



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

## Well-posedness by noise for scalar conservation laws

**Citation for published version:**

Gess, B & Maurelli, M 2019, 'Well-posedness by noise for scalar conservation laws', *Communications in partial differential equations*, vol. 43, no. 12, pp. 1702-1736.  
<https://doi.org/10.1080/03605302.2018.1535604>

**Digital Object Identifier (DOI):**

[10.1080/03605302.2018.1535604](https://doi.org/10.1080/03605302.2018.1535604)

**Link:**

[Link to publication record in Edinburgh Research Explorer](#)

**Document Version:**

Peer reviewed version

**Published In:**

Communications in partial differential equations

**General rights**

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

**Take down policy**

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact [openaccess@ed.ac.uk](mailto:openaccess@ed.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.



# WELL-POSEDNESS BY NOISE FOR SCALAR CONSERVATION LAWS

BENJAMIN GESS AND MARIO MAURELLI

ABSTRACT. We consider stochastic scalar conservation laws with spatially inhomogeneous flux. The regularity of the flux function with respect to its spatial variable is assumed to be low, so that entropy solutions are not necessarily unique in the corresponding deterministic scalar conservation law. We prove that perturbing the system by noise leads to well-posedness.

## 1. INTRODUCTION

In this paper we prove a well-posedness by noise result for the inhomogeneous scalar conservation laws

$$(1.1) \quad \begin{aligned} du(t, x) + b(x, u(t, x)) \cdot \nabla u(t, x) dt + \nabla u(t, x) \circ dW_t &= 0 \quad \text{on } [0, T] \times \mathbb{R}^d, \\ u(0, \cdot) &= u_0. \end{aligned}$$

If the vector field  $b$  lacks sufficient regularity with respect to the space variable  $x$ , ill-posedness can appear. For a counterexample see (1.7) below. In contrast, the main result of this paper (Theorem 3.1 below), here stated in a slightly simplified version, shows that noise restores well-posedness for some of these vector fields:

**Theorem.** *Assume that  $b \in L_{u,loc}^\infty(L_x^\infty) \cap L_{u,loc}^1(W_{x,loc}^{1,1})$  and that  $\operatorname{div} b \in L_{u,loc}^1(L_x^1) \cap L_x^p(L_{u,loc}^\infty)$  for some  $p > d$ ,  $p \leq \infty$ . Then, for every initial datum  $u_0$  in  $(L^1 \cap L^\infty)(\mathbb{R}^d)$ , there exists a unique entropy solution  $u$  to (1.1).*

The question of regularization and well-posedness by noise for SPDE has attracted considerable interest in recent years. One of the driving hopes in this field is to obtain the well-posedness by noise for nonlinear PDE arising in fluid dynamics, for which the deterministic counterpart does not or is not known to allow unique solutions. Despite considerable effort, only partial results in this direction could be obtained so far, cf. e.g. Flandoli [22], Flandoli, Romito [26, 27], Delarue, Flandoli, Vincenzi [16] and the references therein. In the linear setting (i.e. for  $b$  independent of  $u$ ), one of the prominent works in this direction by Flandoli, Gubinelli, Priola is [23] in which the well-posedness by noise for linear transport equations with irregular drift has been shown. More precisely, while weak solutions to

$$(1.2) \quad \partial_t u(t, x) + b(x) \cdot \nabla u(t, x) = 0 \quad \text{on } [0, T] \times \mathbb{R}^d$$

(omitting the initial condition for simplicity of notation) are not necessarily unique if  $\operatorname{div} b \notin L^\infty(\mathbb{R}^d)$  (cf. DiPerna, Lions [17], Ambrosio [1]), it has been shown in [23] that weak solutions to

$$(1.3) \quad du(t, x) + b(x) \cdot \nabla u(t, x) dt + \nabla u(t, x) \circ dW_t = 0 \quad \text{on } [0, T] \times \mathbb{R}^d$$

---

*Date:* April 17, 2018.

*2000 Mathematics Subject Classification.* H6015, 35R60, 35L65.

*Key words and phrases.* Stochastic scalar conservation laws, well-posedness by noise.

M.M. acknowledges funding from the European Research Council under the European Union's Seventh Framework Program (FP7/2007-2013) / ERC grant agreement nr. 258237.

are unique, provided  $b \in C_b^\alpha(\mathbb{R}^d)$  for some  $\alpha \in (0, 1)$ ,  $\operatorname{div} b \in L^p(\mathbb{R}^d)$  for some  $p > 2$  and  $W_t$  denotes a standard  $d$ -dimensional Wiener process. As pointed out in [23] their result yields *the first concrete example of a partial differential equation related to fluid dynamics that may lack uniqueness without noise, but is well-posed with a suitable noise* (cf. [23, p.3, 1.1 ff.]). On the other hand, as observed in [22, 23], in the nonlinear setting ( $d = 1$  for simplicity)

$$(1.4) \quad \partial_t u(t, x) + \partial_x u^2(t, x) = 0 \quad \text{on } [0, T] \times \mathbb{R}$$

the same type of noise seems to be of little use, since the stochastically perturbed equation

$$(1.5) \quad du(t, x) + \partial_x u^2(t, x) dt + \partial_x u(t, x) \circ dW_t = 0 \quad \text{on } [0, T] \times \mathbb{R}$$

reduces to the deterministic case (1.4) via the transformation  $v(t, x) := u(t, x + W_t)$ . That is, if  $u$  is a solution to (1.5) then  $v$  is a solution to (1.4) and vice versa. In particular, shocks and non-uniqueness of weak solutions still appear in (1.5). Hence, no well-posedness by noise, nor regularization by noise seems to be present in this case and it was concluded in [23]: *The generalization to nonlinear transport equations, where  $b$  depends on  $u$  itself, would be a major next step for applications to fluid dynamics but it turns out to be a difficult problem* (cf. [23, p.6, 1.11 ff.]).

The purpose of this work is to shed more light on the effect of linear multiplicative noise on nonlinear scalar conservation laws. In contrast to the above observation, we show that a similar effect of well-posedness by noise as obtained in [23] for (1.2) can be observed for (nonlinear) scalar conservation laws. More precisely, we consider inhomogeneous scalar conservation laws with irregular flux of the type

$$(1.6) \quad \begin{aligned} \partial_t u(t, x) + b(x, u(t, x)) \cdot \nabla u(t, x) &= 0 \quad \text{on } [0, T] \times \mathbb{R}^d, \\ u(0, \cdot) &= u_0, \end{aligned}$$

with a possibly space-irregular  $b$ . In particular, this includes the special case of inhomogeneous Burgers' equations  $b(x, u) = 2b(x)u$ . The model example

$$(1.7) \quad b(x, u) = 2\operatorname{sgn}(x)(\sqrt{|x|} \wedge K)u$$

for some  $K > 0$ ,  $u_0(\cdot) = 1_{[0,1]}(\cdot)$ ,  $d = 1$  shows that entropy solutions to (1.6) are not necessarily unique. Indeed, fix some time  $T > 0$  and choose  $K > T/2 + 1$  for simplicity. Then, there are several entropy solutions to (1.6), including the following two particular ones

$$u^1(t, x) := \begin{cases} 1 & \text{if } 0 \leq x \leq (\frac{t}{2} + 1)^2 \\ 0 & \text{otherwise} \end{cases}, \quad u^2(t, x) := \begin{cases} 1 & \text{if } -(\frac{t}{2})^2 \leq x \leq (\frac{t}{2} + 1)^2 \\ 0 & \text{otherwise} \end{cases},$$

on  $[0, T] \times \mathbb{R}$ . In contrast, the main result of this work (Theorem 3.1) shows that entropy solutions to the stochastically perturbed scalar conservation law (1.1) are unique, under certain assumptions on  $b$  and its divergence. Note that (1.7) satisfies these assumptions. Hence, this demonstrates that linear multiplicative noise has a similar regularizing effect in the case of nonlinear scalar conservation laws with irregular flux as it was obtained in the linear setting (i.e. for linear transport equation (1.3)) in [23]. To the authors' knowledge, this provides the first example of a nonlinear scalar conservation law that becomes well-posed by the inclusion of noise.

Let us comment more on the role of the noise in the presence of spatial inhomogeneities. In this setting, the transformation  $v(t, x) := u(t, x + W_t)$ , applied to (1.5) above, gives

$$\partial_t v(t, x) + b(x + W_t, v(t, x)) \cdot \nabla v(t, x) = 0 \quad \text{on } [0, T] \times \mathbb{R}^d.$$

Hence, the noise induces a random shift of the flux function in its spatial variable. As for the Burgers' equation (1.5), and contrary to the linear transport equation (1.3) this random shift cannot prevent the occurrence of shocks. However, our main result indicates that, as in the linear setting, this random shift still has an averaging and thus regularizing effect on spatial inhomogeneities.

Scalar conservation laws with irregular flux in divergence form have been used in several models, including models of traffic flow, flow in porous media and sedimentation processes (cf. Crasta, De Cicco, De Philippis [12]). In the present work, we choose to consider the non-divergence form in order to allow comparison to the results obtained in [23] and by Beck, Flandoli, Gubinelli, Maurelli in [7]. We expect that related arguments can be also applied to the corresponding divergence type equations, as it was demonstrated in the linear setting in [7], although nontrivial differences with the non-divergence case may arise (see Remark 1.1). This will be treated in a subsequent work. The respective study of conservation laws with irregular flux has attracted considerable interest in recent years, see Andreianov, Karlsen, Risebro [2, 3], Crasta, De Cicco, De Philippis, Ghiraldin [12, 13], Andreianov, Mitrović [4] among many more. Due to the spatial irregularity of the flux, entropy solutions to (1.6) are typically non-unique and several selection criteria to select a unique entropy solution have been introduced, corresponding to different physical phenomena and relative approximation procedures. Therefore, the study of selection methods for (1.6) is of high interest. The well-posedness result for

$$(1.8) \quad \begin{aligned} du(t, x) + b(x, u(t, x)) \cdot \nabla u(t, x) dt + \sigma \nabla u(t, x) \circ dW_t &= 0 \quad \text{on } [0, T] \times \mathbb{R}^d, \\ u(0, \cdot) &= u_0, \end{aligned}$$

with  $\sigma > 0$  obtained in this paper opens the way to study selection principles by vanishing noise  $\sigma \rightarrow 0$ . In the case of linear transport equations with irregular drift such vanishing noise selection methods have been analyzed by Attanasio, Flandoli [5], Delarue, Flandoli [5, 15] and it should be noted that in general the vanishing viscosity selection does not coincide with the vanishing noise selection. In analogy to linear stochastic transport equations (1.3), stochastic scalar conservation laws (1.1) model the evolution of passive scalars in turbulent fluids, so-called Kraichnan models.

The literature on regularization (i.e. improvement of regularity) and well-posedness (i.e. existence, uniqueness and possibly stability) by noise is vast and giving a complete survey at this point would exceed the purpose of this paper. Therefore, we will restrict to those that seem most relevant for the content of this work and refer to Flandoli [22], Flandoli, Romito [26, 27], Gyöngy, Pardoux [32], Butkovsky, Mytnik, Leonid [8] for further references and a more complete account of the literature. Concerning the case of transport equations with irregular drift (1.2), we mention the works by Flandoli, Gubinelli, Priola [23, 25], Flandoli, Fedrizzi [21], Beck, Flandoli, Gubinelli, Maurelli [7] and the references therein. In particular, we would like to emphasize the work [6] by Attanasio, Flandoli which provides a purely analytic approach to the effect of well-posedness by noise for (1.3), since the proof has served as an inspiration for some of the steps of the proof presented in this paper. A regularization by noise effect for (1.3) has been first obtained in [21] where it has been shown that solutions to (1.3) are smooth if the initial condition is, assuming that  $b$  satisfies certain integrability conditions, slightly more restrictive than the Ladyzhenskaya-Prodi-Serrin condition. A PDE-based approach and a generalization of these results to drifts  $b$  satisfying the Ladyzhenskaya-Prodi-Serrin condition and to divergence-type equations has been given in [7]. A path-by-path approach to well-posedness by noise has been introduced by Catellier, Gubinelli in [10] and was

used by Catellier in [9] for transport equations. Another approach based on Malliavin calculus has been introduced by Menoukeu-Pamen, Meyer-Brandis, Nilssen, Proske in [42] and developed in a series of papers, cf. e.g. Mohammed, Nilssen, Proske [43] on transport equations. We also refer to Duboscq, Réveillac [18] for a generalization of [23] to SDE with random drift.

In some (typically nonlinear) situations, the spatial dependence of the noise coefficients has proven to be crucial in order to obtain well-posedness by noise results. More precisely, in Flandoli, Gubinelli, Priola [24] well-posedness by spatially dependent linear transport noise for point vortex dynamics informally related to stochastic 2D-Euler equations has been shown. In [16] it has been shown that the same type of noise can prevent the collapse of point charges in Vlasov-Poisson equations. More recently, regularizing effects of nonlinear noise in the setting of (nonlinear) scalar conservation laws has been observed by Gess, Souganidis in [31] and in the setting of fully nonlinear PDE by Gassiat, Gess in [28]. Well-posedness of stochastic scalar conservation laws with random flux has been considered by Lions, Perthame, Souganidis [37, 38], Gess, Souganidis, [30], Mariani [41].

We next present the idea and an outline of the proof. Our treatment of (1.1) is based on the kinetic formulation of (stochastic) scalar conservation laws as introduced by Lions, Perthame, Tadmor in [39]. For a function  $u : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}$  we introduce the kinetic function  $\chi(t, x, \xi) : [0, T] \times \mathbb{R}^d \times \mathbb{R} \rightarrow \mathbb{R}$  by

$$(1.9) \quad \chi(t, x, \xi) = \chi(u(t, x), \xi) := \mathbf{1}_{\xi < u(t, x)} - \mathbf{1}_{\xi < 0}.$$

In the case of a smooth spatial inhomogeneity  $b$  and smooth driving signal  $W$ ,  $u$  is an entropy solution to (1.1) iff  $\chi$  solves the following equation, in the sense of distributions,

$$(1.10) \quad \partial_t \chi = -b(x, \xi) \cdot \nabla \chi - \nabla \chi \cdot \dot{W}_t + \partial_\xi m,$$

where  $m$  is a nonnegative bounded random measure on  $[0, T] \times \mathbb{R}^d \times \mathbb{R}$  and the derivatives are intended with respect to  $x$  unless differently specified. In the general case of (1.1), we take (1.9), (1.10) as the definition of an entropy solution to (1.1), where now the term  $\nabla \chi \cdot \dot{W}_t$  should be interpreted as a Stratonovich integral, or more precisely,

$$(1.11) \quad \begin{aligned} \partial_t \chi &= -b(x, \xi) \cdot \nabla \chi - \nabla \chi \circ dW_t + \partial_\xi m \\ &= -b(x, \xi) \nabla \chi - \nabla \chi \cdot dW_t + \frac{1}{2} \Delta \chi + \partial_\xi m, \end{aligned}$$

see Definition 2.8 below for details. As in the deterministic case, the notion of a generalized kinetic solution is convenient in the construction of an entropy solution since, roughly speaking, the class of generalized kinetic solutions is stable under weak limits. Roughly speaking, a function  $f$  is said to be a generalized kinetic solution to (1.1) if  $f$  solves (1.10) for some nonnegative measure  $m$  and  $|f| = \text{sgn}(\xi)f \leq 1$ ,  $\partial_\xi f = \delta_0 - \nu$  for some nonnegative measure  $\nu$ . The key difference to an entropy solution is that  $f$  is not assumed to be of the form of a kinetic function (1.9) for some function  $u$ .

The main difficulty then lies in proving that generalized kinetic solutions are in fact entropy solutions, which boils down to proving  $|f| = 1$  a.e.. In order to prove this we aim to estimate the difference  $|f| - f^2$  based on (1.11). The proof now consists of two steps. In the first step, an (in)equality for  $|f| - f^2$  is derived based on renormalization techniques (cf. [1, 17]) using the assumption  $b \in L^1_{\xi, \text{loc}}(W^{1,1}_{x, \text{loc}})$ . Informally, this leads to the equality

$$\partial_t (|f| - f^2) + b(x, \xi) \cdot \nabla (|f| - f^2) + \nabla (|f| - f^2) \circ dW_t = (\text{sgn}(\xi) - 2f) \partial_\xi m.$$

Passing to the Itô formulation and taking the expectation, we informally “gain a Laplacian” similarly to [6]. The main difficulty at this point that is due to the nonlinearity of (1.1) is the additional singular term  $\partial_\xi m$ . To handle this term, in the second step, we integrate in both  $\omega$  and  $\xi$  (in the Itô formulation), which informally yields

$$\begin{aligned} & \partial_t \int \mathbb{E}(|f| - f^2) d\xi + \int b(x, \xi) \cdot \nabla \mathbb{E}(|f| - f^2) d\xi + \frac{1}{2} \Delta \int \mathbb{E}(|f| - f^2) d\xi \\ &= \mathbb{E} \int \varphi(\operatorname{sgn}(\xi) - 2f) \partial_\xi m d\xi. \end{aligned}$$

Since  $\partial_\xi f = \delta_0 - \nu \leq \delta_0$  this implies

$$\partial_t \int \mathbb{E}(|f| - f^2) d\xi + \int b(x, \xi) \cdot \nabla \mathbb{E}(|f| - f^2) d\xi + \frac{1}{2} \Delta \int \mathbb{E}(|f| - f^2) d\xi \leq 0$$

This is a linear *parabolic* PDE in  $\int \mathbb{E}(|f| - f^2) d\xi$ , but, in contrast to the linear setting (i.e. for linear transport equation) in [6], this PDE is not closed, since it involves both  $\int \mathbb{E}(|f| - f^2) d\xi$  and  $\int b(x, \xi) \cdot \nabla \mathbb{E}(|f| - f^2) d\xi$ . The rigorous analysis is carried out by passing to the distributional form. The problem that the above PDE is non-closed then relates to finding a nonnegative test function  $\varphi$ , independent of  $\xi$ , that satisfies for every  $\xi$ ,

$$\partial_t \varphi + \operatorname{div}(b(x, \xi)\varphi) + \Delta \varphi \leq C,$$

for some constant  $C > 0$ . In the analysis of this PDE, we rely on the boundedness assumption on  $b$  and the integrability assumption on  $\operatorname{div} b$ . We conclude (Lemma 3.3) that

$$\mathbb{E} \int_{\mathbb{R}^d \times \mathbb{R}} (|f_t| - f_t^2) d\xi dx \leq C \int_{\mathbb{R}^d \times \mathbb{R}} (|f_0| - f_0^2) d\xi dx,$$

which implies that  $|f| = 1$  and ends the proof.

*Remark 1.1.* Scalar conservation laws in divergence form

$$\partial_t u + \operatorname{div}(b(x, u)u) = 0,$$

with spatially irregular drift  $b$ , can have different pathological behaviour than equation (1.6), such as concentration of mass. For further details we refer to [7, 29] and the references therein.

**1.1. Notation.** We let  $(\Omega, \mathcal{A}, P)$  be a measurable space,  $(\mathcal{F}_t)_t$  be a normal filtration on  $(\Omega, \mathcal{A}, P)$  (i.e.  $(\mathcal{F}_t)_t$  is right-continuous and  $\mathcal{F}_0$  contains the null sets of  $(\Omega, \mathcal{A}, P)$ ) and  $W = (W_t)_t$  be a  $d$ -dimensional Brownian motion on  $\Omega$  with respect to the filtration  $(\mathcal{F}_t)_t$ . For a  $\sigma$ -finite measure space  $(E, \mathcal{E}, \mu)$ , we say that a function  $f : E \rightarrow \mathbb{R}$  is measurable if, for every Borel subset  $A$  of  $\mathbb{R}$ ,  $f^{-1}(A)$  is in  $\mathcal{E}$ . Given a Banach space  $V$ , we define  $L^0(E; V) = L^0(E, \mathcal{E}, \mu; V)$  in two cases:

- (1) if  $V = U^*$  is the dual space of a separable Banach space  $U$ ,  $L^0(E; V)$  is defined as the space of classes of equivalence, under the relation “ $f = g$   $\mu$ -a.e.”, of weakly- $*$  measurable functions  $f : E \rightarrow V$ , i.e., for every  $\varphi$  in  $U$ ,  $x \mapsto \langle f(x), \varphi \rangle_{V, U}$  is measurable. This applies to the case of  $V = \mathcal{M}(D)$ , the space of finite signed measure over a domain  $D$  of  $\mathbb{R}^n$ ,  $L^\infty(D)$ ,  $L^p(D)$  for  $1 < p < \infty$ ;
- (2) if  $V$  is separable,  $L^0(E; V)$  is defined as the space of classes of equivalence, under the relation “ $f = g$   $\mu$ -a.e.”, of weakly measurable functions  $f : E \rightarrow V$ , i.e., for every  $\varphi$  in  $V^*$ ,  $x \mapsto \langle f(x), \varphi \rangle_{V, V^*}$  is measurable. This applies to the case of  $V = C_0(D)$ , the space of continuous bounded function on a domain  $D$  of  $\mathbb{R}^n$  vanishing at infinity,  $L^1(D)$ ,  $L^p(D)$  for  $1 < p < \infty$ .

Similarly, one can define  $L^p(E; V)$ . When  $V = \mathbb{R}$  we simply write  $L^0(E)$ ,  $L^p(E)$  (the usual  $L^p$  spaces). For a metric, locally compact,  $\sigma$ -compact space  $S$ , the space  $\mathcal{M}(S)$  denotes the space of finite signed Borel measures on  $S$ ,  $\mathcal{M}_+(S)$  the subset of finite nonnegative Borel measures. More details on these spaces and on measurability issues are given in the appendix. When not otherwise stated, the spaces  $\Omega$ , resp.  $[0, T] \times \Omega$  are considered endowed with the  $\sigma$ -algebrae  $\mathcal{A}$ , resp.  $\mathcal{B}([0, T]) \otimes \mathcal{A}$ ;  $\mathcal{P}$  denotes the progressive  $\sigma$ -algebra on  $[0, T] \times \Omega$ , progressive measurability is measurability with respect to  $\mathcal{P}$ . We use progressive measurability instead of predictability because càdlàg adapted processes are  $\mathcal{P}$ -measurable. The stochastic Fubini theorem [45, Exercise 5.17], that we put to use later-on, is stated for predictable processes but can be immediately extended to progressively measurable processes in our context (where the martingale part of the integrator is a Brownian motion or an integral with respect to Brownian motion). The concepts of entropy solutions, kinetic solutions, generalized kinetic solutions, kinetic measures have always to be understood in the sense of equivalence classes, although we will often consider them as functions when this does not create confusion. In cases where we need to work with representatives this will be indicated, although we will often use the same symbol for the class and the representative.

The variables  $t, \omega, x, \xi$  denote elements resp. in  $[0, T], \Omega, \mathbb{R}^d, \mathbb{R}$ . We often use the short notation  $L_x^p, L_{t,\omega,x}^p, \mathcal{M}_x, \dots$  for the spaces  $L^p(\mathbb{R}^d), L^p([0, T] \times \Omega \times \mathbb{R}^d), \dots$  and  $L_{\xi,[-R,R]}^p$  for the space  $L^p([-R, R])$ . We also use the notation  $b \in L_x^p(L_{\xi,loc}^\infty)$ ,  $b \in L_{\xi,loc}^1(W_{x,loc}^{1,1})$ , ... to state that  $b \in L_x^p(L_{\xi,[-R,R]}^\infty)$  for every  $R > 0$ ,  $b \in L_{\xi,[-R,R]}^1(W_{x,B_R}^{1,1})$  for every  $R > 0, \dots$ . The symbols  $\nabla, \operatorname{div}, \Delta$ , if not differently specified, are referred to derivatives in  $x$ , while derivatives in  $t$  and  $\xi$  are denoted by  $\partial_t, \partial_\xi$ . As usual in probability theory,  $\varphi_t$  denotes the evaluation at time  $t$ , that is,  $\varphi_t = \varphi(t)$  [note however that the subscript  $t$  in  $\partial_t$  does only denote the time derivative and not its evaluation at  $t$ : in particular,  $\int_0^t g \partial_t \varphi \, dr = \int_0^t g(r) \partial_t \varphi(r) \, dr$  denotes the integral from 0 to  $t$  of  $g$  times the time derivative of  $\varphi$ ]. The symbol  $\langle \cdot, \cdot \rangle$  denotes the scalar product in  $L_{x,\xi}^2$ , unless differently specified. For example,  $\langle \cdot, \cdot \rangle_{t,x,\xi}$  denotes the scalar product in  $L_{t,x,\xi}^2$ . Sometimes, for a measure  $m$  on  $[0, T] \times \mathbb{R}^d \times \mathbb{R}$ , we use the notation  $\langle m, \varphi \rangle dt$  for  $\int_{\mathbb{R}^d \times \mathbb{R}} \varphi(t, x, \xi) m(dt, dx, d\xi)$ . The convolution operator is denoted by  $*_{var}$ , where  $var$  stands for the variable (usually  $x$  or  $\xi$  or both) for which the convolution is performed. The function  $\rho$  denotes a smooth nonnegative compactly supported even function on  $\mathbb{R}^d$  such that  $\int_{\mathbb{R}^d} \rho(x) \, dx = 1$ , and  $\rho^\epsilon := \epsilon^{-d} \rho(\epsilon^{-1} \cdot)$ . Similarly  $\bar{\rho}$  denotes a smooth nonnegative compactly supported even function on  $\mathbb{R}$  such that  $\int_{\mathbb{R}} \bar{\rho}(\xi) \, d\xi = 1$ , and  $\bar{\rho}^\delta = \delta^{-1} \bar{\rho}(\delta^{-1} \cdot)$ . In statements and proofs, the letter  $C$  denotes a generic positive constant, which can change from line to line and can depend on  $d$  (dimension) and  $p$  (integrability exponent assumed for  $\operatorname{div} b$ ). In accordance to (1.9) we use the notation  $\chi(\xi, u) = 1_{\xi < u} - 1_{\xi < 0}$ . When we use the kinetic formulation, we write  $b$  for  $b(x, \xi)$ .

**1.2. Organization of the paper.** In Section 2 we introduce the notions of entropy, kinetic and generalized kinetic solutions to (1.1), prove a flow-transformation result linking (1.1) to a scalar conservation law with random coefficients and prove the existence of generalized entropy solutions based on stable  $L^p$ -estimates. Some subtle measurability properties are postponed to the second Appendix 5. The results and definitions in Section 2 are applicable under mild assumptions on  $b$  and, in particular, apply without change to the non-perturbed case. In Section 3 it is shown that generalized entropy solutions are entropy solutions and their uniqueness is deduced using certain parabolic PDE estimates given in the first Appendix 4.

## 2. DEFINITIONS AND THE EXISTENCE OF GENERALIZED KINETIC SOLUTIONS

In this section we give some general definitions and results, which hold also without noise. In the case of a smooth vector field  $b$ , there exists a unique entropy solution. In the general case, even the existence of an entropy solution may not hold in general. However, one can get the existence of a so-called generalized kinetic solution.

We start defining the concept of an entropy solution.

**Definition 2.1.** A (stochastic) bounded kinetic measure is a map  $m : \Omega \rightarrow \mathcal{M}([0, T] \times \mathbb{R}^d \times \mathbb{R})$ , weakly- $*$  measurable, satisfying the following properties:

- (i)  $m \in L^\infty(\Omega; \mathcal{M}([0, T] \times \mathbb{R}^d \times \mathbb{R}))$ ;
- (ii)  $m$  is a.s. non-negative and supported on  $[0, T] \times \mathbb{R}^d \times [-R, R]$  for some  $R > 0$  independent of  $\omega$ ;
- (iii) for every  $\varphi \in C_c^\infty([0, T] \times \mathbb{R}^d \times \mathbb{R})$ , the process  $(t, \omega) \mapsto \int_{[0, t] \times \mathbb{R}^d \times \mathbb{R}} \varphi dm$  is progressively measurable.

Here and in what follows, we can extend definitions and formulations to test functions  $\varphi$  which are not necessarily compactly supported in the  $\xi$  variable, because of the assumption that  $m$  is supported on  $[0, T] \times \mathbb{R}^d \times [-R, R]$ .

**Definition 2.2.** Let  $b \in L^1_{loc}(\mathbb{R}^{d+1})$  with  $\operatorname{div} b \in L^1_{loc}(\mathbb{R}^{d+1})$  and let  $u_0 \in (L^1 \cap L^\infty)(\mathbb{R}^d)$ . An entropy solution to (1.1) is a measurable function  $u : [0, T] \times \Omega \times \mathbb{R}^d \rightarrow \mathbb{R}$ , such that  $\chi(t, \omega, x, \xi) = \chi(u(t, \omega, x), \xi) = 1_{\xi < u(t, \omega, x)} - 1_{\xi < 0}$  satisfies the following properties:

- (i)  $\chi \in L^\infty([0, T] \times \Omega; L^1(\mathbb{R}^d \times \mathbb{R}))$  and is supported on  $[0, T] \times \Omega \times \mathbb{R}^d \times [-R, R]$  for some  $R > 0$ ;
- (ii)  $\chi$  is a weakly- $*$  progressively measurable  $L^\infty_{x, \xi}$ -valued process;
- (iii) there exists a bounded kinetic measure  $m$  such that, for every test-function  $\varphi \in C_c^\infty([0, T] \times \mathbb{R}^d \times \mathbb{R})$ , it holds, for a.e.  $(t, \omega)$ ,

$$(2.1) \quad \begin{aligned} \langle \chi_t, \varphi_t \rangle &= \langle \chi_0, \varphi_0 \rangle + \int_0^t \langle \chi, \partial_t \varphi + \operatorname{div}(b(x, \xi)\varphi) \rangle dr + \int_0^t \langle \chi, \nabla \varphi \rangle dW \\ &\quad + \frac{1}{2} \int_0^t \langle \chi, \Delta \varphi \rangle dr - \int_{[0, t] \times \mathbb{R}^d \times \mathbb{R}} \partial_\xi \varphi dm, \end{aligned}$$

$$\text{with } \chi_0(x, \xi) = 1_{\xi < u_0(x)} - 1_{\xi < 0}.$$

The function  $\chi$  is called a kinetic solution.

The well-known definitions of entropy solutions, kinetic solutions and kinetic measures in the case of deterministic scalar conservation laws are recovered in the above definitions by removing the  $\omega$  dependence, the progressive measurability assumptions as well as the second order term and stochastic integral in (2.1).

*Remark 2.3.* (i) For every kinetic solution  $\chi$  and test function  $\varphi \in C_c^\infty([0, T] \times \mathbb{R}^d \times \mathbb{R})$ ,  $(t, \omega) \mapsto \langle \chi_t, \varphi_t \rangle$  is a semimartingale admitting a càdlàg version. More precisely, it admits a version which is the sum of a continuous martingale and a process with BV paths. Indeed, for every  $\varphi$  and every representative of  $m$ , for a.e.  $\omega$ , the function  $t \mapsto \int_{[0, t] \times \mathbb{R}^d \times \mathbb{R}} \partial_\xi \varphi dm$  is of finite variation.

The processes  $\int_{[0, t] \times \mathbb{R}^d \times \mathbb{R}} \partial_\xi \varphi dm$  and  $\int_{[0, t] \times \mathbb{R}^d \times \mathbb{R}} \partial_\xi \varphi dm$  are progressively measurable and resp. càdlàg, càglàd. Moreover, the (random) times where the paths are discontinuous depend only on  $m$  and not on  $\varphi$ .

(ii) More general, let  $\varphi : [0, T] \times \Omega \times \mathbb{R}^d \times \mathbb{R}^m \rightarrow \mathbb{R}$  be a measurable bounded function such that: 1) for every  $(x, \xi, z)$ ,  $(t, \omega) \mapsto \varphi(t, \omega, x, \xi, z)$  is progressively



measurable; 2) for a.e.  $\omega$ ,  $(t, x, \xi, z) \mapsto \varphi(t, \omega, x, \xi, z)$  is continuous. Then, for every representative of  $m$ , the maps

$$\begin{aligned} (t, z, \omega) &\mapsto \int_{[0,t] \times \mathbb{R}^d \times \mathbb{R}} \varphi(r, \omega, x, \xi, z) m(r, x, \xi) dr dx d\xi, \\ (t, z, \omega) &\mapsto \int_{[0,t] \times \mathbb{R}^d \times \mathbb{R}} \varphi(r, \omega, x, \xi, z) m(r, x, \xi) dr dx d\xi \end{aligned}$$

are measurable and:

- 1) for each  $z$  fixed, progressively measurable in  $(t, \omega)$ ;
- 2) for a.e.  $\omega$ , with zero set independent of  $z$ , and each  $z$  fixed, càdlàg, resp. càglàd, in  $t$ ;
- 3) for a.e.  $\omega$ , with zero set independent of  $t$ , and each  $t$  fixed, continuous in  $z$ .

This result follows from the combination of the following two facts: a) We can apply Remark 5.4 (ii) below to  $z$  fixed to get the progressive measurability and càdlàg/càglàd property, the latter depending only on  $m$  and not on  $z$ . b) The integral is continuous with respect to  $z$ , as consequence of the dominated convergence theorem.

*Remark 2.4.* By equation (2.1), for every test function  $\varphi \in C_c^\infty([0, T] \times \mathbb{R}^d \times \mathbb{R})$ , the quadratic covariance between  $\langle \chi, \nabla \varphi \rangle$  and  $W$  is  $[\langle \chi, \nabla \varphi \rangle, W]_t = \int_0^t \langle \chi, \Delta \varphi \rangle dr$  for a.e.  $(t, \omega)$ . Note that by abuse of notation we here use  $\langle \chi, \nabla \varphi \rangle$  to also denote its càdlàg version.

Therefore, the Stratonovich integral  $\int_0^t \langle \chi, \nabla \varphi \rangle \circ dW$  makes sense and equation (2.1) can be rewritten as

$$\begin{aligned} \langle \chi_t, \varphi_t \rangle &= \langle \chi_0, \varphi_0 \rangle + \int_0^t \langle \chi, \partial_t \varphi + \operatorname{div}(b\varphi) \rangle dr + \int_0^t \langle \chi, \nabla \varphi \rangle \circ dW \\ &\quad - \int_{[0,t] \times \mathbb{R}^d \times \mathbb{R}} \partial_\xi \varphi dm. \end{aligned}$$

In particular, we see here that equation (2.1) is of hyperbolic type.

*Remark 2.5.* By the definition of  $\chi$  we have immediately that, for every  $1 \leq p < \infty$ , for a.e.  $(t, \omega, x)$ ,

$$\begin{aligned} u(t, \omega, x) &= \int_{\mathbb{R}} \chi(t, \omega, x, \xi) d\xi, \\ (2.2) \quad \frac{1}{p} |u(t, \omega, x)|^p &= \int_{\mathbb{R}} |\xi|^{p-1} \operatorname{sgn}(\xi) \chi(t, \omega, x, \xi) d\xi. \end{aligned}$$

Therefore, the weakly-\* progressive measurability of  $\chi$  implies that of  $u$  and  $|u|^p$ .

Conversely, if  $u$  is an  $L_x^\infty$ -valued weakly-\* progressively measurable process, then by Proposition 5.2 below  $u$  is  $\mathcal{P} \otimes \mathcal{B}(\mathbb{R}^d)$ -measurable as a real-valued function of  $(t, \omega, x)$  (more precisely, there exists a version of  $u$  which is  $\mathcal{P} \otimes \mathcal{B}(\mathbb{R}^d)$ -measurable). Since  $(v, \xi) \mapsto 1_{\xi < v} - 1_{\xi < 0}$  is Borel measurable, the function  $(t, \omega, x, \xi) \mapsto \chi(t, \omega, x, \xi)$  is  $\mathcal{P} \otimes \mathcal{B}(\mathbb{R}^d) \otimes \mathcal{B}(\mathbb{R})$ -measurable; that is,  $\chi$  is a  $L_{x, \xi}^\infty$ -valued weakly-\* progressively measurable process.

From the formulas above and the fact that  $\chi = 0$  for  $|\xi| > R$ , it follows that  $u$  is in  $L^\infty([0, T] \times \Omega; L^\infty(\mathbb{R}^d)) \cap L^\infty([0, T] \times \Omega; L^1(\mathbb{R}^d))$ . Hence,  $u$  is in  $L^\infty([0, T] \times \Omega; L^p(\mathbb{R}^d))$  and  $\chi$  is in  $L^\infty([0, T] \times \Omega; L^p(\mathbb{R}^d \times \mathbb{R}))$  for every  $p \in [1, \infty]$ .

**2.1. A flow transformation.** Before giving the existence result, we recall the following transformation that links equation (1.1) to a scalar conservation law with random coefficients.

**Proposition 2.6.** *Let  $b \in L^1_{loc}(\mathbb{R}^{d+1})$  with  $\operatorname{div} b$  in  $L^1_{loc}(\mathbb{R}^{d+1})$ . A function  $u$  is an entropy solution to (1.1) iff the function  $\tilde{u}(t, x) := u(t, x + W_t)$  is  $L^\infty_x$ -valued weakly-\* progressively measurable and is a.s. an entropy solution to*

$$(2.3) \quad \partial_t \tilde{u}(t, x) + b(x + W_t, \tilde{u}(t, x)) \cdot \nabla \tilde{u}(t, x) = 0.$$

More precisely,  $\chi = \chi(u)$  is a kinetic solution to (1.1) with kinetic measure  $m$  iff:

- (i)  $\tilde{\chi}(t, x, \xi) := 1_{\xi < \tilde{u}(t, \omega, x)} - 1_{\xi < 0} = \chi(t, x + W_t, \xi)$  is  $L^\infty_{x, \xi}$ -valued weakly-\* progressively measurable.
- (ii)  $\tilde{m}(t, x, \xi) = m(t, x + W_t, \xi)$  is weakly-\* progressively measurable, that is, for every  $\psi \in C_c^\infty([0, T] \times \mathbb{R}^d \times \mathbb{R})$ , the process  $(t, \omega) \mapsto \int_{[0, t] \times \mathbb{R}^d \times \mathbb{R}} \psi d\tilde{m}$  is progressively measurable.
- (iii) For a.e.  $\omega$ ,  $\tilde{\chi}^\omega$  is a kinetic solution to (2.3) with kinetic measure  $\tilde{m}^\omega$ . In particular, in the sense of distributions,

$$(2.4) \quad \partial_t \tilde{\chi} + b(x + W_t, \xi) \cdot \nabla \tilde{\chi} = \partial_\xi \tilde{m}.$$

*Proof. Step 1: Progressive measurability.*

Progressive measurability of  $\tilde{\chi}$  can be deduced from progressive measurability of  $\chi$  and vice versa. Indeed, for every  $\varphi$  in  $C_c^\infty(\mathbb{R}^d \times \mathbb{R})$ ,  $\langle \tilde{\chi}, \tilde{\varphi} \rangle = \langle \chi, \tilde{\varphi}(x - W_t, \xi) \rangle$  is progressively measurable, by Remark 5.4. A similar reasoning applies to  $\tilde{m}$ : For every  $\tilde{\varphi} \in C_c^\infty([0, T] \times \mathbb{R}^d \times \mathbb{R})$ , the process  $(t, \omega) \mapsto \int_{[0, t] \times \mathbb{R}^d \times \mathbb{R}} \tilde{\varphi} d\tilde{m} = \int_{[0, t] \times \mathbb{R}^d \times \mathbb{R}} \varphi(r, x - W_r, \xi) dm$  is progressively measurable, again by Remark 5.4.

*Step 2: Equation (2.1) implies (2.4).*

Since  $\langle \tilde{\chi}, \tilde{\varphi} \rangle = \langle \chi, \tilde{\varphi}(x - W_t, \xi) \rangle$  for any (deterministic) test function  $\tilde{\varphi}$ , the statement would follow if we could take  $\tilde{\varphi}(x - W_t)$  as a test function. Unfortunately, this is not possible, since  $\tilde{\varphi}(x + W_t)$  is not deterministic. Therefore we use a regularization procedure: We consider  $\chi^\epsilon$ , a regularization of  $\chi$  with respect to  $x$  and  $\xi$ . Then, for fixed  $x$  and  $\xi$ , we multiply  $\chi^\epsilon$  by  $\tilde{\varphi}(x - W_t)$  using Itô's formula, integrate in  $x$  and  $\xi$  and pass to the limit  $\epsilon \rightarrow 0$ .

We consider a regularization of  $\chi$  in both  $x$  and  $\xi$ , i.e.  $\chi_t^\epsilon(x, \xi) := \langle \chi_t, \rho_\epsilon(x - \cdot) \bar{\rho}_\epsilon(\xi - \cdot) \rangle$ . For every  $(x, \xi)$ , we have the following equation, outside a null set possibly depending on  $(x, \xi)$  and  $\epsilon$ :

$$\begin{aligned} \chi^\epsilon(t, x, \xi) &= \chi^\epsilon(0, x, \xi) + \int_0^t \langle \chi, \operatorname{div}(b(\cdot, \cdot) \rho_\epsilon(x - \cdot) \bar{\rho}_\epsilon(\xi - \cdot)) \rangle dr \\ &\quad + \int_0^t \langle \chi, \nabla \rho_\epsilon(x - \cdot) \bar{\rho}_\epsilon(\xi - \cdot) \rangle dW + \frac{1}{2} \int_0^t \langle \chi, \Delta \rho_\epsilon(x - \cdot) \bar{\rho}_\epsilon(\xi - \cdot) \rangle dr \\ &\quad - \int_{[0, t]} \langle m, \partial_\xi \rho_\epsilon(x - \cdot) \bar{\rho}_\epsilon(\xi - \cdot) \rangle dr. \end{aligned}$$

We multiply  $\chi^\epsilon$  by  $\tilde{\varphi}(t, x - W_t, \xi)$  and use Itô's formula for càdlàg processes (see, for example, [44, Chapter II Theorem 33]), applied to  $f(x, y) = xy$ . Note that no jump term appears here because the function  $f$  is bilinear and thus, with the notation of Protter [44, Chapter II Theorem 33],  $f(x_s, y_s) - f(x_{s-}, y_s) - \partial_x f(x_{s-}, y_s) \cdot \Delta x_s = 0$ .

Hence, we get, outside a null set as above,

$$\begin{aligned}
& \chi^\epsilon(t, x, \xi) \tilde{\varphi}(t, x - W_t, \xi) \\
&= \chi^\epsilon(0, x, \xi) \tilde{\varphi}(0, x, \xi) + \int_0^t \chi^\epsilon(r, x, \xi) \partial_t \tilde{\varphi}(r, x - W, \xi) dr \\
&+ \int_0^t \langle \chi, \operatorname{div}(b(\cdot, \cdot) \rho_\epsilon(x - \cdot) \bar{\rho}_\epsilon(\xi - \cdot)) \rangle \tilde{\varphi}(r, x - W, \xi) dr \\
&+ \int_0^t \langle \chi, \nabla \rho_\epsilon(x - \cdot) \bar{\rho}_\epsilon(\xi - \cdot) \rangle \tilde{\varphi}(r, x - W, \xi) dW \\
&+ \frac{1}{2} \int_0^t \langle \chi, \Delta \rho_\epsilon(x - \cdot) \bar{\rho}_\epsilon(\xi - \cdot) \rangle \tilde{\varphi}(r, x - W, \xi) dr \\
&- \int_{[0, t]} \langle m, \rho_\epsilon(x - \cdot) \partial_\xi \bar{\rho}_\epsilon(\xi - \cdot) \rangle \tilde{\varphi}(r, x - W, \xi) dr \\
&- \int_0^t \chi^\epsilon(r, x, \xi) \nabla \tilde{\varphi}(r, x - W, \xi) dW + \frac{1}{2} \int_0^t \chi^\epsilon(r, x, \xi) \Delta \tilde{\varphi}(r, x - W, \xi) dr \\
&+ \int_0^t \langle \chi, \nabla \rho_\epsilon(x - \cdot) \bar{\rho}_\epsilon(\xi - \cdot) \rangle \cdot \nabla \tilde{\varphi}(r, x - W, \xi) dr.
\end{aligned}$$

By the stochastic Fubini theorem (see for example Revuz, Yor [45, Exercise 5.17]) and Remark 2.3, all the addends have measurable versions in  $(t, \omega, x, \xi)$ ; moreover, for these versions the equality above is true for a.e.  $(t, \omega, x, \xi)$ , we can integrate in  $x$  and in  $\xi$  and exchange the order of integration. We do so and bring the convolution on  $\varphi$ : we get, with  $\varphi(t, x, \xi) = \tilde{\varphi}(t, x - W_t, \xi)$ ,

$$\begin{aligned}
\langle \chi_t, \varphi_t^\epsilon \rangle &= \langle \chi_0, \varphi_0^\epsilon \rangle + \int_0^t \langle \chi, \partial_t \varphi^\epsilon \rangle dr + \int_0^t \langle \chi, \operatorname{div}(b\varphi^\epsilon) \rangle dr \\
&+ \int_0^t \langle \chi, \nabla \varphi^\epsilon \rangle dW + \frac{1}{2} \int_0^t \langle \chi, \Delta \varphi^\epsilon \rangle dr - \int_{[0, t]} \langle m, \partial_\xi \varphi^\epsilon \rangle dr \\
&- \int_0^t \langle \chi, \nabla \varphi^\epsilon \rangle dW + \frac{1}{2} \int_0^t \langle \chi, \Delta \varphi^\epsilon \rangle dr - \int_0^t \langle \chi, \Delta \varphi^\epsilon \rangle dr \\
&= \langle \chi_0, \varphi_0^\epsilon \rangle + \int_0^t \langle \chi, \partial_t \varphi^\epsilon \rangle dr + \int_0^t \langle \chi, \operatorname{div}(b\varphi^\epsilon) \rangle dr - \int_{[0, t]} \langle m, \partial_\xi \varphi^\epsilon \rangle dr.
\end{aligned}$$

Finally, we let  $\epsilon$  go to 0 and use the change of variable  $\tilde{x} = x - W_t$ , to obtain

$$(2.5) \quad \langle \tilde{\chi}_t, \tilde{\varphi}_t \rangle = \langle \tilde{\chi}_0, \tilde{\varphi}_0 \rangle + \int_0^t \langle \tilde{\chi}, \partial_t \tilde{\varphi} \rangle dr + \int_0^t \langle \tilde{\chi}, \operatorname{div}(\tilde{b}\tilde{\varphi}) \rangle dr - \int_{[0, t]} \langle \tilde{m}, \partial_\xi \tilde{\varphi} \rangle dr.$$

This formula is valid for every  $\tilde{\varphi}$  smooth test function (with compact support), on a full measure set in  $(t, \omega)$  which can depend on  $\tilde{\varphi}$ . To make this set independent of  $\tilde{\varphi}$ , we use a density argument. Let  $D$  be a countable dense set in  $C_c^\infty(\mathbb{R}^d \times \mathbb{R})$  and let  $F$  be a full measure set in  $(t, \omega)$  satisfying: for every  $(t, \omega)$  in  $F$  and for every  $\tilde{\varphi}$  in  $D$ ,  $\tilde{m}(\omega)$  is a bounded measure and (2.5) holds. Now, for a given test function  $\tilde{\varphi}$ , we take a sequence  $(\tilde{\varphi}_n)_n$  in  $D$  converging to  $\tilde{\varphi}$  in  $C_b^2$ ; we pass to the limit in (2.5) for  $\tilde{\varphi}_n$  (using that  $\tilde{\chi}$  is bounded for every  $(t, \omega)$ ) and we get (2.5) for  $\tilde{\varphi}$  for every  $(t, \omega)$  in  $F$ . The proof of the first part is complete.

*Step 3:* Equation (2.4) and weak-\* progressive measurability imply (2.1).

Since the strategy is similar to that of the first part, we will only sketch it. We regularize  $\tilde{\chi}$  by convolving it with an approximate identity, obtaining  $\tilde{\chi}^\epsilon$ . The progressive measurability hypothesis implies that  $\tilde{\chi}^\epsilon$  is an Itô process. Therefore, for every test function  $\varphi$ , we can multiply it by  $\varphi(t, x + W_t, \xi)$  and apply Itô's

formula. By Fubini's theorem, the stochastic Fubini theorem and Remark 2.3 we can integrate in  $x$  and in  $\xi$  and exchange the order of integration. Then we bring the convolution on  $\varphi$ , let  $\epsilon$  go to 0 and change variable to get finally (2.1).  $\square$

**2.2. The case of smooth coefficients.** In this section we consider the case of a smooth coefficient  $b \in C_c^\infty(\mathbb{R}^{d+1})$  and a smooth initial condition  $u_0 \in C_c^\infty(\mathbb{R}^d)$  and derive stable a priori bounds. For simplicity of notation, we set  $R_0 = \|u_0\|_{L^\infty}$ .

**Proposition 2.7.** *Let  $u_0 \in C_c^\infty(\mathbb{R}^d)$  and  $b \in C_c^\infty(\mathbb{R}^{d+1})$ . Then there is a unique entropy solution  $u$  to (1.1). Moreover, we have*

$$(2.6) \quad \operatorname{ess\,sup}_{\omega \in \Omega} \sup_{t \in [0, T]} \|u(t)\|_{L^\infty} \leq \|u_0\|_{L^\infty}$$

and, for every  $p \geq 1$  finite,

$$(2.7) \quad \begin{aligned} \operatorname{ess\,sup}_{\omega \in \Omega} \sup_{t \in [0, T]} \|u(t)\|_{L^p}^p &+ p(p-1) \int_{[0, T]} \int \int |\xi|^{p-2} m d\xi dx dr \\ &\leq \|u_0\|_{L^p}^p + p \|u_0\|_{L^\infty}^{p-1} \|\operatorname{div} b\|_{L^1([0, T] \times \mathbb{R}^d \times [-R_0, R_0])}. \end{aligned}$$

Moreover,  $\chi$  and  $m$  are supported a.s. on  $[0, T] \times \mathbb{R}^d \times [-R_0, R_0]$ .

*Proof. Step 1:* We start with the equation

$$(2.8) \quad \partial_t v + g(t, x, v) \cdot \nabla v = 0$$

for some  $g \in C([0, T]; C_b^3(\mathbb{R}^{d+1}))$ , i.e. three times continuously differentiable with bounded derivatives, and initial condition  $u_0 \in C_c^\infty(\mathbb{R}^d)$ .

We first note that, due to the regularity of  $g$  one may rewrite (2.8) in divergence form with a force. Following Kruřkov [33, Theorem 5 and Section 5], there exists a bounded entropy solution  $v = v^g$  to (2.8). This solution can be constructed by first approximating  $g$  by a smooth  $g^\delta$  and then considering a vanishing viscosity approximation. That is,  $v$  can be obtained as an a.e. limit of the solutions  $v^{\epsilon, \delta}$  to

$$\partial_t v^{\epsilon, \delta} + g^\delta(t, x, v^{\epsilon, \delta}) \cdot \nabla v^{\epsilon, \delta} = \epsilon \Delta v^{\epsilon, \delta}.$$

The maximum principle applied to these equations yields  $\|v^{\epsilon, \delta}\|_{L_{t,x}^\infty} \leq \|u_0\|_{L_x^\infty}$ . Passing to the limit, we obtain the bound

$$(2.9) \quad \|v\|_{L_{t,x}^\infty} \leq \|u_0\|_{L_x^\infty}.$$

As a consequence of Lécureux-Mercier [36, Corollary 2.5 and Theorem 2.6], the solution  $v^g$  is in  $C([0, T]; L^1(\mathbb{R}^d))$  and is unique. Moreover, the map  $C([0, T]; C_b^3(\mathbb{R}^{d+1})) \ni g \mapsto v^g \in C([0, T]; L^1(\mathbb{R}^d))$  is locally Lipschitz continuous.

Denote by  $\chi = \chi^g$  the associated kinetic solution. The bound (2.9) implies that, for every  $t$  and for a.e.  $x$ ,  $\chi(t, x, \cdot) = \chi^g(t, x, \cdot)$  is supported on  $[-\|u_0\|_{L_x^\infty}, \|u_0\|_{L_x^\infty}] = [-R_0, R_0]$ . We have also  $\|\chi_t - \chi_s\|_{L_{x,\xi}^1} = \|v_t - v_s\|_{L_x^1}$  and  $\|\chi_t^{g^1} - \chi_t^{g^2}\|_{L_{x,\xi}^1} = \|v_t^{g^1} - v_t^{g^2}\|_{L_x^1}$ . Consequently,  $\chi$  is in  $C([0, T]; L^1(\mathbb{R}^d \times \mathbb{R}))$  and the map  $C([0, T]; C_b^3(\mathbb{R}^{d+1})) \ni g \mapsto \chi^g \in C([0, T]; L^1(\mathbb{R}^d \times \mathbb{R}))$  is locally Lipschitz continuous. As a consequence, the maps

$$\begin{aligned} [0, T] \times C([0, T]; C_b^3(\mathbb{R}^{d+1})) &\ni (t, g) \mapsto v_t^g \in L^1(\mathbb{R}^d) \\ [0, T] \times C([0, T]; C_b^3(\mathbb{R}^{d+1})) &\ni (t, g) \mapsto \chi_t^g \in L^1(\mathbb{R}^d \times \mathbb{R}) \end{aligned}$$

are continuous.

The existence of a kinetic measure  $m^g$  associated to  $\chi^g$  can be derived as by Dalibard in [14, Section 2.2] extended to the time dependent and non conservative case; that is, for every  $\varphi$  compactly supported we have

$$(2.10) \quad \int_{[0,T] \times \mathbb{R}^d \times \mathbb{R}} \partial_\xi \varphi dm^g = - \langle \chi_T^g, \varphi_T \rangle + \langle \chi_0^g, \varphi_0 \rangle \\ + \int_0^T \langle \chi^g, \partial_t \varphi + \operatorname{div}(g(r, x, \xi)\varphi) \rangle dr.$$

Therefore,  $m^g$  is uniquely determined and supported on  $[0, T] \times \mathbb{R}^d \times [-R_0, R_0]$  and (2.10) is satisfied for all smooth  $\varphi$  compactly supported in  $x$ .

In order to obtain the estimate (2.7), we consider the test functions given by  $(\operatorname{sgn}(\xi)|\xi|^{p-1})^\epsilon \psi_{1/\epsilon}(x)$ ; here  $\psi_{1/\epsilon}$  is an increasing sequence of smooth functions,  $[0, 1]$ -valued, with values 1 on  $B_{1/\epsilon}$ , 0 on  $B_{2/\epsilon}^c$  and such that  $|\nabla \psi_{1/\epsilon}(x)| \leq 2\epsilon$  for every  $x$  and  $(\operatorname{sgn}(\xi)|\xi|^{p-1})^\epsilon := \operatorname{sgn}(\cdot) \cdot |\cdot|^{p-1} *_{\xi} \bar{\rho}_\epsilon$ . In particular,  $\operatorname{sgn}(\xi)(\operatorname{sgn}(\xi)|\xi|^{p-1})^\epsilon$  is a sequence of nonnegative functions converging pointwise on  $\mathbb{R} \setminus \{0\}$  to  $|\xi|^{p-1}$ . Moreover, in the case  $p > 1$ ,  $\partial_\xi (\operatorname{sgn}(\xi)|\xi|^{p-1})^\epsilon$  converges pointwise on  $\mathbb{R}$  to  $(p-1)|\xi|^{p-2}$ , with the convention  $|0|^{p-2} = +\infty$  for  $p < 2$  and  $|0|^0 = 1$  for  $p = 2$ . Due to (2.10), we have

$$\langle \chi_t^g, (\operatorname{sgn}(\xi)|\xi|^{p-1})^\epsilon \psi_{1/\epsilon} \rangle + \int_{[0,t] \times \mathbb{R}^d \times \mathbb{R}} \frac{d}{d\xi} (\operatorname{sgn}(\xi)|\xi|^{p-1})^\epsilon \psi_{1/\epsilon} dm^g \\ = \langle \chi_0^g, (\operatorname{sgn}(\xi)|\xi|^{p-1})^\epsilon \psi_{1/\epsilon} \rangle + \int_0^t \langle \chi^g, \operatorname{div} g(r, x, \xi) (\operatorname{sgn}(\xi)|\xi|^{p-1})^\epsilon \psi_{1/\epsilon} \rangle dr \\ + \int_0^t \langle \chi^g, g(r, x, \xi) (\operatorname{sgn}(\xi)|\xi|^{p-1})^\epsilon \cdot \nabla \psi_{1/\epsilon} \rangle dr.$$

In the case  $p > 1$ , we take the  $\liminf$  for  $\epsilon \rightarrow 0$  and, recalling that  $\chi_t^g (\operatorname{sgn}(\xi)|\xi|^{p-1})^\epsilon = |\chi_t^g| \operatorname{sgn}(\xi) (\operatorname{sgn}(\xi)|\xi|^{p-1})^\epsilon$ , we apply Fatou's lemma for the second term on the left hand side and the dominated convergence theorem for the remaining terms: we get

$$\langle |\chi_t^g|, |\xi|^{p-1} \rangle + (p-1) \int_{[0,t] \times \mathbb{R}^d \times \mathbb{R}} |\xi|^{p-2} dm^g \\ \leq \langle |\chi_0^g|, |\xi|^{p-1} \rangle + \int_0^t \langle |\chi^g|, |\operatorname{div} g| |\xi|^{p-1} \rangle dr.$$

Recalling (2.2) and (2.9) we obtain

$$(2.11) \quad \|v^g\|_{L^p}^p + p(p-1) \int_{[0,t] \times \mathbb{R}^d \times \mathbb{R}} |\xi|^{p-2} dm^g \\ \leq \|u_0\|_{L^p}^p + p \int_0^t \int \int (\operatorname{div} g) |\xi|^{p-1} \mathbf{1}_{|\xi| \leq R_0} dx d\xi dr \\ \leq \|u_0\|_{L^p}^p + p \|u_0\|_{L_x^\infty}^{p-1} \|\operatorname{div} g\|_{L^1([0,T] \times \mathbb{R}^d \times [-R_0, R_0])}$$

In particular, taking  $p = 2$ , we see that  $\|m\|_{\mathcal{M}_{t,x,\xi}}$  is bounded in terms of  $u_0$  and  $\|\operatorname{div} g\|_{L^1([0,T] \times \mathbb{R}^d \times [-R_0, R_0])}$ .

In the case  $p = 1$ , proceeding as before we get

$$\|v^g\|_{L^1} + \liminf_{\epsilon \rightarrow 0} \int_{[0,t] \times \mathbb{R}^d \times \mathbb{R}} \frac{d}{d\xi} (\operatorname{sgn}(\xi))^\epsilon \psi_{1/\epsilon}(x) dm^g \\ \leq \|u_0\|_{L_x^1} + \|\operatorname{div} g\|_{L^1([0,T] \times \mathbb{R}^d \times [-R_0, R_0])}.$$

In particular, recalling again (2.2), this gives the global  $L_{x,\xi}^1$  bound

$$(2.12) \quad \sup_{t \in [0,T]} \|\chi^g(t)\|_{L_{x,\xi}^1} \leq \|u_0\|_{L_x^1} + \|\operatorname{div} g\|_{L^1([0,T] \times \mathbb{R}^d \times [-R_0, R_0])}.$$

*Step 2:* We apply the previous results to  $g = \tilde{b}^\omega = b(x + W_t(\omega), u)$  and, by Proposition 2.6, get the existence of an entropy solution with the desired estimates. The technical details are not difficult but not immediate, since we have to pass from a process with values in a space of functions of  $x$  to a measurable function of  $(t, \omega, x)$ :

1) The map  $(t, \omega) \mapsto \bar{u}(t, \omega) = v_t^{\tilde{b}^\omega}$  is measurable bounded from  $\mathcal{B}([0, T]) \otimes \mathcal{F}_T$  to  $\mathcal{B}(L_x^1)$ , for every  $T$ , since it is the composition of the measurable map  $(t, \omega) \mapsto (t, \tilde{b}^\omega)$  from  $\mathcal{B}([0, T]) \otimes \mathcal{F}_T$  to  $\mathcal{B}([0, T]) \otimes \mathcal{B}(C([0, T]; C_b^3(\mathbb{R}^{d+1})))$  and the continuous map  $(t, g^\omega) \mapsto v_t^g$ . Moreover the  $L_x^1$ -valued process  $\bar{u}$  has time-continuous paths. So  $\bar{u}$  is actually measurable bounded from  $\mathcal{P}$  (the progressive  $\sigma$ -algebra) to  $\mathcal{B}(L_x^1)$ ; in particular, it is weakly measurable with respect to  $\mathcal{P}$ . Therefore, by Proposition 5.2, there exists  $\tilde{u}$  in  $L^1([0, T] \times \Omega \times \mathbb{R}^d, \mathcal{P} \otimes \mathcal{B}(\mathbb{R}^d))$  version of  $\bar{u}$  (in the sense that, for a.e.  $(t, \omega)$ ,  $\bar{u}(t, \omega) = \tilde{u}(t, \omega)$ ). By the  $L^\infty$  bounds,  $\tilde{u}$  is in  $L_{t, \omega, x}^\infty$ .

Similarly the map  $(t, \omega) \mapsto \bar{\chi}(t, \omega) = \chi_t^{\tilde{b}^\omega}$  is measurable bounded from  $\mathcal{B}([0, T]) \otimes \mathcal{F}_T$  to  $\mathcal{B}(L_{x, \xi}^1)$ , for every  $T$ , and time-continuous and thus weakly measurable with respect to  $\mathcal{P}$ . Setting  $\tilde{\chi} = 1_{\bar{u} < \xi} - 1_{0 < \xi}$ ,  $\tilde{\chi}$  is a  $\mathcal{P} \otimes \mathcal{B}(\mathbb{R}^d) \otimes \mathcal{B}(\mathbb{R})$ -measurable version of  $\bar{\chi}$ , it is in  $L^1$  and supported on  $[-R_0, R_0]$ .

2) The map  $\omega \mapsto \bar{m}(\omega) = m^{\tilde{b}^\omega}$  is bounded as an  $\mathcal{M}_{t, x, \xi}$ -valued function, nonnegative and supported on  $[0, T] \times \mathbb{R}^d \times [-R_0, R_0]$ . Concerning progressive measurability, for  $\psi$  in  $C_c^\infty([0, T] \times \mathbb{R}^d \times \mathbb{R})$ , call  $\varphi$  a primitive function of  $\psi$ . Then  $\int_{[0, T] \times \mathbb{R}^d \times \mathbb{R}} \psi d\bar{m} = \int \partial_\xi \varphi d\bar{m}$  is  $\mathcal{F}_T$ -measurable for every  $T$  by equation (2.10) with  $g = \tilde{b}^\omega$  (where the right-hand side is  $\mathcal{F}_T$ -measurable), moreover it has càdlàg paths; hence it is progressively measurable. By equation (2.10),  $\tilde{\chi}$  and  $\bar{m}$  satisfy equation (2.4).

3) We call  $u(t, \omega, x) = \tilde{u}(t, \omega, x - W_t(\omega))$ ,  $\chi(t, \omega, x, \xi) = \tilde{\chi}(t, \omega, x - W_t(\omega), \xi)$  and  $m(t, \omega, x, \xi) = \bar{m}(t, \omega, x - W_t(\omega), \xi)$ : more precisely, for  $\omega$  fixed, we define  $m(\omega)$  as the image measure of  $\bar{m}(\omega)$  under  $(t, x, \xi) \mapsto (t, x + W_t(\omega), \xi)$ . Then, by Proposition 2.6,  $u$  and  $\chi$  are also  $L_{x, \xi}^\infty$ -valued weakly-\* progressively measurable,  $m$  is a kinetic measure and  $u$  is an entropy solution of (1.1), with kinetic function  $\chi$  and kinetic measure  $m$ .

4) Changing variable  $x' = x - W_t$  in (2.11) and in (2.12), we get the estimates (2.6) and (2.7).  $\square$

**2.3. Existence of generalized kinetic solutions.** We introduce the notion of a generalized kinetic solution.

**Definition 2.8.** Let  $f_0 \in (L^1 \cap L^\infty)(\mathbb{R}^d \times \mathbb{R})$ . A generalized kinetic solution to (1.1) is a measurable function  $f : [0, T] \times \Omega \times \mathbb{R}^d \times \mathbb{R} \rightarrow \mathbb{R}$  with the following properties:

- (i)  $f \in L^\infty([0, T] \times \Omega; L^1(\mathbb{R}^d \times \mathbb{R}))$  and is supported on  $[0, T] \times \Omega \times \mathbb{R}^d \times [-R, R]$  for some  $R > 0$ ;
- (ii)  $f$  is a weakly-\* progressively measurable  $L_{x, \xi}^\infty$ -valued process;
- (iii) there exists a kinetic bounded measure  $m$  on  $[0, T] \times \Omega \times \mathbb{R}^d \times \mathbb{R}$  satisfying: for every  $\varphi \in C_c^\infty([0, T] \times \mathbb{R}^d \times \mathbb{R})$ , it holds for a.e.  $(t, \omega)$ ,

$$(2.13) \quad \begin{aligned} \langle f_t, \varphi_t \rangle &= \langle f_0, \varphi_0 \rangle + \int_0^t \langle f, \partial_t \varphi + \operatorname{div}(b(x, \xi)\varphi) \rangle dr + \int_0^t \langle f, \nabla \varphi \rangle dW \\ &\quad + \frac{1}{2} \int_0^t \langle f, \Delta \varphi \rangle dr - \int_{[0, t] \times \mathbb{R}^d \times \mathbb{R}} \partial_\xi \varphi dm. \end{aligned}$$

- (iv) there exists a kinetic bounded measure  $\nu$  on  $[0, T] \times \Omega \times \mathbb{R}^d \times \mathbb{R}$ , which moreover is in  $L^\infty([0, T] \times \Omega, \mathcal{M}(\mathbb{R}^d \times \mathbb{R}))$ , satisfying: for every  $\varphi \in$

$$\begin{aligned}
& C_c^\infty([0, T] \times \mathbb{R}^d \times \mathbb{R}), \text{ it holds for a.e. } (t, \omega), \\
(2.14) \quad & |f(t, x, \xi)| = \text{sgn}(\xi) f(t, x, \xi) \leq 1, \quad \text{for a.e. } (x, \xi), \\
& \langle f_t, -\partial_\xi \varphi_t \rangle = \int \varphi(t, x, 0) dx - \int_{\mathbb{R}^d \times \mathbb{R}} \varphi_t d\nu_t.
\end{aligned}$$

A formal, short-hand notation for (2.13) and (2.14) is

$$\partial_t f + b(x, \xi) \cdot \nabla f + \nabla f \circ dW_t = \partial_\xi m,$$

and

$$\begin{aligned}
|f|(t, x, \xi) &= \text{sgn}(\xi) f(t, x, \xi) \leq 1, \\
\frac{\partial f}{\partial \xi} &= \delta(\xi) - \nu(t, x, \xi).
\end{aligned}$$

*Remark 2.9.* Kinetic solutions are a particular type of generalized kinetic solutions. Indeed, if  $f_0(x, \xi) := \chi(u_0(x), \xi)$  and  $\chi$  is a kinetic solution to (1.1), with associated kinetic measure  $m$ , then  $\chi$  is also a generalized solution with kinetic measure  $m$  and  $\nu = \delta_{\xi=u(t, \omega, x)}$ .

The following theorem asserts the existence of a generalized kinetic solution.

**Theorem 2.10.** *Let  $b \in L_{loc}^1(\mathbb{R}^{d+1})$  with  $\text{div } b$  in  $L_{\xi, loc}^1(L_x^1)$  and  $u_0 \in (L^1 \cap L^\infty)(\mathbb{R}^d)$ . Then there exists a generalized kinetic solution  $f$  to (1.1) starting from  $f_0(x, \xi) := \chi(u_0(x), \xi)$ .*

*Proof. Step 1: Approximation of  $f$  and convergence.*

We introduce smooth approximations:  $b^\varepsilon \in C_c^\infty(\mathbb{R}^{d+1})$  of  $b$  with  $b^\varepsilon \rightarrow b$  in  $L_{x, \xi, loc}^1$  and  $\text{div } b^\varepsilon \rightarrow \text{div } b$  in  $L_{\xi, loc}^1(L_x^1)$ ;  $u_0^\varepsilon \in C_c^\infty(\mathbb{R}^d)$  of  $u_0$  with  $\|u_0^\varepsilon\|_{L^p} \leq \|u_0\|_{L^p}$  for all  $p \geq 1$  and  $u_0^\varepsilon \rightarrow u_0$  in  $L_x^1$ . We consider the corresponding unique entropy solution  $u^\varepsilon$  (see Proposition (2.7)) to

$$\partial_t u^\varepsilon(t, x) + b^\varepsilon(x, u^\varepsilon(t, x)) \cdot \nabla u^\varepsilon(t, x) + \nabla u^\varepsilon(t, x) \circ dW_t = 0;$$

that is  $\chi^\varepsilon = \chi(u^\varepsilon)$  solves

$$(2.15) \quad \partial_t \chi^\varepsilon + b^\varepsilon(x, \xi) \cdot \nabla \chi^\varepsilon + \nabla \chi^\varepsilon \circ dW_t = \partial_\xi m^\varepsilon.$$

Since  $|\chi^\varepsilon| \leq 1$  and  $\chi^\varepsilon$  are  $\mathcal{P} \otimes \mathcal{B}(\mathbb{R}^{d+1})$ -measurable, the sequence  $\chi^\varepsilon$  converges weakly-\*, up to taking a subsequence, to a limit  $f$  in  $L^\infty([0, T] \times \Omega \times \mathbb{R}^d \times \mathbb{R}, \mathcal{P} \otimes \mathcal{B}(\mathbb{R}^{d+1}))$ . In particular,  $f$  is weakly-\* progressively measurable as an  $L_{x, \xi}^\infty$ -valued process.

Note that the sequence  $\chi^\varepsilon$ , as  $\mathcal{B}([0, T]) \otimes \mathcal{A} \otimes \mathcal{B}(\mathbb{R}^{d+1})$ -measurable processes, is also weakly-\* compact in  $L^\infty([0, T] \times \Omega \times \mathbb{R}^d \times \mathbb{R}, \mathcal{B}([0, T]) \otimes \mathcal{A} \otimes \mathcal{B}(\mathbb{R}^{d+1}))$ . Therefore, up to taking a sub-subsequence, we can assume that  $\chi^\varepsilon$  converges weakly-\* to  $f$  also in  $L^\infty([0, T] \times \Omega \times \mathbb{R}^d \times \mathbb{R}, \mathcal{B}([0, T]) \otimes \mathcal{A} \otimes \mathcal{B}(\mathbb{R}^{d+1}))$ . In particular, we can allow test functions of the form  $F(\omega)\varphi(t, x, \xi)$ .

*Step 2: Bounds and support of  $f$ .*

Using Proposition 2.7 and  $\text{div } b \in L_{\xi, loc}^1(L_x^1)$ , we obtain that  $(\chi^\varepsilon)^+ = \chi^\varepsilon \vee 0$  is uniformly bounded in  $L_{t, \omega}^\infty(L_{x, \xi}^1)$ . Therefore, identifying  $\chi^\varepsilon(x, \xi)$  with  $\chi^\varepsilon(x, \xi) dx d\xi$ ,  $(\chi^\varepsilon)^+$  is uniformly bounded in  $L_{t, \omega}^\infty(\mathcal{M}_{x, \xi, +})$ . By Theorem 5.3, up to the selection of a subsequence,  $(\chi^\varepsilon)^+$  converges weakly-\* in  $L_{t, \omega}^\infty(\mathcal{M}_{x, \xi, +})$  to an element  $g^+ \in L_{t, \omega}^\infty(\mathcal{M}_{x, \xi, +})$ . Similarly  $(\chi^\varepsilon)^- = (-\chi^\varepsilon) \vee 0$  converges weakly-\*, up to the selection of a subsequence, to an element  $g^- \in L_{t, \omega}^\infty(\mathcal{M}_{x, \xi, +})$ . Moreover, we can take the same subsequence for the weakly-\* convergence of  $\chi^\varepsilon$  in  $L_{t, \omega, x, \xi}^\infty$  and of  $(\chi^\varepsilon)^+$  and

$(\chi^\epsilon)^-$  in  $L_{t,\omega}^\infty(\mathcal{M}_{x,\xi,+})$ . By a density argument (on the test functions), we see that  $g := g^+ - g^- = f$ . In particular,

$$\begin{aligned} \|f\|_{L_{t,\omega}^\infty(L_{x,\xi}^1)} &= \|g\|_{L_{t,\omega}^\infty(\mathcal{M}_{x,\xi})} \leq \|g^+\|_{L_{t,\omega}^\infty(\mathcal{M}_{x,\xi})} + \|g^-\|_{L_{t,\omega}^\infty(\mathcal{M}_{x,\xi})} \\ &\leq 2 \sup_\epsilon \|\chi^\epsilon\|_{L_{t,\omega}^\infty(L_{x,\xi}^1)}. \end{aligned}$$

For the support property of  $f$ , again by Proposition 2.7 the functions  $\chi^\epsilon$  are concentrated a.s. on  $[0, T] \times \mathbb{R}^d \times [-\|u_0\|_\infty, \|u_0\|_\infty]$ . Therefore,

$$\mathbb{E}[F\langle\chi^\epsilon, \varphi\rangle_{t,x,\xi}] = 0$$

for every  $\varphi$  in  $L^1([0, T] \times \mathbb{R}^d \times \mathbb{R})$  with support outside  $[0, T] \times \mathbb{R}^d \times [-\|u_0\|_\infty, \|u_0\|_\infty]$  and every  $F$  in  $L_\omega^1$ . Passing to the limit in the above equality, we conclude that  $f$  is concentrated a.s. on  $[0, T] \times \mathbb{R}^d \times [-\|u_0\|_\infty, \|u_0\|_\infty]$ .

*Step 3: Convergence of  $m^\epsilon$ .*

By Proposition 2.7 (applied with  $p = 2$ ),  $m^\epsilon$  is a bounded sequence in the space  $L^\infty(\Omega; \mathcal{M}_+([0, T] \times \mathbb{R}^d \times \mathbb{R}))$ . Therefore, by Theorem 5.3, it converges weakly-\*, up to subsequences, to a limit  $m$  in  $L^\infty(\Omega; \mathcal{M}_+([0, T] \times \mathbb{R}^d \times \mathbb{R}))$ . The support property of  $m$  follows from Proposition 2.7 as for  $f$ , replacing  $L^1([0, T] \times \mathbb{R}^d \times \mathbb{R})$  with  $C_0([0, T] \times \mathbb{R}^d \times \mathbb{R})$ .

Concerning progressive measurability of  $m$ , we have to prove that, for every  $\varphi$  in  $C_c^\infty([0, T] \times \mathbb{R}^d \times \mathbb{R})$ , the process  $(\langle m, 1_{[0,t]} \varphi \rangle)_t$  is  $\mathcal{P}$ -measurable. By right-continuity of the process and of the filtration, it is enough to show that, for every  $t_0$ , for every positive integer  $n$ , the random variable  $\langle m, \varphi^{t_0, n} \rangle$  is  $\mathcal{F}_{t_0+1/n}$ -measurable, where  $\varphi^{t_0, n} = \varphi 1_{[0, t_0]} * 1_{[0, 1/n]}$ . The functions  $\langle m^\epsilon, \varphi^{t_0, n} \rangle$  are  $\mathcal{F}_{t_0+1/n}$ -measurable and, by continuity of  $\varphi^{t_0, n}$ , converge to  $\langle m, \varphi^{t_0, n} \rangle$  weakly-\* in  $L_\omega^\infty$ , in particular weakly in  $L_\omega^2$ . Now the space of  $\mathcal{F}_{t_0+1/n}$ -measurable  $L_\omega^2$  functions is (isomorphic to) a closed (and thus weakly closed) subspace of  $L_\omega^2$ . Hence,  $\langle m, \varphi^{t_0, n} \rangle$  is  $\mathcal{F}_{t_0+1/n}$ -measurable (precisely,  $\mathcal{F}_{t_0+1/n}$ -measurable up to  $P$ -null sets which implies  $\mathcal{F}_{t_0+1/n}$ -measurability by completeness of  $\mathcal{F}_0$ ). This shows that  $m$  is a kinetic measure.

*Step 4: Equation (2.13).*

Equation (2.13) is obtained passing to the limit in (2.15) for  $\varphi$  in  $C_c^\infty([0, T] \times \mathbb{R}^d \times \mathbb{R})$ , exploiting the linearity of the equation. More precisely, we multiply (2.15) by a measurable bounded function  $G = G(t, \omega)$ , we integrate in  $t$  and  $\omega$ ; we can then pass to the limit as  $\epsilon \rightarrow 0$ , thanks to the weak-\* convergence of  $\chi^\epsilon$  and  $m^\epsilon$  and the fact that  $b \cdot \nabla \varphi$  and  $\varphi \operatorname{div} b$  are in  $L_{t,\omega,x,\xi}^1$ . By arbitrariness of  $G$  we get (2.13).

*Step 5: Properties (2.14).*

The bound  $\|f\|_{L_{t,\omega,x,\xi}^\infty} \leq 1$  follows from the same bound for  $\chi^\epsilon$ . For the property  $|f| = \operatorname{sgn}(\xi)f$ , we notice that  $\mathbb{E}[\langle \chi^\epsilon, \operatorname{sgn}(\xi)G \rangle_{t,x,\xi}] \geq 0$  for every  $G$  nonnegative function in  $L^1([0, T] \times \Omega \times \mathbb{R}^d \times \mathbb{R})$  and we pass to the limit as  $\epsilon \rightarrow 0$ , getting  $|f| = \operatorname{sgn}(\xi)f$ .

Further, for  $\nu^\epsilon$ , we have for a.e.  $(t, \omega)$ ,

$$(2.16) \quad \langle \chi_t^\epsilon, -\partial_\xi \varphi_t \rangle = \int_{\mathbb{R}^d} \varphi(t, x, 0) dx - \int_{\mathbb{R}^d \times \mathbb{R}} \varphi_t d\nu_t^\epsilon,$$

where  $d\nu^\epsilon = \delta_{\xi=u^\epsilon(t,x)} dx dt$ . In particular,  $\nu^\epsilon$  is a bounded sequence in  $L^\infty(\Omega \times [0, T]; \mathcal{M}(\mathbb{R}^d \times \mathbb{R}))$  of kinetic measures. Proceeding as for  $m^\epsilon$ , we get that  $\nu^\epsilon$  converges weakly-\*, up to subsequences, to a bounded kinetic measure  $\nu$ . We then pass to the limit in (2.16) in a way similar to the proof of equation (2.13) and we obtain (2.14).  $\square$



*Remark 2.11.* For any generalized kinetic solution  $f$ , which by definition is in  $L^\infty(\Omega \times [0, T]; L^1(\mathbb{R}^d \times \mathbb{R})) \cap L^\infty(\Omega \times [0, T]; L^\infty(\mathbb{R}^d \times \mathbb{R}))$ , we have by interpolation  $f \in L^\infty(\Omega \times [0, T]; L^p(\mathbb{R}^d \times \mathbb{R}))$  for every  $1 \leq p \leq \infty$ . Moreover the global  $L^1_{x,\xi}$  bound allows to consider also bounded test functions, independent of  $\xi$ , which are in  $L^\infty([0, T]; W^{2,\infty}(\mathbb{R}^d)) \cap L^\infty(\mathbb{R}^d; W^{1,\infty}([0, T]))$ .

The following lemma will be useful in the next section.

**Lemma 2.12.** *Let  $f$  be a generalized kinetic solution to equation (1.1). For every test function  $\psi$  in  $C_c^\infty(\mathbb{R}^d \times \mathbb{R})$ , there exist measurable functions  $f(\psi)^+$ ,  $f(\psi)^-$  on  $[0, T] \times \Omega \times \mathbb{R}^d \times \mathbb{R}$ , such that:*

- (i)  $f(\psi)^+$ ,  $f(\psi)^-$  are versions of  $f *_{x,\xi} \psi$  (that is, for every  $(x, \xi)$ ,  $f(\psi)^+(x, \xi)$  and  $f(\psi)^-(x, \xi)$  coincide with  $f *_{x,\xi} \psi$  on a full-measure set in  $[0, T] \times \Omega$ , possibly depending on  $(x, \xi)$  and  $\psi$ );
- (ii) for every  $(x, \xi)$ ,  $f(\psi)^+(x, \xi)$ ,  $f(\psi)^-(x, \xi)$  are progressively measurable processes;
- (iii) for a.e.  $\omega$  it holds: for every  $(x, \xi)$ ,  $f(\psi)^+(x, \xi)$  is càdlàg,  $f(\psi)^-(x, \xi)$  is càglàd;
- (iv) for a.e.  $\omega$  it holds: for every  $t$ ,  $f(\psi)^+$  is  $C^1_{x,\xi}$  and  $\nabla_{x,\xi} f(\psi)^+ = f(\nabla_{x,\xi} \psi)^+$  and similarly for  $f(\psi)^-$ .

The above lemma is similar to Remark 2.3 but with the additional property (iv). The existence of such versions is needed when dealing with terms of the form  $\int \partial_\xi f(\psi)^+ dm$ , since for these terms both the precise version in time and the differentiability in  $\xi$  are needed. Note that properties (ii) and (iv) imply  $\mathcal{P} \otimes \mathcal{B}(\mathbb{R}^{d+1})$ -measurability.

*Proof of Lemma 2.12.* We call  $\varphi^{x,\xi}(y, \zeta) = \psi(x - y, \xi - \zeta)$ . We know that, for every  $(x, \xi)$ , it holds for a.e.  $(t, \omega)$ ,

$$(2.17) \quad \begin{aligned} \langle f_t, \varphi_t^{x,\xi} \rangle &= \langle f_0, \varphi_0^{x,\xi} \rangle + \int_0^t \langle f, \partial_t \varphi^{x,\xi} + \operatorname{div}_y(b\varphi^{x,\xi}) \rangle dr + \int_0^t \langle f, \nabla_y \varphi^{x,\xi} \rangle dW \\ &+ \frac{1}{2} \int_0^t \langle f, \Delta_y \varphi^{x,\xi} \rangle dr - \int_{[0,t] \times \mathbb{R}^d \times \mathbb{R}} \partial_\zeta \varphi^{x,\xi} dm. \end{aligned}$$

For the integrals  $\int_0^t \langle f, \partial_t \varphi^{x,\xi} + \operatorname{div}_y(b\varphi^{x,\xi}) \rangle dr$ ,  $\frac{1}{2} \int_0^t \langle f, \Delta_y \varphi^{x,\xi} \rangle dr$ , there exist resp. versions  $A(t, \omega, x, \xi)$ ,  $B(t, \omega, x, \xi)$  which satisfy the second and the fourth property above and are continuous (a.s.) in  $(t, x, \xi)$  (these versions are simply the Lebesgue integrals for a fixed version of  $f$ ). Such a version  $C(t, \omega, x, \xi)$  exists also for the stochastic integral  $-\int_0^t \langle f, \nabla_y \varphi^{x,\xi} \rangle dW$ , by Theorem 10.6 in Kunita [35, Chapter 1]. Finally, for  $-\int_{[0,t] \times \mathbb{R}^d \times \mathbb{R}} \partial_\zeta \varphi^{x,\xi} dm$ , by Remark 2.3 there exist versions  $D^+(t, \omega, x, \xi)$ ,  $D^-(t, \omega, x, \xi)$  which satisfy the second and the fourth property above and are resp. càdlàg, càglàd for fixed  $(x, \xi)$  (these versions are simply the Lebesgue integrals resp. on  $[0, t] \times \mathbb{R}^d \times \mathbb{R}$  and  $[0, t[ \times \mathbb{R}^d \times \mathbb{R}$  for a fixed version of  $f$ ). Therefore  $f(\psi)^+ = \langle f_0, \varphi_0^{x,\xi} \rangle + A + B + C + D^+$  and  $f(\psi)^- = \langle f_0, \varphi_0^{x,\xi} \rangle + A + B + C + D^-$  are measurable versions of  $f *_{x,\xi} \psi$  with the desired properties.  $\square$

From now on, when this does not create confusion, the first three integrals in formula (2.17) will denote their continuous versions. The càdlàg version  $D^+$  of the last integral will be denoted still by  $-\int_{[0,t] \times \mathbb{R}^d \times \mathbb{R}} \partial_\zeta \varphi^{x,\xi} dm$ , while the càglàd version  $D^-$  by  $-\int_{[0,t[ \times \mathbb{R}^d \times \mathbb{R}} \partial_\zeta \varphi^{x,\xi} dm$ , coherently with the continuity property in  $t$  of the integral on  $[0, t]$ .

*Remark 2.13.* Consider  $f^{\epsilon,\delta} = f *_{x,\xi}(\rho_\epsilon \bar{\rho}_\delta)$ , where  $\rho = \rho(x)$ ,  $\bar{\rho} = \bar{\rho}(\xi)$  are two  $C_c^\infty$  even functions and  $\rho_\epsilon(x) = \epsilon^{-d} \rho(\epsilon^{-1}x)$ ,  $\bar{\rho}_\delta(\xi) = \delta^{-1} \bar{\rho}(\delta^{-1}\xi)$ . We call  $f^{\epsilon,\delta,+}$ ,  $f^{\epsilon,\delta,-}$  the versions of  $f^{\epsilon,\delta}$  as in the previous Lemma. Note that, by construction, for a.e.  $\omega$ , it holds for every  $(t, x, \xi)$  (with the above convention on the integrals),

$$\begin{aligned} f_t^{\epsilon,\delta,+}(x, \xi) &= f_0(x, \xi) + \int_0^t \frac{1}{2} \Delta f^{\epsilon,\delta}(x, \xi) dr - \int_0^t (b \cdot \nabla f)^{\epsilon,\delta}(x, \xi) dr \\ &\quad - \int_0^t \nabla f^{\epsilon,\delta}(x, \xi) dW_r + \int_{[0,t] \times \mathbb{R}^d \times \mathbb{R}} \rho^\epsilon(x-y) (\bar{\rho}^\delta)'(\xi-\zeta) m(r, y, \zeta) dy d\zeta dr, \end{aligned}$$

where  $(b \cdot \nabla f)^{\epsilon,\delta} = \nabla(\rho^\epsilon \bar{\rho}^\delta) *_{x,\xi}(bf) + (\rho^\epsilon \bar{\rho}^\delta) *_{x,\xi}((\operatorname{div} b)f)$ . The integrands  $\int_{\mathbb{R}^d \times \mathbb{R}} \rho^\epsilon(x-y) \bar{\rho}^\delta(\xi-\zeta) m(\cdot, y, \zeta) dy d\zeta$ ,  $\int_{\mathbb{R}^d \times \mathbb{R}} \rho^\epsilon(x-y) (\bar{\rho}^\delta)'(\xi-\zeta) m(\cdot, y, \zeta) dy d\zeta$  will be denoted resp. by  $m^{\epsilon,\delta}(\cdot, x, \xi)$  and  $\partial_\xi m^{\epsilon,\delta}(\cdot, x, \xi)$ ; they are measures on  $[0, T]$  parametrized by  $(\omega, x, \xi)$ .

Moreover, for every fixed representative of  $m$  and every test function  $\psi$ , it holds:

- 1) the function  $(t, \omega, x, \xi) \mapsto \int_{[0,t]} f(\psi)^+(x, \xi) \partial_\xi m^{\epsilon,\delta}(r, x, \xi) dr$  is measurable in  $(t, \omega, x, \xi)$ , càdlàg in  $t$  and continuous in  $(x, \xi)$  for a.e.  $\omega$ ;
- 2) for a.e.  $\omega$ , we have for every  $t \geq 0$ ,

$$\begin{aligned} &\int_{\mathbb{R}^d \times \mathbb{R}} \int_{[0,t]} f(\psi)^+(r, x, \xi) \partial_\xi m^{\epsilon,\delta}(r, x, \xi) dr dx d\xi \\ &= \int_{[0,t] \times \mathbb{R}^d \times \mathbb{R}} \int_{\mathbb{R}^d \times \mathbb{R}} f(\psi)^+(r, x, \xi) \rho^\epsilon(x-y) (\bar{\rho}^\delta)'(\xi-\zeta) dx d\xi m(r, y, \zeta) dy d\zeta dr \\ &= - \int_{[0,t] \times \mathbb{R}^d \times \mathbb{R}} \partial_\xi f(\psi)^+(r, x, \xi) m^{\epsilon,\delta}(r, x, \xi) dx d\xi dr. \end{aligned}$$

Indeed, the measurability follows from Remark 5.4 below applied at  $(x, \xi)$  fixed and from the continuity property of the integral with respect to  $(x, \xi)$  (at  $(t, \omega)$  fixed). The above equality follows from Fubini's theorem, Lemma 2.12 and the càdlàg property of the integrals. An analogous property holds replacing  $f(\psi)^+$  with  $f(\psi)^+ \varphi$  or  $f(\psi)^- \varphi$  for regular test functions  $\varphi$ .

### 3. WELL-POSEDNESS OF ENTROPY SOLUTIONS

In this section we prove the well-posedness by noise result, namely the existence, uniqueness and stability of entropy solutions:

**Theorem 3.1.** *Assume that  $b \in L_{\xi,loc}^\infty(L_x^\infty) \cap L_{\xi,loc}^1(W_{x,loc}^{1,1})$  and that  $\operatorname{div} b \in L_{\xi,loc}^1(L_x^1) \cap L_{\xi,loc}^p(L_{\xi,loc}^\infty)$  for some  $p > d$ ,  $p \leq \infty$ . For every initial datum  $u_0$  in  $(L^1 \cap L^\infty)(\mathbb{R}^d)$ , there exists a unique entropy solution  $u$  to (1.1). Moreover, for every initial data  $u_0^1, u_0^2$  in  $(L^1 \cap L^\infty)(\mathbb{R}^d)$ , the two corresponding entropy solutions  $u^1, u^2$  satisfy*

$$\mathbb{E} \int |u_t^1 - u_t^2| dx \leq C \int |u_0^1 - u_0^2| dx,$$

for a.e.  $t \in [0, T]$  and some constant  $C > 0$ , depending only on  $T$ ,  $\|b\|_{L_{\xi,[-M,M]}^\infty(L_x^\infty)}$  and  $\|\operatorname{div} b\|_{L_x^p(L_{\xi,[-M,M]}^\infty)}$ , where  $M = \max\{\|u_0^1\|_{L_x^\infty}, \|u_0^2\|_{L_x^\infty}\}$ .

*Remark 3.2.* As it will be clear from the proof, the result can be generalized to fluxes with  $b(x, u)$  replaced by

$$\sum_{k=1}^N b_k(x, u),$$

where  $b_k$  satisfy the assumptions of Theorem 3.1 with integrability exponents  $p_k > d$  (i.e.  $\operatorname{div} b_k \in L_x^{p_k}(L_{\xi,loc}^\infty)$ ) which can vary with  $k$ .

Another generalization concerns the condition  $\operatorname{div} b \in L_{\xi,loc}^1(L_x^1) \cap L_x^p(L_{\xi,loc}^\infty)$ , which can be relaxed to  $\operatorname{div} b \in L_{x,\xi,loc}^1$ ,  $(\operatorname{div} b)_+ \in L_{\xi,loc}^1(L_x^1) \cap L_x^p(L_{\xi,loc}^\infty)$ . Indeed, only the bound on the positive part of  $\operatorname{div} b$  is required for the a priori estimates in the proof of the existence of generalized solutions as well as in the proof of uniqueness.

The proof of Theorem 3.1 follows from the following two preliminary results, the key estimate being the following

**Lemma 3.3.** *Assume that  $b \in L_{\xi,loc}^\infty(L_x^\infty) \cap L_{\xi,loc}^1(W_{x,loc}^{1,1})$  and that  $\operatorname{div} b \in L_x^p(L_{\xi,loc}^\infty)$  for some  $p > d$ ,  $p \leq \infty$ . Let  $f$  be a generalized kinetic solution to (1.1), supported on  $[0, T] \times \Omega \times \mathbb{R}^d \times [-R, R]$  for some  $R \geq 0$ . Then,*

$$(3.1) \quad \mathbb{E} \int_{\mathbb{R}^d \times \mathbb{R}} (|f_t| - f_t^2) d\xi dx \leq C \int_{\mathbb{R}^d \times \mathbb{R}} (|f_0| - f_0^2) d\xi dx,$$

for a.e.  $t \in [0, T]$  and some constant  $C > 0$ , depending only on  $T$ ,  $\|b\|_{L_{\xi,[-R,R]}^\infty(L_x^\infty)}$  and  $\|\operatorname{div} b\|_{L_x^p(L_{\xi,[-R,R]}^\infty)}$ .

Note that  $|f_t| - f_t^2 \geq 0$  for any generalized kinetic solution, since  $|f| \leq 1$  by definition. When the initial datum  $f_0$  is the kinetic function of some  $u_0$ , that is, if  $f_0(x, \xi) = \chi(u_0)(x, \xi)$ , then Lemma 3.3 implies that  $f$  takes values in  $\{0, \pm 1\}$ . In this case  $f$  is a true kinetic function:

**Proposition 3.4.** *Assume that  $b$  satisfies the assumptions of Lemma 3.3 and let  $f$  be a generalized kinetic solution to (1.1) starting from  $f_0 = \chi(u_0)$ , for some  $u_0$  in  $(L^1 \cap L^\infty)(\mathbb{R}^d)$ . Then there exists an entropy solution  $u$  to (1.1) such that  $f(x, \xi, t) = \chi(\xi, u(x, t))$  a.e. in  $(t, \omega, x, \xi)$ .*

Lemma 3.3 and Proposition 3.4, together with Theorem 2.10, imply the well-posedness result Theorem 3.1:

*Proof of Theorem 3.1.* Concerning the existence of an entropy solution, Theorem 2.10 yields the existence of a generalized kinetic solution  $f$  to (1.1). Proposition 3.4 then implies the existence of an entropy solution to (1.1).

For stability, let  $\chi^i = \chi(u^i, \xi)$  be the kinetic functions associated to  $u^i$ ,  $i = 1, 2$ . Note that  $|\chi^1 - \chi^2|^2 = |\chi^1 - \chi^2| = 1_{u^1 \leq \xi < u^2} + 1_{u^2 \leq \xi < u^1}$  for a.e.  $\xi$  and, in particular,  $\int |\chi_t^1 - \chi_t^2|^2 d\xi = |u_t^1 - u_t^2|$ . Therefore, the statement is equivalent to

$$(3.2) \quad \mathbb{E} \int |\chi_t^1 - \chi_t^2|^2 dx d\xi \leq C \int |\chi_0^1 - \chi_0^2|^2 dx d\xi.$$

Now consider  $f := \frac{1}{2}(\chi^1 + \chi^2)$ . Then  $f$  is a generalized kinetic solution, with associated Young measure  $\nu = \delta_0 - \frac{1}{2}(\delta_{\xi=u^1} + \delta_{\xi=u^2})$ . Moreover,

$$\begin{aligned} |f| - f^2 &= \frac{1}{2} \operatorname{sgn}(\xi)(\chi^1 + \chi^2) - \frac{1}{4}((\chi^1)^2 + (\chi^2)^2 + 2\chi^1\chi^2) \\ &= \frac{1}{2}(|\chi^1| + |\chi^2|) - \frac{1}{4}(|\chi^1| + |\chi^2| + 2\chi^1\chi^2) \\ &= \frac{1}{4}|\chi^1 - \chi^2|^2. \end{aligned}$$

Therefore, Lemma 3.3 implies (3.2). Uniqueness follows from stability, thus, the proof is complete.  $\square$

In order to prove Lemma 3.3, we will use the equations (more precisely, certain inequalities) satisfied by  $|f|$  and  $f^2$ . We recall that, since  $f$  satisfies a transport-type equation, for any function  $\beta$  regular enough, informally  $\beta(f)$  also satisfies a transport-type equation. This property is known as renormalization. When coming to a rigorous proof, however, problems can appear from the drift term, when  $b$  is not regular enough, and from the kinetic measure term  $m$ . The Sobolev assumption on  $b$ , as in the theory of DiPerna, Lions [17] and Ambrosio [1], ensures that the drift term behaves nicely. The presence of the kinetic measure  $m$  does not allow to write an equation for  $|f|$  and  $f^2$  themselves but is enough for the following inequality:

**Lemma 3.5.** *Assume that  $b \in L^1_{\xi,loc}(W^{1,1}_{x,loc})$ . Let  $f$  be a generalized kinetic solution to (1.1). Then, for every nonnegative test function  $\varphi$  in  $C_c^\infty([0, T] \times \mathbb{R}^d)$  independent of  $\xi$ , it holds for a.e.  $(t, \omega)$ ,*

$$\begin{aligned} \int_{\mathbb{R}^d \times \mathbb{R}} (|f_t| - f_t^2) \varphi_t dx d\xi &\leq \int_0^t \int_{\mathbb{R}^d \times \mathbb{R}} [\partial_t \varphi + \frac{1}{2} \Delta \varphi + \operatorname{div}(b\varphi)] (|f| - f^2) dx d\xi dr \\ &\quad + \int_0^t \int_{\mathbb{R}^d \times \mathbb{R}} \nabla \varphi (|f| - f^2) dx d\xi dW_r. \end{aligned}$$

For the proof of this Lemma we need the following commutator lemma. Recall that  $f^{\epsilon, \delta} = f *_{x, \xi} (\rho^\epsilon \bar{\rho}^\delta)$ .

**Lemma 3.6.** *Assume that  $b \in L^1_{\xi,loc}(W^{1,1}_{x,loc})$ . Then it holds, for every finite  $m \geq 1$ ,*

$$\begin{aligned} \lim_{\epsilon \rightarrow 0} \lim_{\delta \rightarrow 0} \mathbb{E} \int_0^T & \left| \int_{\mathbb{R}^d \times \mathbb{R}} \int_{\mathbb{R}^d \times \mathbb{R}} f_r^{\epsilon, \delta}(x, \xi) f_r(y, \zeta) \bar{\rho}_\delta(\xi - \zeta) \left( \nabla \rho_\epsilon(x - y) \cdot (b(x, \xi) - b(y, \zeta)) \right. \right. \\ & \left. \left. + \rho_\epsilon(x - y) \operatorname{div}_y b(y, \zeta) \right) \varphi_r(x) dy d\zeta dx d\xi \right|^m dr \\ &= 0. \end{aligned}$$

*Remark 3.7.* This is the only point where we need the Sobolev assumption on  $b$ . One may note that, in Ambrosio [1], the renormalization property for the linear transport equation is proved only assuming  $BV$  regularity for  $b$ , roughly speaking by showing the above commutator estimate for a carefully chosen kernel  $\rho$ . One may expect that this strategy also works here, but we do not investigate this issue any further.

*Proof of Lemma 3.6.* The proof is obtained by adapting the classical commutator lemma (see for example [1, 17]) to this anisotropic regularization in  $x$  and  $\xi$ , which was also used by Chen, Perthame in [11]. Since  $b$  is weakly differentiable in the  $x$  variable, we have for a.e.  $(x, y, \xi)$

$$b(x, \xi) - b(y, \xi) = \int_0^1 D_x b(y + a(x - y), \xi) (x - y) da.$$

This formula can be obtained by approximation of  $b$  in  $L^1_{\xi,loc}(W^{1,1}_{x,loc})$  with regular  $b^n$ . By the change of variable  $z = (x - y)/\epsilon$ ,  $\eta = (\xi - \zeta)/\delta$ , we obtain

$$\begin{aligned}
& \int_{\mathbb{R}^d \times \mathbb{R}} \int_{\mathbb{R}^d \times \mathbb{R}} f_r^{\epsilon,\delta}(x, \xi) f_r(y, \zeta) \bar{\rho}_\delta(\xi - \zeta) \left( \nabla \rho_\epsilon(x - y) \cdot (b(x, \xi) - b(y, \zeta)) \right. \\
& \quad \left. + \rho_\epsilon(x - y) \operatorname{div}_y b(y, \zeta) \right) \varphi_r(x) dy d\zeta dx d\xi \\
&= \int_0^1 \int \bar{\rho}(\eta) \nabla \rho(z) \cdot \int f_r^{\epsilon,\delta}(x, \xi) f_r(x - \epsilon z, \xi - \delta \eta) \cdot D_x b(x - \epsilon z, \xi) z \varphi_r(x) dx d\xi dz d\eta da \\
&+ \int \bar{\rho}(\eta) \rho(z) \int f_r^{\epsilon,\delta}(x, \xi) f_r(x - \epsilon z, \xi - \delta \eta) \operatorname{div} b(x - \epsilon z, \xi - \delta \eta) \varphi_r(x) dx d\xi dz d\eta \\
&+ \frac{1}{\epsilon} \int \bar{\rho}(\eta) \nabla \rho(z) \cdot \int f_r^{\epsilon,\delta}(x, \xi) f_r(x - \epsilon z, \xi - \delta \eta) (b(x, \xi - \delta \eta) - b(x, \xi)) \varphi_r(x) dx d\xi dz d\eta \\
(3.3) \quad &=: A + B + C
\end{aligned}$$

Here and in the following we can suppose without loss of generality that all the integrals range over a compact set independent of  $\epsilon$ ,  $\delta$ ,  $r$  and  $\omega$ , since the test functions  $\varphi$ ,  $\rho$ ,  $\bar{\rho}$  are compactly supported and  $f_r(x, \xi)$  and  $f_r^{\epsilon,\delta}(x, \xi)$  are compactly supported in  $\xi$  uniformly in  $\epsilon$ ,  $\delta$ ,  $r$  and  $\omega$ .

We start with the first integral  $A$  on the right hand side of (3.3). We first take the  $L^m_{t,\omega}$ -limit as  $\delta \rightarrow 0$  and we find that  $A$  converges to

$$(3.4) \quad \int_0^1 \int \nabla \rho(z) \cdot \int f_r^\epsilon(x, \xi) f_r(x - \epsilon z, \xi) \cdot D_x b(x - \epsilon z, \xi) z \varphi_r(x) dx d\xi dz da,$$

where  $f^\epsilon(x, \xi) = f(\cdot, \xi) * \rho^\epsilon(x)$ . The proof of this fact is standard and relies on arguments similar to, but simpler than, those for the limit as  $\epsilon \rightarrow 0$ , so we omit it. Now we take the  $L^m_{t,\omega}$ -limit of (3.4) as  $\epsilon \rightarrow 0$ . First we fix  $z$ ,  $a$ ,  $r$  and  $\omega$ . For the inner integral, we have

$$\begin{aligned}
& \int f_r^\epsilon(x, \xi) f_r(x - \epsilon z, \xi) D_x b(x - \epsilon z, \xi) z \varphi_r(x) dx d\xi - \int f_r(x, \xi)^2 \cdot D_x b(x, \xi) z \varphi_r(x) dx d\xi \\
&= \int f_r^\epsilon(x, \xi) f_r(x - \epsilon z, \xi) (D_x b(x - \epsilon z, \xi) - D_x b(x, \xi)) z \varphi_r(x) dx d\xi \\
&+ \int f_r^\epsilon(x, \xi) (f_r(x - \epsilon z, \xi) D_x b(x - \epsilon z, \xi) - f_r(x, \xi) D_x b(x, \xi)) z \varphi_r(x) dx d\xi \\
&+ \int (f_r^\epsilon(x, \xi) - f_r(x, \xi)) f_r(x, \xi) D_x b(x, \xi) z \varphi_r(x) dx d\xi
\end{aligned}$$

The first addend on the right hand side above goes to 0 for  $\epsilon \rightarrow 0$ : indeed both  $D_x b(x - \epsilon z, \xi)$  and  $D_x b(x, \xi)$  tend to  $D_x b(x, \xi)$  in  $L^1_{x,\xi}$  by continuity of translation and  $f^\epsilon(x, \xi) f(x - \epsilon z, \xi) \varphi_r(x)$  is bounded in  $L^\infty_{x,\xi}$  uniformly in  $\epsilon$ . The second addend also goes to 0 for  $\epsilon \rightarrow 0$ :  $f(x - \epsilon z, \xi) D_x b(x - \epsilon z, \xi)$  tends to  $f(x, \xi) D_x b(x, \xi)$  in  $L^1_{x,\xi}$  by continuity of translation and  $f^\epsilon(x, \xi) \varphi_r(x)$  is bounded in  $L^\infty_{x,\xi}$  uniformly in  $\epsilon$ . Finally, the third addend goes to 0 by dominated convergence:  $f^\epsilon(x, \xi) - f(x, \xi)$  tends to 0 for a.e.  $(x, \xi)$  and the integrand is bounded by  $C |D_x b|(x, \xi)$ , for some  $C > 0$ . Therefore, for fixed  $z$ ,  $\eta$ ,  $r$  and  $a$ , the inner integral in the addend  $A$  of (3.3) converges to  $\int f_r^2(x, \xi) D_x b(x, \xi) z \varphi_r(x) dx d\xi$ . Moreover this inner integral is

bounded uniformly in  $z$ ,  $a$ ,  $r$  and  $\omega$ , therefore dominated convergence implies for  $A$

$$\begin{aligned} & \lim_{\varepsilon \rightarrow 0} \left( \lim_{\delta \rightarrow 0} \left( \int_0^1 \int \bar{\rho}(\eta) \nabla \rho(z) \cdot \int f_r^{\varepsilon, \delta}(x, \xi) f_r(x - \varepsilon z, \xi - \delta \eta) \cdot \right. \right. \\ & \quad \left. \left. \cdot D_x b(x - a\varepsilon z, \xi) z \varphi_r(x) dx d\xi dz d\eta da \right) \right) \\ &= \int_0^1 \int \nabla \rho(z) \cdot \int f_r(x, \xi)^2 \cdot D_x b(x, \xi) z \varphi_r(x) dx d\xi dz da \\ &= - \int f_r(x, \xi)^2 \operatorname{div} b(x, \xi) \varphi_r(x) dx d\xi, \end{aligned}$$

where the limits are taken in  $L_{t, \omega}^m$  and we have used that  $\int \partial_i \rho(z) z_j dz = -\delta_{ij}$ .

Similarly, for the second integral  $B$  on the right hand side of (3.3) we have

$$\begin{aligned} & \lim_{\varepsilon \rightarrow 0} \left( \lim_{\delta \rightarrow 0} \left( \int \bar{\rho}(\eta) \rho(z) \int f_r^{\varepsilon, \delta}(x, \xi) f_r(x - \varepsilon z, \xi - \delta \eta) \cdot \right. \right. \\ & \quad \left. \left. \cdot \operatorname{div} b(x - \varepsilon z, \xi - \delta \eta) \varphi_r(x) dx d\xi dz d\eta \right) \right) \\ &= \int f_r(x, \xi)^2 \operatorname{div} b(x, \xi) \varphi_r(x) dx d\xi, \end{aligned}$$

where again the limits are taken in  $L_{t, \omega}^m$ . For the third integral  $C$ , again with similar reasoning but now taking only the limit  $\delta \rightarrow 0$ , we get

$$\begin{aligned} & \lim_{\delta \rightarrow 0} \mathbb{E} \int_0^T \left| \frac{1}{\varepsilon} \int \bar{\rho}(\eta) \nabla \rho(z) \cdot \right. \\ & \quad \left. \cdot \int f^{\varepsilon, \delta}(x, \xi) f(x - \varepsilon z, \xi - \delta \eta) (b(x, \xi - \delta \eta) - b(x, \xi)) \varphi(x) dx d\xi dz d\eta \right|^m dr = 0. \end{aligned}$$

Putting together these limits we obtain the desired statement.  $\square$

Now we prove Lemma 3.5.

*Proof of Lemma 3.5. Step 1:* We start with the equation for  $|f|$ . Since, by (2.14),  $|f| = f \operatorname{sgn}(\xi)$ , we aim to use  $\operatorname{sgn}(\xi)$  as a test function in (2.13). To do so, we regularize  $\operatorname{sgn}$  via  $\operatorname{sgn} *_{\xi} \bar{\rho}^\delta =: \operatorname{sgn}^\delta$ . Note that  $\partial_\xi \operatorname{sgn}^\delta = 2\bar{\rho}^\delta$ . For technical reasons that will become clear in the second step, we write an equation for  $\int_{\mathbb{R}^d \times \mathbb{R}} f^{\varepsilon, \delta} \operatorname{sgn}^\delta(\xi) \varphi dx d\xi$  (where  $f^{\varepsilon, \delta} = f *_{x, \xi} (\rho^\varepsilon \bar{\rho}^\delta)$ ), that is, we take  $(\operatorname{sgn}^\delta(\xi) \varphi) *_{x, \xi} (\rho^\varepsilon \bar{\rho}^\delta)$  as a test function in (2.13). Moreover, again in (2.13) we take the càdlàg version of the integral and thus get, for a.e.  $\omega$  (on a full-measure set independent of  $t$ ), for every  $t$ ,

$$\begin{aligned} & \int_{\mathbb{R}^d \times \mathbb{R}} f_t^{\varepsilon, \delta, +} \operatorname{sgn}^\delta(\xi) \varphi dx d\xi \\ &= \int_{\mathbb{R}^d \times \mathbb{R}} f_0^{\varepsilon, \delta} \operatorname{sgn}^\delta(\xi) \varphi_0 dx d\xi + \int_0^t \int f^{\varepsilon, \delta} \operatorname{sgn}^\delta(\xi) (\partial_t \varphi + \frac{1}{2} \Delta \varphi) dx d\xi dr \\ (3.5) \quad & + \int_0^t \int f b \cdot \nabla (\operatorname{sgn}^\delta(\xi) \varphi)^{\varepsilon, \delta} dx d\xi dr + \int_0^t \int f \operatorname{div} b (\operatorname{sgn}^\delta(\xi) \varphi)^{\varepsilon, \delta} dx d\xi dr \\ & - \int_0^t \int f^{\varepsilon, \delta} \operatorname{sgn}^\delta(\xi) \nabla \varphi dx d\xi dW_r - 2 \int_{[0, t] \times \mathbb{R}^d \times \mathbb{R}} \varphi \bar{\rho}^\delta(\xi) m^{\varepsilon, \delta} dx d\xi dr, \end{aligned}$$

where  $f^{\varepsilon, \delta, +}$  is the càdlàg version of  $f^{\varepsilon, \delta}$  and  $m^{\varepsilon, \delta} = m *_{x, \xi} (\rho^\varepsilon \bar{\rho}^\delta)$  (see Remark 2.13).

*Step 2:* For  $f^2$ , we would like to take  $f\varphi$  as a test function in (2.13). Since  $f$  is not regular, we regularize it in both  $x$  and  $\xi$ . More precisely we take  $f^{\varepsilon, \delta, +}$ ,  $f^{\varepsilon, \delta, -}$  resp.

càdlàg, càglàd versions of  $f^{\varepsilon, \delta}$ , as in Lemma 2.12 and Remark 2.13. Itô's formula for càdlàg processes (cf. [44, Chapter II Theorem 33]) yields

$$\begin{aligned}
& (f_t^{\varepsilon, \delta, +}(x, \xi))^2 \varphi_t(x) - (f_0^{\varepsilon, \delta}(x, \xi))^2 \varphi_0(x) \\
&= \int_0^t f_r^{\varepsilon, \delta, +}(x, \xi)^2 \partial_t \varphi_r(x) dr + \int_{[0, t]} (f_r^{\varepsilon, \delta, +}(x, \xi) + f_r^{\varepsilon, \delta, -}(x, \xi)) \varphi_r(x) df_r^{\varepsilon, \delta} \\
&\quad + \int_0^t \varphi_r(x) d[f^{\varepsilon, \delta}]_r \\
&= \int_0^t f_r^{\varepsilon, \delta}(x, \xi)^2 \partial_t \varphi_r(x) dr + \int_0^t f_r^{\varepsilon, \delta}(x, \xi) \Delta f_r^{\varepsilon, \delta}(x, \xi) \varphi_r(x) dr \\
&\quad - \int_0^t 2f_r^{\varepsilon, \delta}(x, \xi) (b \cdot \nabla f_r)^{\varepsilon, \delta}(x, \xi) \varphi_r(x) dr - \int_0^t 2f_r^{\varepsilon, \delta}(x, \xi) \nabla f_r^{\varepsilon, \delta}(x, \xi) \varphi_r(x) dW_r \\
&\quad + \int_{[0, t]} (f_r^{\varepsilon, \delta, +}(x, \xi) + f_r^{\varepsilon, \delta, -}(x, \xi)) \varphi_r(x) \partial_\xi m^{\varepsilon, \delta}(r, x, \xi) dx d\xi dr \\
&\quad + \int_0^t |\nabla f_r^{\varepsilon, \delta}(x, \xi)|^2 \varphi_r(x) dr.
\end{aligned}$$

This formula is valid for each  $(x, \xi)$  for a.e.  $(t, \omega)$ , where the exceptional set may depend on  $(x, \xi)$ . However, by Remark 2.13, for a fixed representative of  $m$ , the integral with the measure  $\partial_\xi m^{\varepsilon, \delta}$  is measurable in  $(t, \omega, x, \xi)$ , càdlàg in  $t$  for  $(\omega, x, \xi)$  fixed and continuous in  $(x, \xi)$  for  $(t, \omega)$  fixed. Also the other integrals have versions that are continuous in  $(t, x, \xi)$  for  $\omega$  fixed and, in particular, are measurable in  $(t, \omega, x, \xi)$ . For such versions, for a.e.  $\omega$ , the above equality above holds for every  $(t, x, \xi)$ .

The idea at this point is first to integrate in  $x$  and  $\xi$ , then to use integration by parts to bring the derivatives onto  $\varphi$  and thereby to get an equation for  $f^{\varepsilon, \delta, +}$  which is similar to the one satisfied by  $f$  itself, plus a remainder. Indeed we integrate in  $x$  and  $\xi$  and use Remark 2.13, Fubini's theorem and the stochastic Fubini theorem: we obtain the following equality, valid for every  $t$  and for every  $\omega$  in a full-measure set independent of  $t$ ,

$$\begin{aligned}
& \int_{\mathbb{R}^d \times \mathbb{R}} (f_t^{\varepsilon, \delta, +}(x, \xi))^2 \varphi_t(x) dx d\xi - \int_{\mathbb{R}^d \times \mathbb{R}} (f_0^{\varepsilon, \delta}(x, \xi))^2 \varphi_0(x) dx d\xi \\
&= \int_0^t \int_{\mathbb{R}^d \times \mathbb{R}} f_r^{\varepsilon, \delta}(x, \xi)^2 \left( \partial_t \varphi_r(x) + \frac{1}{2} \Delta \varphi_r(x) + \operatorname{div}(b(x, \xi) \varphi_r(x)) \right) dx d\xi dr \\
&\quad + 2 \int_0^t \int_{\mathbb{R}^d \times \mathbb{R}} \int_{\mathbb{R}^d \times \mathbb{R}} f_r^{\varepsilon, \delta}(x, \xi) f_r(y, \zeta) \bar{\rho}_\delta(\xi - \zeta) \left( \nabla \rho_\varepsilon(x - y) \cdot (b(x, \xi) - b(y, \zeta)) \right. \\
&\quad \left. + \rho_\varepsilon(x - y) \operatorname{div}_y b(y, \zeta) \right) \varphi_r(x) dy d\zeta dx d\xi dr \\
&\quad - \int_{[0, t] \times \mathbb{R}^d \times \mathbb{R}} (\partial_\xi f_r^{\varepsilon, \delta, +} + \partial_\xi f_r^{\varepsilon, \delta, -})(x, \xi) m^{\varepsilon, \delta}(r, x, \xi) \varphi_r(x) dx d\xi dr \\
&\quad - \int_0^t \int_{\mathbb{R}^d \times \mathbb{R}} f_r^{\varepsilon, \delta}(x, \xi)^2 \nabla \varphi_r(x) dx d\xi dW_r.
\end{aligned}$$

For the third addend, note that, for every  $(x, \xi)$ , it holds, for a.e.  $(t, \omega)$ ,  $\partial_\xi f^{\varepsilon, \delta}(x, \xi) = \bar{\rho}_\delta(\xi) - \nu *_\xi \bar{\rho}_\delta(\xi)$  (the convolution being in the  $\xi$  direction) and so  $\partial_\xi f^{\varepsilon, \delta}(x, \xi) \geq \bar{\rho}_\delta(\xi)$ . Therefore, by the càdlàg/càglàd properties of  $\partial_\xi f^{\varepsilon, \delta, +}$  and  $\partial_\xi f^{\varepsilon, \delta, -}$ , it holds for a.e.  $\omega$ : for every  $(t, x, \xi)$ ,  $\partial_\xi f^{\varepsilon, \delta, +}(x, \xi) \geq \bar{\rho}_\delta(\xi)$  and  $\partial_\xi f^{\varepsilon, \delta, -}(x, \xi) \geq \bar{\rho}_\delta(\xi)$ . So

we obtain

$$\begin{aligned} & - \int_{[0,t] \times \mathbb{R}^d \times \mathbb{R}} (\partial_\xi f_r^{\epsilon,\delta,+} + \partial_\xi f_r^{\epsilon,\delta,-})(x, \xi) m^{\epsilon,\delta}(r, x, \xi) \varphi_r(x) dx d\xi dr \\ & \leq -2 \int_{[0,t] \times \mathbb{R}^d \times \mathbb{R}} \bar{\rho}^\delta(\xi) m^{\epsilon,\delta}(r, x, \xi) \varphi_r(x) dx d\xi dr. \end{aligned}$$

Here we see the reason for the additional regularization of  $f^{\epsilon,\delta}$  in the first step: in this way the right hand side of the above inequality is equal to the last term in formula (3.5). In conclusion we get, for a.e.  $\omega$  (on a full-measure set independent of  $t$ ), for every  $t$ ,

$$\begin{aligned} & \int_{\mathbb{R}^d \times \mathbb{R}} (f_t^{\epsilon,\delta,+} \operatorname{sgn}^\delta(\xi) - (f_t^{\epsilon,\delta,+})^2) \varphi_t dx d\xi - \int_{\mathbb{R}^d \times \mathbb{R}} (f_0^{\epsilon,\delta} \operatorname{sgn}^\delta(\xi) - (f_0^{\epsilon,\delta})^2) \varphi_0 dx d\xi \\ & \leq \int_0^t \int f^{\epsilon,\delta} \operatorname{sgn}^\delta(\xi) (\partial_t \varphi_r + \frac{1}{2} \Delta \varphi_r) dx d\xi dr \\ & + \int_0^t \int f b \cdot \nabla (\operatorname{sgn}^\delta(\xi) \varphi)^{\epsilon,\delta} dx d\xi dr + \int_0^t \int f \operatorname{div} b (\operatorname{sgn}^\delta(\xi) \varphi)^{\epsilon,\delta} dx d\xi dr \\ (3.6) \quad & - \int_0^t \int_{\mathbb{R}^d \times \mathbb{R}} (f^{\epsilon,\delta})^2 (\partial_t \varphi + \frac{1}{2} \Delta \varphi + \operatorname{div}(b\varphi)) dx d\xi dr \\ & - \int_0^t \int_{\mathbb{R}^d \times \mathbb{R}} (f^{\epsilon,\delta} \operatorname{sgn}^\delta(\xi) - (f^{\epsilon,\delta})^2) \nabla \varphi dx d\xi dW_r \\ & - 2 \int_0^t \int_{\mathbb{R}^d \times \mathbb{R}} \int_{\mathbb{R}^d \times \mathbb{R}} f_r^{\epsilon,\delta}(x, \xi) f_r(y, \zeta) \bar{\rho}_\delta(\xi - \zeta) (\nabla \rho_\epsilon(x - y) \cdot (b(x, \xi) - b(y, \zeta))) \\ & + \rho^\epsilon(x - y) \operatorname{div}_y b(y, \zeta) \varphi_r(x) dy d\zeta dx d\xi dr. \end{aligned}$$

*Step 3:* The last addend in the right hand side above is the commutator error, which by Lemma 3.6 goes to zero in  $L_{t,\omega}^2$  letting first  $\delta \rightarrow 0$  and then  $\epsilon \rightarrow 0$ . Therefore, taking the  $L_{t,\omega}^2$ -limit in (3.6) first for  $\delta \rightarrow 0$  then for  $\epsilon \rightarrow 0$ , we obtain the statement.  $\square$

We are ready to prove the key Lemma 3.3.

*Proof of Lemma 3.3.* By Lemma 3.5 we have, for every nonnegative test function  $\varphi$  in  $C_c^\infty([0, T] \times \mathbb{R}^d)$  independent of  $\xi$ , for a.e.  $t$  (with the exceptional set possibly depending on  $\varphi$ ),

$$\begin{aligned} \mathbb{E} \int_{\mathbb{R}^d \times \mathbb{R}} (|f_t| - f_t^2) \varphi_t dx d\xi & \leq \int_{\mathbb{R}^d \times \mathbb{R}} (|f_0| - f_0^2) \varphi_0 dx d\xi \\ (3.7) \quad & + \mathbb{E} \int_0^t \int_{\mathbb{R}^d \times \mathbb{R}} [\partial_t \varphi + \frac{1}{2} \Delta \varphi + \operatorname{div}(b(x, \xi) \varphi)] (|f| - f^2) dx d\xi dr; \end{aligned}$$

here we used that  $\int_0^t \int_{\mathbb{R}^d \times \mathbb{R}} \nabla \varphi (|f| - f^2) dx d\xi dW_r$  is an  $L^2$  martingale with zero mean, since  $\nabla \varphi (|f| - f^2)$  is bounded and compactly supported.

The idea at this point is to use duality; that is, we would like to take a test function  $\varphi$ , independent of  $\xi$ , nonnegative and sufficiently regular, with  $\varphi_T > 0$ , such that, for every  $\xi$  in a bounded interval  $[-R, R]$ ,

$$(3.8) \quad \partial_t \varphi + \frac{1}{2} \Delta \varphi + \operatorname{div}(b(x, \xi) \varphi) \leq C.$$



Then we could conclude by Gronwall's inequality. To do so, the strategy is as follows. First we take  $\varphi$  as a nonnegative solution to

$$\partial_t \varphi + \frac{1}{2} \Delta \varphi + F(x) \varphi = 0, \quad \varphi(t_{fin}, x) = 1,$$

with  $F(x) = \|\operatorname{div} b(x, \cdot)\|_{L_{\xi, B_R}^\infty}$  (measurable function),  $t_{fin}$  a given time and  $R$  such that the support of  $f$  is in  $[0, T] \times \Omega \times \mathbb{R}^d \times [-R, R]$ . Then we use a bound on the transport term  $b \cdot \nabla \varphi$  to obtain (3.8).

For technical reasons, we take, for  $\epsilon, t_{fin} > 0$  fixed,  $\varphi^\epsilon$  to be a solution on  $[0, t_{fin}]$  to

$$(3.9) \quad \partial_t \varphi^\epsilon + \frac{1}{2} \Delta \varphi^\epsilon + F^\epsilon \varphi^\epsilon = 0, \quad \varphi^\epsilon(t_{fin}, x) = \psi_{1/\epsilon}(x);$$

here  $\psi_{1/\epsilon}$  is a  $C_c^\infty$  nonnegative function, with values in  $[0, 1]$ , equal to 1 on  $B_{1/\epsilon}(0)$  and uniformly bounded (in  $\epsilon$ ) in the  $W^{1,\infty}(\mathbb{R}^d)$  norm;  $F^\epsilon$  is a compactly supported regularization of  $F$ , converging to  $F$  a.e. and in  $L^p$ , if  $p < \infty$ , or a.e. and with uniform  $L^\infty$  bound, if  $p = \infty$ . We extend  $\varphi^\epsilon$  to the whole interval  $[0, T]$  by taking  $\varphi^\epsilon(t, x) = \psi_{1/\epsilon}(x)$  for  $t \in [t_{fin}, T]$ . By Remark 4.1 below,  $\varphi^\epsilon$  is nonnegative and in  $L_t^\infty(W_x^{2,\infty}) \cap L_x^\infty(W_t^{1,\infty})$  for every  $\epsilon > 0$ . Therefore, reasoning as in Remark 2.11,  $\varphi^\epsilon$  can be used as test function in (3.7). Consequently, we have, for a.e.  $t \leq t_{fin}$ , with the exceptional set  $N^{\epsilon, t_{fin}}$  possibly depending on  $\epsilon$  and  $t_{fin}$ ,

$$(3.10) \quad \begin{aligned} & \mathbb{E} \int \varphi_t^\epsilon (|f_t| - f_t^2) dx d\xi \\ & \leq \int_{\mathbb{R}^d \times \mathbb{R}} \varphi_0^\epsilon (|f_0| - f_0^2) dx d\xi \\ & \quad + \int_0^t \mathbb{E} \int [\partial_t \varphi^\epsilon + \frac{1}{2} \Delta \varphi^\epsilon + F^\epsilon \varphi^\epsilon] (|f| - f^2) dx d\xi dr \\ & \quad + \int_0^t \mathbb{E} \int [b \cdot \nabla \varphi^\epsilon + (\operatorname{div} b) \varphi^\epsilon - F \varphi^\epsilon] (|f| - f^2) dx d\xi dr \\ & \quad + \int_0^t \mathbb{E} \int [F \varphi^\epsilon - F^\epsilon \varphi^\epsilon] (|f| - f^2) dx d\xi dr \\ & \leq \int_0^t \mathbb{E} \int (b \cdot \nabla \varphi^\epsilon) (|f| - f^2) dx d\xi dr \\ & \quad + \int_0^t \mathbb{E} \int (F - F^\epsilon) \varphi^\epsilon (|f| - f^2) dx d\xi dr, \end{aligned}$$

where we have used that  $|f| - f^2 \geq 0$  and that  $f$  is supported on  $[0, T] \times \Omega \times \mathbb{R}^d \times [-R, R]$ .

Before passing to the limit  $\epsilon \rightarrow 0$ , we aim to replace  $t$  by  $t_{fin}$  in the above inequality. This is not immediate, since the function  $t \mapsto \mathbb{E}[|f_t| - f_t^2]$  is not known to be (even weakly) continuous. To overcome this difficulty, we fix a version of the map  $[0, T] \rightarrow L^1(\mathbb{R}^d \times \mathbb{R})$  given by  $t \mapsto \mathbb{E}[|f_t| - f_t^2]$  and we use Lusin's theorem for separable Banach space-valued functions, see for example Loeb, Talvila [40]: for every  $\delta > 0$ , there exists a measurable set  $A_\delta \subseteq [0, T]$  with Lebesgue measure  $|A_\delta| \geq T - \delta$ , such that  $t \mapsto \mathbb{E}[|f_t| - f_t^2]$  is continuous on  $A_\delta$  as an  $L^1(\mathbb{R}^d \times \mathbb{R})$ -valued map. We can also assume that  $A_\delta$  has no points which are isolated from the left (here we say that  $t_0$  is isolated from the left in  $A_\delta$  if  $(t_0 - \eta, t_0) \cap A_\delta = \emptyset$  for some  $\eta > 0$ ): indeed, the set of points of  $A_\delta$  which are isolated from the left is at most countable and thus has zero Lebesgue measure. Therefore, for  $t_{fin} \in A_\delta$ , we can find a sequence  $t_n \leq t_{fin}$  in  $A_\delta \setminus N^{\epsilon, t_{fin}}$  converging to  $t_{fin}$  (as  $n \rightarrow \infty$ ) and such that (3.10) holds for  $t_n$  and  $\mathbb{E}[|f_{t_n}| - f_{t_n}^2] \rightarrow \mathbb{E}[|f_{t_{fin}}| - f_{t_{fin}}^2]$  in  $L^1(\mathbb{R}^d \times \mathbb{R})$ .

Moreover, by Remark 4.1  $\varphi^\epsilon$  is in  $L_x^\infty(W_t^{1,\infty})$  and so the map  $[0, T] \rightarrow L^\infty(\mathbb{R}^d \times \mathbb{R})$  given by  $t \mapsto \varphi_t^\epsilon$  is continuous. Hence, by Hölder's inequality,

$$\mathbb{E} \int \varphi_{t_n}^\epsilon (|f_{t_n}| - f_{t_n}^2) dx d\xi \rightarrow \mathbb{E} \int \varphi_{t_{fin}}^\epsilon (|f_{t_{fin}}| - f_{t_{fin}}^2) dx d\xi.$$

Since the right hand side of (3.10) is continuous in time, we can pass to the limit in (3.10) for  $t_n \rightarrow t_{fin}$  and obtain (3.10) for  $t_{fin} \in A_\delta$ . Since this is true for any  $\delta > 0$ , we obtain (3.10) for a.e.  $t = t_{fin}$ .

Now we let  $\epsilon$  go to 0. By Lemma 4.3, applied to the backward PDE (3.9), and the uniform bound on  $F^\epsilon$  in  $L^p$ , we have a uniform (in  $\epsilon$ ) bound on  $\|\varphi^\epsilon\|_{L_t^\infty(W_x^{1,\infty})}$ . Therefore, we can bound the first addend of the right hand side in (3.10) by

$$\begin{aligned} & \limsup_{\epsilon \rightarrow 0} \int_0^{t_{fin}} \mathbb{E} \int (b \cdot \nabla \varphi^\epsilon) (|f| - f^2) d\xi dx \\ & \leq \|b\|_{L_{\xi, [-R, R]}^\infty(L_x^\infty)} \sup_\epsilon \|\varphi^\epsilon\|_{L_t^\infty(W_x^{1,\infty})} \int_0^{t_{fin}} \mathbb{E} \int (|f| - f^2) dx d\xi dr \\ (3.11) \quad & \leq C \int_0^{t_{fin}} \mathbb{E} \int (|f| - f^2) dx d\xi dr. \end{aligned}$$

Concerning the second addend in (3.10), in the case  $p < \infty$ ,  $F - F^\epsilon$  converges to 0 in  $L_x^p$  and thus in  $L^p([0, T] \times \Omega \times \mathbb{R}^d \times [-R, R])$ ;  $\varphi^\epsilon$  is uniformly bounded in  $L_{t,x}^\infty$  and  $(|f| - f^2)$  is in  $L^{p'}([0, T] \times \Omega \times \mathbb{R}^d \times [-R, R])$  by Remark 2.11. Therefore, by Hölder's inequality,

$$\limsup_{\epsilon \rightarrow 0} \int_0^{t_{fin}} \mathbb{E} \int (F - F^\epsilon) \varphi^\epsilon (|f| - f^2) dx d\xi dr = 0.$$

In the case  $p = \infty$  we get the same result: here we exploit (via dominated convergence theorem) the a.e. convergence to 0 and the uniform bound of  $\varphi^\epsilon(F - F^\epsilon)$  and the fact that  $(|f| - f^2)$  is in  $L^1([0, T] \times \Omega \times \mathbb{R}^d \times [-R, R])$ . Finally, concerning the initial condition, using again the uniform bound from Lemma 4.3 we get  $\int_{\mathbb{R}^d \times \mathbb{R}} \varphi_0^\epsilon (|f_0| - f_0^2) dx d\xi \leq C \int_{\mathbb{R}^d \times \mathbb{R}} (|f_0| - f_0^2) dx d\xi$ . Putting all together we have, for a.e.  $t_{fin} > 0$ ,

$$\begin{aligned} \mathbb{E} \int (|f_{t_{fin}}| - f_{t_{fin}}^2) d\xi dx & \leq C \int_{\mathbb{R}^d \times \mathbb{R}} (|f_0| - f_0^2) dx d\xi \\ & \quad + C \int_0^{t_{fin}} \mathbb{E} \int (|f| - f^2) d\xi dx dr. \end{aligned}$$

We conclude by Gronwall's lemma for discontinuous functions (cf. Ethier, Kurtz [19, Theorem 5.1 in the Appendix]) that, for a.e.  $t \in [0, T]$ ,

$$\mathbb{E} \int (|f_t| - f_t^2) d\xi dx \leq C \int_{\mathbb{R}^d \times \mathbb{R}} (|f_0| - f_0^2) dx d\xi,$$

where  $C$  is a constant that depends only on the bound (3.11) and on the a priori estimates in Lemma 4.3, applied to the backward PDE (3.9). Therefore,  $C$  depends only on  $T$ ,  $\|b\|_{L_{\xi, [-R, R]}^\infty(L_x^\infty)}$  and  $\|\operatorname{div} b\|_{L_x^p(L_{\xi, [-R, R]}^\infty)}$ . The proof is complete.  $\square$

Finally we prove Proposition 3.4.

*Proof of Proposition 3.4.* Since  $f_0$  takes values in  $\{0, \pm 1\}$ , we have  $|f_0| - f_0^2 = 0$ . Therefore, Lemma 3.3 implies  $f^2 - |f| = 0$  a.s. (recall  $|f| \leq 1$  by definition) and thus  $f$  takes values in  $\{0, \pm 1\}$  for a.e.  $(t, \omega, x, \xi)$ . We then define  $u(t, \omega, x) := \int_{\mathbb{R}} f(t, \omega, x, \xi) d\xi$ . Note that  $u$  is well-defined since  $f$  is compactly supported in  $\xi$  and measurable by Fubini's theorem.

Now we claim that, for every  $h > 0$ , for a.e.  $(t, \omega, x, \xi)$ ,

$$(3.12) \quad (f(t, \omega, x, \xi) - f(t, \omega, x, \xi + h))(1_{-\infty < \xi < -h} + 1_{h < \xi < +\infty}) \geq 0,$$

$$(3.13) \quad f(t, \omega, x, \xi) - f(t, \omega, x, \xi + h) + 1 \geq 0.$$

Leaving the proof of these inequality for later, we use them to conclude. Since the pushforward of the Lebesgue measure via the map  $(\xi, h) \mapsto (\xi, \xi + h)$  is equivalent to the Lebesgue measure, the two inequalities above imply, for a.e.  $(t, \omega, x, \xi, \eta)$ ,

$$(3.14) \quad (f(t, \omega, x, \xi) - f(t, \omega, x, \eta))(1_{\xi < \eta < 0} + 1_{0 < \xi < \eta}) \geq 0,$$

$$(3.15) \quad (f(t, \omega, x, \xi) - f(t, \omega, x, \eta) + 1)1_{\xi < \eta} \geq 0.$$

Now we fix a version of  $f$  and we consider, for fixed  $(t, \omega, x)$ , the set  $A = A(t, \omega, x) = \{\xi < 0 : (f(t, \omega, x, \xi) - f(t, \omega, x, \eta))\text{sgn}(\xi - \eta) \leq 0 \text{ for a.e. } \eta < 0\}$ . By Fubini's theorem, (3.14) implies that, for a.e.  $(t, \omega, x)$ ,  $A(t, \omega, x)$  is a full-measure set on  $(-\infty, 0)$ . Moreover, for any  $(t, \omega, x)$ ,  $f$  is non-increasing on  $A(t, \omega, x)$ . Indeed, if this would not be true, we could find  $\xi < \eta$  in  $A$  with  $f(t, \omega, x, \xi) - f(t, \omega, x, \eta) < 0$ . Thus, since  $f(t, \omega, x, \xi) - f(t, \omega, x, \zeta) \geq 0$  for a.e.  $\zeta > \xi$ , we obtain  $f(t, \omega, x, \zeta) - f(t, \omega, x, \eta) < 0$  for a.e.  $\zeta \in (\xi, \eta)$ , in contradiction to  $\eta \in A$ . Similarly, for a.e.  $(t, \omega, x)$ ,  $B(t, \omega, x) = \{\xi > 0 : (f(t, \omega, x, \xi) - f(t, \omega, x, \eta))\text{sgn}(\xi - \eta) \leq 0 \text{ for a.e. } \eta > 0\}$  is a full-measure set on  $(0, +\infty)$  on which  $f$  is non-increasing. Since  $f$  is compactly supported in  $\xi$  and takes values a.e. in  $\{0, \pm 1\}$ , we conclude for a.e.  $(t, \omega, x)$ ,  $f = -1_{\{a < \xi < 0\}} + 1_{\{0 < \xi < b\}}$  for some  $a \leq 0 \leq b$  (depending on  $(t, \omega, x)$ ) on the full-measure set  $A(t, \omega, x) \cup B(t, \omega, x)$ . By (3.15) this yields that either  $f = -1_{\{a < \xi < 0\}}$  a.e. or  $f = -1_{\{0 < \xi < b\}}$  a.e. and thus  $f = \chi(u)$  a.e.. Progressive measurability of  $u$  follows from the respective property of  $f = \chi(u)$ , by Remark 2.5.

It remains to prove the claim above, that is, (3.12) and (3.13). To prove (3.12) we take a nonnegative test function  $\psi$  in  $C_c^\infty([0, T] \times \mathbb{R}^d \times \mathbb{R})$  with support contained in  $(-\infty, -h)$ . We call  $\varphi$  the function such that  $\psi = -\partial_\xi \varphi$  and that  $\varphi(-a) = 0$  for  $a$  large enough;  $\varphi$  is a nonpositive nonincreasing function, constant on  $[-h, +\infty)$ . We then have by (2.14), for a.e.  $\omega$ ,

$$\begin{aligned} & \int (f(t, x, \xi) - f(t, x, \xi + h))\psi(t, x, \xi) dx d\xi dt \\ &= \int f(t, x, \xi)(\psi(t, x, \xi) - \psi(t, x, \xi - h)) dx d\xi dt \\ &= \int (\varphi(t, x, 0) - \varphi(t, x, -h)) dx dt - \int (\varphi(t, x, \xi) - \varphi(t, x, \xi - h)) \nu(dx, d\xi, dt) \\ &= - \int (\varphi(t, x, \xi) - \varphi(t, x, \xi - h)) \nu(dx, d\xi, dt) \geq 0. \end{aligned}$$

This proves that  $f(t, \omega, x, \xi) - f(t, \omega, x, \xi + h) \geq 0$  on  $\{-\infty < \xi < -h\}$ ; similarly for  $\{h < \xi < +\infty\}$ . This proves (3.12).

For (3.13), we take a nonnegative test function  $\psi$  in  $C_c^\infty([0, T] \times \mathbb{R}^d \times \mathbb{R})$  and we call  $\varphi$  the nonpositive, nonincreasing function such that  $\psi = -\partial_\xi \varphi$  and that  $\varphi(-a) = 0$  for  $a$  large enough. Again we have by (2.14)

$$\begin{aligned} & \int (f(t, x, \xi) - f(t, x, \xi + h) + 1)\psi(t, x, \xi) dx d\xi dt \\ &= \int f(t, x, \xi)(\psi(t, x, \xi) - \psi(t, x, \xi - h)) dx d\xi dt + \int (\int \psi(t, x, \xi) d\xi) dx dt \\ &= \int (\varphi(t, x, 0) - \varphi(t, x, -h)) dx dt - \int (\varphi(t, x, \xi) - \varphi(t, x, \xi - h)) \nu(dx, d\xi, dt) \\ & \quad + \int (\varphi(t, x, -R_1) - \varphi(t, x, R_1)) dx dt, \end{aligned}$$

for some  $R_1$  such that the support of  $\psi$  is contained in  $[0, T] \times \mathbb{R}^d \times [-R_1, R_1]$ . Now the monotonicity property of  $\varphi$  gives that  $\varphi(t, x, \xi) - \varphi(t, x, \xi - h) \leq 0$  for every  $\xi$ , and also  $\varphi(t, x, -R_1) - \varphi(t, x, -h) \geq 0$  and  $\varphi(t, x, 0) - \varphi(t, x, R_1) \geq 0$ . Therefore, the right hand side of the formula above is  $\geq 0$ . This proves (3.13) and concludes the proof of the claim.  $\square$

#### 4. APPENDIX A: A PRIORI ESTIMATES ON PARABOLIC PDES

In this section we provide a priori estimates for a linear parabolic PDE on  $\mathbb{R}^d$  of the form

$$(4.1) \quad \partial_t \varphi = \frac{1}{2} \Delta \varphi + g \varphi + h \varphi,$$

where  $g \in L_x^p$  for some finite  $p > d$  and  $h \in L_x^\infty$  and the initial datum  $\varphi_0$  is nonnegative. Since we are interested in a priori estimates in this section, we suppose that  $g, h$  and  $\varphi_0$  are smooth and compactly supported. The estimates can be applied also to the backward PDE, by a change of time. The methods used in this section are essentially classical, see for example (among many other references) Krylov [34] for heat kernel estimates in  $L^p$  spaces and Fedrizzi, Flandoli [20] for estimates on Kolmogorov-type PDEs.

*Remark 4.1.* The existence of a nonnegative solution  $\varphi$  in  $L_t^\infty(W_x^{2,\infty})$  to (4.1) in the case of smooth compactly supported coefficients and nonnegative initial datum is ensured, for example, by the representation formula

$$(4.2) \quad \varphi(t, x) = \mathbb{E}[\exp[\int_0^t (g(x + W_r - W_t) + h(x + W_r - W_t)) dr] \varphi_0(x - W_t)],$$

where the expectation  $\mathbb{E}$  and Brownian motion  $W$  are defined on some generic probability space, not related to the one used before. The equation also implies, again for smooth compactly supported data, that such a solution is in  $L_x^\infty(W_t^{1,\infty})$ .

We start by recalling the regularizing properties of the heat kernel, of easy (and classical) proof:

**Lemma 4.2.** *Let  $p_t(x) = t^{-d/2} p_1(t^{-1/2}x)$  be the heat kernel on  $\mathbb{R}^d$ , i.e.  $p_1(x) = (2\pi)^{-d/2} e^{-|x|^2/2}$ . Then we have, for  $m \in [1, \infty]$ ,*

$$\|p_t\|_{L_x^m} \leq C_{m,d} t^{-(d-d/m)/2} \quad \text{and} \quad \|\nabla p_t\|_{L_x^m} \leq C_{m,d} t^{-(1+d-d/m)/2}.$$

*Proof.* We only prove the second inequality, the proof of the first one being similar. The case  $m = \infty$  is obvious, thus let  $m \in [1, \infty)$ . Note that  $\nabla p_t(x) = t^{-(1+d)/2} \nabla p_1(t^{-1/2}x)$ . By the change of variable  $y = t^{-1/2}x$ , we get

$$\int_{\mathbb{R}^d} |\nabla p_t(x)|^m dx = t^{-(1+d)m/2} t^{d/2} \int_{\mathbb{R}^d} |\nabla p_1(y)|^m dy = t^{-m(1+d-d/m)/2} \|\nabla p_1\|_{L_x^m}^m,$$

which is the desired estimates.  $\square$

We write the PDE (4.1) using the variational formulation:

$$\varphi_t = p_t * \varphi_0 + \int_0^t p_{t-s} * (g \varphi_s) ds + \int_0^t p_{t-s} * (h \varphi_s) ds.$$

**Lemma 4.3.** *There exists a locally bounded function  $c = c(T, \|g\|_{L_x^p}, \|h\|_{L_x^\infty})$  such that, for every  $\varphi_0$  in  $C_c^\infty$ , it holds*

$$\|\varphi_t\|_{W_x^{1,\infty}} \leq \|\varphi_0\|_{W_x^{1,\infty}} c(T, \|g\|_{L_x^p}, \|h\|_{L_x^\infty}).$$

*Proof.* Here  $C$  denotes any positive constant, which can change from line to line, possibly depending on  $T$ ,  $p$  and  $d$ . We start with the  $L_x^\infty$  estimate. Using Young's inequality for convolutions we get

$$\begin{aligned} \|\varphi_t\|_{L^\infty} &\leq \|p_t * \varphi_0\|_{L^\infty} + \int_0^t \|p_{t-s} * (g\varphi_s)\|_{L^\infty} ds + \int_0^t \|p_{t-s} * (h\varphi_s)\|_{L^\infty} ds \\ &\leq C\|\varphi_0\|_{L^\infty} + C \int_0^t (t-s)^{-d/2p} \|g\varphi_s\|_{L^p} ds + C \int_0^t \|h\varphi_s\|_{L^\infty} ds \\ &\leq C\|\varphi_0\|_{L^\infty} + C \int_0^t (t-s)^{-d/2p} \|g\|_{L^p} \|\varphi_s\|_{L^\infty} ds + C \int_0^t \|h\|_{L^\infty} \|\varphi_s\|_{L^\infty} ds. \end{aligned}$$

Since  $p > d$ ,  $(t-s)^{-d/2p}$  is locally in  $L^2$ , Hölder's inequality yields

$$\|\varphi_t\|_{L^\infty} \leq C\|\varphi_0\|_{L^\infty} + C\|g\|_{L^p} \left( \int_0^t \|\varphi_s\|_{L^\infty}^2 ds \right)^{1/2} + C\|h\|_{L^\infty} \left( \int_0^t \|\varphi_s\|_{L^\infty}^2 ds \right)^{1/2}$$

and thus

$$\|\varphi_t\|_{L^\infty}^2 \leq C\|\varphi_0\|_{L^\infty}^2 + C\|g\|_{L^p}^2 \int_0^t \|\varphi_s\|_{L^\infty}^2 ds + C\|h\|_{L^\infty}^2 \int_0^t \|\varphi_s\|_{L^\infty}^2 ds.$$

Gronwall's inequality implies

$$(4.3) \quad \|\varphi_t\|_{L^\infty} \leq C\|\varphi_0\|_{L^\infty} \exp[C(\|g\|_{L^p}^2 + \|h\|_{L^\infty}^2)].$$

We continue with the  $L_x^\infty$  estimate for  $\nabla\varphi_t$ . Using again Young's inequality we get

$$\begin{aligned} \|\nabla\varphi_t\|_{L^\infty} &\leq \|p_t * \nabla\varphi_0\|_{L^\infty} + \int_0^t \|\nabla p_{t-s} * (g\varphi_s)\|_{L^\infty} ds + \int_0^t \|\nabla p_{t-s} * (h\varphi_s)\|_{L^\infty} ds \\ &\leq C\|\nabla\varphi_0\|_{L^\infty} + C \int_0^t (t-s)^{-(1+d/p)/2} \|g\varphi_s\|_{L^p} ds + C \int_0^t (t-s)^{-1/2} \|h\varphi_s\|_{L^\infty} ds \\ &\leq C\|\nabla\varphi_0\|_{L^\infty} + C \int_0^t (t-s)^{-(1+d/p)/2} \|g\|_{L^p} \|\varphi_s\|_{L^\infty} ds \\ &\quad + C \int_0^t (t-s)^{-1/2} \|h\|_{L^\infty} \|\varphi_s\|_{L^\infty} ds. \end{aligned}$$

Since  $p > d$ ,  $(t-s)^{-(1+d/p)/2}$  is locally integrable and we obtain, with (4.3),

$$\|\nabla\varphi_t\|_{L^\infty} \leq C\|\nabla\varphi_0\|_{L^\infty} + C(\|g\|_{L^p} + \|h\|_{L^\infty})\|\varphi_0\|_{L^\infty} \exp[C(\|g\|_{L^p}^2 + \|h\|_{L^\infty}^2)].$$

The proof is complete.  $\square$

## 5. APPENDIX B: MEASURABILITY

In the following, let  $(E, \mathcal{E}, \mu)$  be a  $\sigma$ -finite measure space. For a function  $f : E \rightarrow \mathbb{R}$  recall the definition (given in the introduction) of measurability. Given a Banach space  $V$  and a function  $f : E \rightarrow V$ , we recall the following three definitions of measurability of  $f$ :

- we say that  $f$  is strongly measurable if it is the pointwise (everywhere) limit of a sequence of  $V$ -valued simple measurable functions (i.e. of the form  $\sum_{i=1}^N v_i 1_{A_i}$  for  $A_i$  in  $\mathcal{E}$  and  $v_i$  in  $V$ );
- we say that  $f$  is weakly measurable if, for every  $\varphi$  in  $V^*$ ,  $x \mapsto \langle f(x), \varphi \rangle_{V, V^*}$  is measurable;
- if  $V = U^*$  is the dual space of a Banach space  $U$ , we say that  $f$  is weakly-\* measurable if, for every  $\varphi$  in  $U$ ,  $x \mapsto \langle f(x), \varphi \rangle_{V, U}$  is measurable;
- we say that  $f$  is Borel measurable if, for every open set  $A$  in  $V$  (endowed with the strong topology),  $f^{-1}(A)$  is in  $\mathcal{E}$ .

The following result is morally Pettis measurability theorem. The present version is a consequence of [46, Chapter I Propositions 1.9 and 1.10].

**Proposition 5.1.** *Let  $V$  be a separable Banach space. Then the notions of strong measurability, weak measurability and Borel measurability coincide. They also coincide with the weak-\* measurability if moreover  $V$  is reflexive (in particular if  $V = \mathbb{R}$ ).*

As mentioned in the introduction, in the definition of  $L^p$  spaces we only consider two cases: (1)  $V = U^*$  is the dual space of a separable Banach space, where  $L^0(E; V)$  is the space of equivalent classes of weakly-\* measurable functions; (2)  $V$  is a separable Banach space, where  $L^0(E; V)$  is the space of equivalent classes of weakly (or strongly or Borel) measurable functions. In both cases, for any function  $f$  in  $L^0(E; V)$ , the function  $x \mapsto \|f(x)\|_V$  is measurable: in the case (1) because  $\|f(x)\| = \sup_{\varphi \in D} |\langle f(x), \varphi \rangle|$ , where  $D$  is a countable dense set of  $B_1^U$  (the unit centered ball in  $U$ ); in the case (2) as composition of the Borel map  $f$  and the continuous map  $\|\cdot\|_V$ . Therefore, it makes sense to define the spaces  $L^p(E; V)$  for  $1 \leq p \leq \infty$ .

**Proposition 5.2.** *Let  $D$  be a domain of  $\mathbb{R}^n$ . For every  $1 \leq p \leq \infty$ , the space  $L^p(E \times D, \mathcal{E} \otimes \mathcal{B}(D))$  is canonically embedded in  $L^p(E; L^p(D))$  (whose functions are weakly measurable for  $1 \leq p < \infty$ , weakly-\* measurable for  $p = \infty$ ). This embedding is a surjective isomorphism.*

*Proof.* The embedding result is easy to show using Fubini's theorem, we prove only the surjectivity. We start with the case  $p < \infty$ . To prove this, let  $F$  be an element (more precisely, a representative of an element) in  $L^p(E; L^p(D))$ . By Proposition 5.1,  $F$  is strongly measurable, i.e. there exists a sequence  $(F_n)_n$  of simple functions in  $L^p(E; L^p(D))$  which converges to  $F$  in  $L^p(D)$  for every  $x$  and, without loss of generality, in  $L^p(E; L^p(D))$ . We can write  $F_n$  as

$$F_n(x) = \sum_{k=1}^{N(n)} F_{n,k} 1_{A_{n,k}}(x)$$

for some measurable sets  $A_{n,k}$  and some elements  $F_{n,k}$  in  $L^p(D)$ . Now we define, for each  $n$ , the map  $G_n : E_x \times E_y \rightarrow V$  by

$$G_n(x, y) = \sum_{k=1}^{N(n)} G_{n,k}(y) 1_{A_{n,k}}(x),$$

where  $G_{n,k}$  is a representative of  $F_{n,k}$ . The function  $G_n$  is measurable in  $(x, y)$ ; since  $\|G_n - G_m\|_{L^p(E \times D)} = \|F_n - F_m\|_{L^p(E; L^p(D))}$ , the sequence  $(G_n)_n$  is Cauchy in  $L^p(E \times D)$ , therefore it converges to some  $G$  in  $L^p(E \times D)$ . In particular  $x \mapsto [y \mapsto G_n(x, y)]$  (where  $[y \mapsto G_n(x, y)]$  is the equivalence class of  $y \mapsto G_n(x, y)$ ) converges to  $x \mapsto [y \mapsto G(x, y)]$  in  $L^p(E; L^p(D))$ . It follows that  $x \mapsto [y \mapsto G(x, y)]$  coincides with  $F$   $\mu$ -a.e.. Hence  $G$  is the desired representative in  $L^p(E \times D)$  of  $F$ . This concludes the proof in the case  $p < \infty$ .

The case  $p = \infty$  can be reduced to the case  $p < \infty$ . Indeed, call  $(E_n)_n$  an increasing sequence of sets with finite measure and with  $E_n \nearrow E$ ; then any function  $f$  in  $L^\infty(E; L^\infty(D))$ , restricted to  $L^\infty(E_n; L^\infty(B_R \cap D))$ , is also a weakly measurable function in  $L^2(E_n; L^2(B_R \cap D))$ . Hence, it has a representative in  $L^2(E_n \times (B_R \cap D))$  and thus in  $L^2_{loc}(E \times D)$ , by arbitrariness of  $R$  and  $n$ , and this representative is essentially bounded.  $\square$

The following result is in Valadier [47], Theorem 2 (see also Theorem A.4).

**Theorem 5.3.** *Assume that  $\mu$  is finite and  $\mathcal{E}$  is  $\mu$ -complete. Let  $S$  be a metric  $\sigma$ -compact locally compact space and, for any  $R > 0$ , denote by  $L_R^\infty(E; \mathcal{M}_+(S))$  the subset of  $L^\infty(E; \mathcal{M}(S))$  of nonnegative measure-valued functions  $g$  with  $\|g\|_{L^\infty(E; \mathcal{M}(S))} \leq R$ . Then  $L_R^\infty(E; \mathcal{M}_+(S))$  is (embedded isomorphically in) a bounded sequentially weakly- $*$  closed subset of the dual space of  $L^1(E; C_0(S))$ . In particular, every sequence in  $L_R^\infty(E; \mathcal{M}_+(S))$  admits a subsequence converging weakly- $*$  to an element of  $L_R^\infty(E; \mathcal{M}_+(S))$ .*

We close with a remark on operations on measurable functions:

*Remark 5.4.* (i) Assume that  $V$  is the dual space of a separable space  $U$ . Let  $f : E \rightarrow V$  be a weakly- $*$  measurable map and let  $\varphi : E \rightarrow U$  be a (weakly or equivalently strongly) measurable map. Then the map  $x \mapsto \langle f(x), \varphi(x) \rangle_{V,U}$  is measurable. Indeed, if  $\varphi_k$  are simple measurable functions approximating everywhere  $\varphi$ , then  $x \mapsto \langle f(x), \varphi_k(x) \rangle_{V,U}$  are measurable functions approximating everywhere  $x \mapsto \langle f(x), \varphi(x) \rangle_{V,U}$ .

In particular, take a weakly- $*$  progressively measurable function  $f : [0, T] \times \Omega \rightarrow L_{x,\xi}^\infty$  and a  $\mathcal{P} \otimes \mathcal{B}(\mathbb{R}^d) \otimes \mathcal{B}(\mathbb{R})$ -measurable integrable function  $\varphi : [0, T] \times \Omega \times \mathbb{R}^d \times \mathbb{R} \rightarrow \mathbb{R}$ , so that  $\varphi : [0, T] \times \Omega \rightarrow L_{x,\xi}^1$  is a progressively measurable function. Then  $(t, \omega) \mapsto \langle f(t, \omega), \varphi(t, \omega) \rangle_{x,\xi}$  is a progressively measurable function.

(ii) An analogous property holds for bounded kinetic measures  $m$ . In this case one can consider a more general class of test functions. Let  $\varphi : [0, T] \times \Omega \times \mathbb{R}^d \times \mathbb{R} \rightarrow \mathbb{R}$  be a measurable function (not an equivalence class) such that: 1) for every  $(x, \xi)$ ,  $(t, \omega) \mapsto \varphi(t, \omega, x, \xi)$  is progressively measurable; 2) for a.e.  $\omega$ , the zero set being independent of  $(t, x, \xi)$ , and for every  $(x, \xi)$ ,  $t \mapsto \varphi(t, \omega, x, \xi)$  is càdlàg or càglàd; 3) for a.e.  $\omega$ , the zero set being independent of  $(t, x, \xi)$ , and for every  $t$ ,  $(x, \xi) \mapsto \varphi(t, \omega, x, \xi)$  is continuous; 4) for a.e.  $\omega$ , the map  $(t, x, \xi) \mapsto \varphi(t, \omega, x, \xi)$  is bounded. Then, for every fixed representative of  $m^\omega$ , the map

$$(t, \omega) \mapsto \int_{[0,t] \times \mathbb{R}^d \times \mathbb{R}} \varphi^\omega(r, x, \xi) m^\omega(r, x, \xi) dx d\xi dr$$

is progressively measurable and has a.e. càdlàg paths. The same result, replacing càdlàg with càglàd in the thesis, holds for  $\int_{[0,t] \times \mathbb{R}^d \times \mathbb{R}} \varphi^\omega m^\omega dx d\xi dr$ .

We prove this fact first for  $t \mapsto \varphi(t, \omega, x, \xi)$  càdlàg. We take a regular function  $\psi_n$  on  $\mathbb{R}^d \times \mathbb{R}$  with  $0 \leq \psi_n \leq 1$ ,  $\psi_n = 1$  on  $B_n$  and with support on  $B_{2n}$  and we define  $\varphi_n = n\psi_n \varphi *_t 1_{[0,1/n]}$ . Note that  $\varphi_n$  is a.s. in  $C_0([0, T] \times \mathbb{R}^d \times \mathbb{R})$ : indeed  $\varphi_n^\omega$  is Lipschitz continuous in  $t$  uniformly in  $(x, \xi)$  (by boundedness of  $\varphi^\omega$ ) and continuous in  $(x, \xi)$  at  $t$  fixed. Moreover, for every  $t$ ,  $\omega \mapsto \varphi_n^\omega|_{[0,t]} \in C_0([0, t] \times \mathbb{R}^d \times \mathbb{R})$  is weakly, hence strongly  $\mathcal{F}_{t+1/n}$ -measurable: indeed, for every finite signed measure  $\mu$  on  $[0, t] \times \mathbb{R}^d \times \mathbb{R}$ ,  $\omega \mapsto \int_{[0,t] \times \mathbb{R}^d \times \mathbb{R}} \varphi_n^\omega d\mu = n \int_{[0,t+1/n]} \int_{[(s-1/n) \vee 0, s \wedge t]} \varphi^\omega(s, x, \xi) d\mu(r, x, \xi) ds$  is  $\mathcal{F}_{t+1/n}$ -measurable by Fubini's theorem. The first part of this remark then gives that  $\int_{[0,t] \times \mathbb{R}^d \times \mathbb{R}} \varphi_n^\omega m^\omega dx d\xi dr$  is  $\mathcal{F}_{t+1/n}$ -measurable.

Now, for a.e.  $\omega$ ,  $\varphi_n^\omega$  converges everywhere to  $\varphi^\omega$ , by the càdlàg and continuity properties of  $\varphi$ . Therefore, for every  $t$  fixed,  $\int_{[0,t] \times \mathbb{R}^d \times \mathbb{R}} \varphi_n^\omega m^\omega dx d\xi dr$  is the a.s. limit of  $\int_{[0,t] \times \mathbb{R}^d \times \mathbb{R}} \varphi_n^\omega m^\omega dx d\xi dr$ . Hence  $\int_{[0,t] \times \mathbb{R}^d \times \mathbb{R}} \varphi_n^\omega m^\omega dx d\xi dr$  is  $\mathcal{F}_{t+1/n}$ -measurable for every  $n$  (recall that  $\mathcal{F}_0$  is complete) and thus  $\mathcal{F}_t$ -measurable. Moreover, for any fixed representative of  $m^\omega$ , for a.e.  $\omega$ , the map  $t \mapsto \int_{[0,t] \times \mathbb{R}^d \times \mathbb{R}} \varphi_n^\omega m^\omega dx d\xi dr$  is càdlàg. Therefore,  $(t, \omega) \mapsto \int_{[0,t] \times \mathbb{R}^d \times \mathbb{R}} \varphi_n^\omega m^\omega dx d\xi dr$  has the desired properties.

In the case  $t \mapsto \varphi(t, \omega, x, \xi)$  càglàd, the same proof applies but taking  $\varphi_n = \psi_n \varphi *_t 1_{[-1/n, 0]}$  (with  $\varphi_t$  extended as  $\varphi_0$  for  $-1/n \leq t < 0$ ).

## 6. POTENTIAL CONFLICT OF INTEREST STATEMENT

Conflict of Interest: Benjamin Gess has received research grants from the Max Planck Society via the Max Planck Research Group “Stochastic partial differential equations”. Mario Maurelli has received research grants from the European Research Council under the European Union’s Seventh Framework Program (FP7/2007-2013) / ERC grant agreement nr. 258237.

## REFERENCES

- [1] L. Ambrosio. Transport equation and Cauchy problem for  $BV$  vector fields. *Invent. Math.*, 158(2):227–260, 2004.
- [2] B. Andreianov, K. H. Karlsen, and N. H. Risebro. On vanishing viscosity approximation of conservation laws with discontinuous flux. *Netw. Heterog. Media*, 5(3):617–633, 2010.
- [3] B. Andreianov, K. H. Karlsen, and N. H. Risebro. A theory of  $L^1$ -dissipative solvers for scalar conservation laws with discontinuous flux. *Arch. Ration. Mech. Anal.*, 201(1):27–86, 2011.
- [4] B. Andreianov and D. Mitrović. Entropy conditions for scalar conservation laws with discontinuous flux revisited. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 32(6):1307–1335, 2015.
- [5] S. Attanasio and F. Flandoli. Zero-noise solutions of linear transport equations without uniqueness: an example. *C. R. Math. Acad. Sci. Paris*, 2009.
- [6] S. Attanasio and F. Flandoli. Renormalized solutions for stochastic transport equations and the regularization by bilinear multiplication noise. *Comm. Partial Differential Equations*, 36(8):1455–1474, 2011.
- [7] L. Beck, F. Flandoli, M. Gubinelli, and M. Maurelli. Stochastic odes and stochastic linear pdes with critical drift: regularity, duality and uniqueness. *arXiv:1401.1530*, 2014.
- [8] O. Butkovsky and L. Mytnik. Regularization by noise and flows of solutions for a stochastic heat equation. *arXiv:1610.02553*, 2016.
- [9] R. Catellier. Rough linear transport equation with an irregular drift. *Stoch. Partial Differ. Equ. Anal. Comput.*, 4(3):477–534, 2016.
- [10] R. Catellier and M. Gubinelli. Averaging along irregular curves and regularisation of ODEs. *Stochastic Process. Appl.*, 126(8):2323–2366, 2016.
- [11] G.-Q. Chen and B. Perthame. Well-posedness for non-isotropic degenerate parabolic-hyperbolic equations. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 20(4):645–668, 2003.
- [12] G. Crasta, V. De Cicco, and G. De Philippis. Kinetic formulation and uniqueness for scalar conservation laws with discontinuous flux. *Comm. Partial Differential Equations*, 40(4):694–726, 2015.
- [13] G. Crasta, V. De Cicco, G. De Philippis, and F. Ghiraldin. Structure of solutions of multi-dimensional conservation laws with discontinuous flux and applications to uniqueness. *Arch. Ration. Mech. Anal.*, 221(2):961–985, 2016.
- [14] A.-L. Dalibard. Kinetic formulation for heterogeneous scalar conservation laws. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 23(4):475–498, 2006.
- [15] F. Delarue and F. Flandoli. The transition point in the zero noise limit for a 1D Peano example. *Discrete Contin. Dyn. Syst.*, 34(10):4071–4083, 2014.
- [16] F. Delarue, F. Flandoli, and D. Vincenzi. Noise prevents collapse of Vlasov-Poisson point charges. *Comm. Pure Appl. Math.*, 67(10):1700–1736, 2014.
- [17] R. J. DiPerna and P.-L. Lions. Ordinary differential equations, transport theory and Sobolev spaces. *Invent. Math.*, 98(3):511–547, 1989.
- [18] R. Duboscq and A. Réveillac. Stochastic regularization effects of semi-martingales on random functions. *J. Math. Pures Appl. (9)*, 106(6):1141–1173, 2016.
- [19] S. N. Ethier and T. G. Kurtz. *Markov processes*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York, 1986. Characterization and convergence.
- [20] E. Fedrizzi and F. Flandoli. Pathwise uniqueness and continuous dependence of SDEs with non-regular drift. *Stochastics*, 83(3):241–257, 2011.
- [21] E. Fedrizzi and F. Flandoli. Noise prevents singularities in linear transport equations. *J. Funct. Anal.*, 264(6):1329–1354, 2013.
- [22] F. Flandoli. *Random perturbation of PDEs and fluid dynamic models*, volume 2015 of *Lecture Notes in Mathematics*. Springer, Heidelberg, 2011. Lectures from the 40th Probability Summer School held in Saint-Flour, 2010.
- [23] F. Flandoli, M. Gubinelli, and E. Priola. Well-posedness of the transport equation by stochastic perturbation. *Invent. Math.*, 180(1):1–53, 2010.



- [24] F. Flandoli, M. Gubinelli, and E. Priola. Full well-posedness of point vortex dynamics corresponding to stochastic 2D Euler equations. *Stochastic Process. Appl.*, 121(7):1445–1463, 2011.
- [25] F. Flandoli, M. Gubinelli, and E. Priola. Remarks on the stochastic transport equation with Hölder drift. *Rend. Semin. Mat. Univ. Politec. Torino*, 70(1):53–73, 2012.
- [26] F. Flandoli and M. Romito. Probabilistic analysis of singularities for the 3D Navier-Stokes equations. In *Proceedings of EQUADIFF, 10 (Prague, 2001)*, volume 127, pages 211–218, 2002.
- [27] F. Flandoli and M. Romito. Markov selections for the 3D stochastic Navier-Stokes equations. *Probab. Theory Related Fields*, 140(3-4):407–458, 2008.
- [28] P. Gassiat and B. Gess. Regularization by noise for stochastic Hamilton-Jacobi equations. *arXiv:1609.07074*, 2016.
- [29] B. Gess and S. Smith. Stochastic continuity equations with conservative noise. *arXiv:1710.04906*, 2017.
- [30] B. Gess and P. E. Souganidis. Scalar conservation laws with multiple rough fluxes. *Commun. Math. Sci.*, 13(6):1569–1597, 2015.
- [31] B. Gess and P. E. Souganidis. Long-time behavior and averaging lemmata for stochastic scalar conservation laws. *to appear in: Comm. Pure Appl. Math.*, pages 1–23, 2016.
- [32] I. Gyöngy and É. Pardoux. On the regularization effect of space-time white noise on quasi-linear parabolic partial differential equations. *Probab. Theory Related Fields*, 97(1-2):211–229, 1993.
- [33] S. N. Kružkov. First order quasilinear equations with several independent variables. *Mat. Sb. (N.S.)*, 81 (123):228–255, 1970.
- [34] N. V. Krylov. *Lectures on elliptic and parabolic equations in Sobolev spaces*, volume 96 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2008.
- [35] H. Kunita. Stochastic differential equations and stochastic flows of diffeomorphisms. In *École d’été de probabilités de Saint-Flour, XII—1982*, volume 1097 of *Lecture Notes in Math.*, pages 143–303. Springer, Berlin, 1984.
- [36] M. Lécureux-Mercier. Improved stability estimates for general scalar conservation laws. *J. Hyperbolic Differ. Equ.*, 8(4):727–757, 2011.
- [37] P.-L. Lions, B. Perthame, and P. E. Souganidis. Scalar conservation laws with rough (stochastic) fluxes. *Stoch. Partial Differ. Equ. Anal. Comput.*, 1(4):664–686, 2013.
- [38] P.-L. Lions, B. Perthame, and P. E. Souganidis. Scalar conservation laws with rough (stochastic) fluxes: the spatially dependent case. *Stoch. Partial Differ. Equ. Anal. Comput.*, 2(4):517–538, 2014.
- [39] P.-L. Lions, B. Perthame, and E. Tadmor. A kinetic formulation of multidimensional scalar conservation laws and related equations. *J. Amer. Math. Soc.*, 7(1):169–191, 1994.
- [40] P. A. Loeb and E. Talvila. Lusin’s theorem and Bochner integration. *Sci. Math. Jpn.*, 60(1):113–120, 2004.
- [41] M. Mariani. Large deviations principles for stochastic scalar conservation laws. *Probab. Theory Related Fields*, 147(3-4):607–648, 2010.
- [42] O. Menoukeu-Pamen, T. Meyer-Brandis, T. Nilssen, F. Proske, and T. Zhang. A variational approach to the construction and Malliavin differentiability of strong solutions of SDE’s. *Math. Ann.*, 357(2):761–799, 2013.
- [43] S.-E. A. Mohammed, T. K. Nilssen, and F. N. Proske. Sobolev differentiable stochastic flows for SDEs with singular coefficients: applications to the transport equation. *Ann. Probab.*, 43(3):1535–1576, 2015.
- [44] P. E. Protter. *Stochastic integration and differential equations*, volume 21 of *Applications of Mathematics (New York)*. Springer-Verlag, Berlin, second edition, 2004. Stochastic Modelling and Applied Probability.
- [45] D. Revuz and M. Yor. *Continuous martingales and Brownian motion*, volume 293 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, third edition, 1999.
- [46] N. N. Vakhania, V. I. Tarieladze, and S. A. Chobanyan. *Probability distributions on Banach spaces*, volume 14 of *Mathematics and its Applications (Soviet Series)*. D. Reidel Publishing Co., Dordrecht, 1987. Translated from the Russian and with a preface by Wojbor A. Woyczynski.
- [47] M. Valadier. Young measures. In *Methods of nonconvex analysis (Varenna, 1989)*, volume 1446 of *Lecture Notes in Math.*, pages 152–188. Springer, Berlin, 1990.

MAX-PLANCK INSTITUTE FOR MATHEMATICS IN THE SCIENCES, INSELSTRASSE 22, 04103 LEIPZIG,  
GERMANY

*E-mail address:* [bgess@mis.mpg.de](mailto:bgess@mis.mpg.de)

WEIERSTRASS INSTITUTE FOR APPLIED ANALYSIS AND STOCHASTICS, MOHRENSTRASSE 39, 10117  
BERLIN, GERMANY, & INSTITUT FÜR MATHEMATIK, TECHNISCHE UNIVERSITÄT BERLIN, STRASSE  
DES 17. JUNI 136, 10623 BERLIN, GERMANY

*E-mail address:* [maurelli@wias-berlin.de](mailto:maurelli@wias-berlin.de)