaya–Prodi–Serrin condition*, NoDEA Nonlinear Differential Equations Appl. **22** (2015), no. 5, 1247–1258.

[66] W. Neves and C. Olivera, *Stochastic continuity equations – a general uniqueness result*, Bull. Braz. Math. Soc. (N.S.) **47** (2016), no. 2, 631–639.

[67] T. Nilssen, *Rough path transport equation with discontinuous coefficient - regularization by fractional Brownian motion*, arXiv:1509.01154 (2015).

[68] C. Olivera, *Regularization by noise in one-dimensional continuity equation*, Potential Anal. (2018).

[69] N.I. Portenko, *Generalized Diffusion Processes*, Nauka, Moscow, 1982 In Russian; English translation: Amer. Math. Soc., Providence, Rhode Island, 1990.

[70] E. Priola, *Davie's type uniqueness for a class of SDEs with jumps*, Ann. Inst. Henri Poincaré Probab. Stat. **54** (2018), no. 2, 694–725.

[71] G. Prodi, *Un teorema di unicità per le equazioni di Navier–Stokes*, Ann. Mat. Pura Appl. (4) **48** (1959), 173–182.

[72] D. Revuz and M. Yor, *Continuous Martingales and Brownian Motion*, Springer-Verlag, Berlin, 1994.

[73] J. Serrin, *On the interior regularity of weak solutions of the Navier–Stokes equations*, Arch. Rational Mech. Anal. **9** (1962), 187–195.

[74] A.V. Shaposhnikov, *Some remarks on Davie's uniqueness theorem*, Proc. Edinb. Math. Soc. (2) **59** (2016), no. 4, 1019–1035.

[75] W. Stannat, (*Nonsymmetric*) *Dirichlet operators on L^1: existence, uniqueness and associated Markov processes*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. **28** (1999), 99–140.

[76] D.W. Stroock and S.R.S. Varadhan, *Multidimensional diffusion processes*, Springer-Verlag, Berlin Heidelberg, 1979.

[77] M. Veraar, *The stochastic Fubini theorem revisited*, Stochastics **84** (2012), no. 4, 543–551.

[78] Y.A. Veretennikov, *On strong solution and explicit formulas for solutions of stochastic integral equations*, Math. USSR Sb. **39** (1981) 387–403.

[79] J. Wei, G. Lv, and J.-L. Wu, *On weak solutions of stochastic differential equations with sharp drift coefficients*, arXiv:1711.05058 (2017).