SCALAR CONSERVATION LAWS WITH ROUGH FLUX AND STOCHASTIC FORCING

MARTINA HOFMANOVÁ

Abstract. In this paper, we study scalar conservation laws where the flux is driven by a geometric H"older p-rough path for some $p \in (2,3)$ and the forcing is given by an Itô stochastic integral driven by a Brownian motion. In particular, we derive the corresponding kinetic formulation and define an appropriate notion of kinetic solution. In this context, we are able to establish well-posedness, i.e. existence, uniqueness and the $L^1$-contraction property that leads to continuous dependence on initial condition. Our approach combines tools from rough path analysis, stochastic analysis and theory of kinetic solutions for conservation laws. As an application, this allows to cover the case of flux driven for instance by another (independent) Brownian motion enhanced with Lévy's stochastic area.

1. Introduction

The goal of the present paper is to develop a well-posedness theory for the following scalar rough conservation law

$$
\begin{align*}
\frac{du}{dt} + \text{div}\left(A(x, u)\right) &= g(x, u) \, dz, & t \in (0, T), \ x \in \mathbb{R}^N, \\
\ u(0) &= u_0,
\end{align*}
$$

where $z = (z^1, \ldots, z^M)$ is a deterministic rough driving signal and $W = (W^1, \ldots, W^K)$ is a Wiener process and the stochastic integral is understood in the Itô sense. The coefficients $A : \mathbb{R}^N \times \mathbb{R} \to \mathbb{R}^{N \times M}$, $g : \mathbb{R}^N \times \mathbb{R} \to \mathbb{R}^K$ satisfy a sufficient regularity assumption introduced in Section 2. The above equation can be rewritten by using the Einstein summation convention as follows

$$
\begin{align*}
\frac{du}{dt} + \partial_{x^i}\left(A_{ij}(x, u)\right) \, dz^j &= g_k(x, u) \, dW^k, & t \in (0, T), \ x \in \mathbb{R}^N, \\
\ u(0) &= u_0.
\end{align*}
$$

As an application of our analysis, one can replace $z$, for instance, by another Brownian motion $B$, which is independent of $W$, and give meaning to

$$
\begin{align*}
\frac{du}{dt} + \text{div}\left(A(x, u)\right) \circ dB &= g(x, u) \, dW, & t \in (0, T), \ x \in \mathbb{R}^N, \\
\ u(0) &= u_0.
\end{align*}
$$

Conservation laws and related equations have been paid an increasing attention lately and have become a very active area of research, counting nowadays quite a number of results for deterministic and stochastic setting, that is for conservation laws either of the form

$$
\begin{align*}
\partial_t u + \text{div}\left(A(u)\right) &= 0 \\
\ u(0) &= u_0.
\end{align*}
$$

or

$$
\begin{align*}
\frac{du}{dt} + \text{div}\left(A(u)\right) \, dt &= g(x, u) \, dW
\end{align*}
$$

Date: March 13, 2015.

2010 Mathematics Subject Classification. 60H15, 35R60, 35L65.

Key words and phrases. scalar conservation laws, rough paths, kinetic formulation, kinetic solution, BGK approximation, method of characteristics.
where the Itô stochastic forcing is driven by a finite- or infinite-dimensional Wiener process (see [1], [2], [3], [5], [6], [8], [9], [10], [11], [14], [15], [20], [21], [22], [23], [24], [25], [28], [29], [30], [34], [37], [38]). Since the theory of rough paths viewed as a tool that allows for deterministic treatment of stochastic differential equations has been of growing interest recently, several attempts have already been made to extend this theory to conservation laws as well. First, Lions, Perthame and Souganidis (see [26], [27]) developed a pathwise approach for
\[ du + \text{div} \left( A(x, u) \right) \circ dW = 0 \]
where \( W \) is a continuous signal and \( \circ \) stands for the Stratonovich product in the Brownian case, then Friz and Gess (see [16]) studied
\[ du + \text{div} f(t, x, u)dt = F(t, x, u)dt + \Lambda_k(x, u, \nabla u)dz^k \]
where \( \Lambda_k \) that is affine linear in \( u \) and \( \nabla u \) and \( z = (z^1, \ldots, z^K) \) is a rough driving signal and Gess and Souganidis [19] considered
\[ du + \text{div} \left( A(x, u) \right) dz = 0 \]
where \( z = (z^1, \ldots, z^M) \) is a geometric \( \alpha \)-Hölder path and in [18] they studied the long-time behavior for \( z \) being a Brownian motion.

In order to find a suitable concept of solution for problems of the form (1.1), it was observed already some time ago that, on the one hand, classical \( C^1 \) solutions do not exist in general and, on the other hand, weak or distributional solutions lack uniqueness. The first claim is a consequence of the fact that any smooth solution has to be constant along characteristic lines, which can intersect in finite time (even in the case of smooth data) and shocks can be produced. The second claim demonstrates the inconvenience that often appears in the study of PDEs and SPDEs: the usual way of weakening the equation leads to the occurrence of nonphysical solutions and therefore additional assumptions need to be imposed in order to select the physically relevant ones and to ensure uniqueness. Hence one needs to find some balance that allows to establish existence of a unique (physically reasonable) solution.

Towards this end, we pursue the ideas of kinetic approach, a concept of solution that was first introduced by Lions, Perthame, Tadmor [29] for deterministic hyperbolic conservation laws and further studied in [2], [6], [23], [28], [29], [30], [34]. This direction also appears in several works on stochastic conservation laws and degenerate parabolic SPDEs, see [8], [9], [10], [11], [20], [21] and in the (rough) pathwise works [18], [19], [26], [27]. In comparison to the notion of entropy solution introduced by Kružkov [25] and further developed e.g. in [1], [3], [15], [24], [38], kinetic solutions are more general in the sense that they are well defined even in situations when neither the original conservation law nor the corresponding entropy inequalities can be understood in the sense of distributions which is part of the definition of entropy solution. Usually this happens due to lack of integrability of the flux and entropy-flux terms, e.g. \( A(u) \notin L^1_{\text{loc}} \). Therefore, further assumptions on initial data or the flux function \( A \) are in place in order to overcome this issue and remain in the entropy setting. It will be seen later on that no such restrictions are necessary in the kinetic approach as the equation that is actually to be solved – the so-called kinetic formulation, see (2.3) – is in fact linear. In addition, various proofs simplify as methods for linear PDEs are available.

Let us now shortly present the main ideas of our approach. Apart from the above mentioned difficulties there is another one that originates in the low regularity of driving signals and solution. Namely, the corresponding rough integrals are not well defined so we present a formulation, see (2.4), that does not include any rough path driven terms and therefore provides a suitable notion of kinetic solution in this context. To this end, we adapt the ideas of [26], [27], [19] where the authors introduced a method of modified test functions to eliminate the rough terms. Our method then combines this approach with the ideas for stochastic conservation laws treated in [10], [20]. As usual for this class of problems, we define a second notion of solution – a generalized kinetic solution – which, roughly speaking, takes values in the set of Young measures. The general idea is that in order to get existence of such a solution, weak convergence (in some \( L^p \) space) of the corresponding approximations is sufficient.
which allows for an easier proof. The key result is then the Reduction Theorem that states that any generalized kinetic solution is actually a kinetic one, that is, the Young measure at hand is a parametrized Dirac mass.

Concerning the existence part, we make use of the so-called Bhatnagar-Gross-Krook approximation (BGK for short) which allows to describe the conservation law as the hydrodynamic limit of the corresponding BGK model, as the microscopic scale \( \varepsilon \) vanishes. This is nowadays a standard tool in the deterministic setting where the literature is quite extensive (see [2], [23], [28], [29], [30], [34]). Even though the general concept is analogous in the stochastic case, the techniques are significantly different, the result was established recently by the author [20]. To be more precise, the key point is to solve the corresponding characteristic system which for the general case of (1.1) reads as follows

\[
\begin{align*}
\frac{d\phi_i^0}{dt} &= -\partial_x A_{ij}(\phi_t) \, dz^j + g_k(\phi_t) \, dW^k, \\
\frac{d\phi_i^j}{dt} &= \partial_u A_{ij}(\phi_t) \, dz^j, \quad i = 1, \ldots, N.
\end{align*}
\]

(1.5)

Already at this level one can understand the difficulties coming from the complex structure of (1.1). Namely, in the case of (1.3) the characteristic system reduces to a set of independent equations (note that the flux function \( A \) is independent of the space variable)

\[
\begin{align*}
\frac{d\phi_i^0}{dt} &= 0 \\
\frac{d\phi_i^j}{dt} &= \partial_u A_{ij}(\phi_0^t) \, dt, \quad i = 1, \ldots, N,
\end{align*}
\]

which can be solved immediately and as a consequence the method simplifies. For the stochastic case (1.4) we obtain

\[
\begin{align*}
\frac{d\phi_i^0}{dt} &= g_k(\phi_t) \, dW^k \\
\frac{d\phi_i^j}{dt} &= \sum_{j=1}^{M} \partial_u A_{ij}(\phi_0^t) \, dt, \quad i = 1, \ldots, N,
\end{align*}
\]

hence the first coordinate of the characteristic curve is governed by an SDE and further difficulties arise due to randomness (see [20]). If \( A \) is also \( x \)-dependent we observe an additional term in the first equation of the characteristic system (1.5), however, let us point out that there is a major difference between this and the term coming from the forcing. In particular, if \( g = 0 \) then the flow of diffeomorphisms generated by (1.5) is volume preserving as can be seen easily by calculating divergence of the corresponding vector field. This does not hold true anymore if forcing in nonconservative form is present in the equation, unless \( g \) is independent of the solution, i.e. we consider additive noise in (1.1).

The method of BGK approximation has another advantage especially when dealing with rough path driven conservation laws. As already mentioned above, the problem boils down to solving the characteristic system (1.5), i.e. an ordinary (stochastic, rough) differential equation and the theory for those problems is well-established, unlike the theory for rough path driven equations in infinite dimension, i.e. rough PDEs. Furthermore, the BGK model also provides an explicit formula for the approximate solutions and therefore the necessary estimates independent of the microscopic scale \( \varepsilon \) come rather naturally.

The exposition is organized as follows. In Section 2, we introduce the basic setting, define the notion of kinetic solution and state our main result, Theorem 2.7. In order to make the paper more self-contained, Section 3 provides a brief overview of the relevant concepts from rough path theory. Section 4 is devoted to the proof of uniqueness, reduction of a generalized kinetic solution to a kinetic one and the \( L^1 \)-contraction property, Theorem 4.6 and Corollary 4.7. The remainder of the paper deals with the existence part of Theorem 2.7 which is divided into several parts and finally established through Theorem 5.9. Subsections 5.1 and 5.2 give a rough-pathwise existence of unique solutions to the BGK approximation. In these two sections we work with a fixed realization of the driving signals.
or, to be more precise, a fixed realization of a joint lift of \((t, z, W)\). In Subsection 5.3, the stochastic approach is resumed and we pass to the limit and obtain a kinetic solution to (1.1).

2. Definitions and the main result

Let us now introduce the precise setting of (1.1). We work on a finite time interval \([0, T]\), \(T > 0\) and on the whole space concerning the space variable \(x \in \mathbb{R}^N\). Throughout the paper, we assume that \(z\) can be lifted to a geometric Hölder \(p\)-rough path for some \(p \in (2, 3)\) and denote its lift by \(z = (z^1, z^2)\). Next we consider the following joint lift of \((z, W)\): we define \(\Lambda = (\Lambda^1, \Lambda^2)\) by

\[
\Lambda^1_{i,t} = \begin{cases} 
    z^1_i, & \text{if } i \in \{1, \ldots, M\}, \\
    W^1_{i-M}, & \text{if } i \in \{M+1, \ldots, M+K\}, 
\end{cases}
\]

\[
\Lambda^2_{i,j,t} = \begin{cases} 
    z^2_{i,j}, & \text{if } i, j \in \{1, \ldots, M\}, \\
    \int_0^t W^2_{i-M} \circ dW^2_{j-M}, & \text{if } i, j \in \{M+1, \ldots, M+K\}, \\
    \int_0^t z^2_{i,j} dW^2_{j-M}, & \text{if } i \in \{1, \ldots, M\}, j \in \{M+1, \ldots, M+K\}, \\
    z^2_{i,M+K} - \int_0^t z^2_{i,j} dW^2_{j-M}, & \text{if } i \in \{M+1, \ldots, M+K\}, j \in \{1, \ldots, M\}.
\end{cases}
\]

It was shown in [12] that such a stochastic process exists and could be considered as the canonical joint lift of \(z\) and \(W\). Note that although in the original equation (1.1) we consider Itô stochastic integral, the lift of \(W\) used in the construction of \(\Lambda\) above corresponds to the Stratonovich version.

Regarding the coefficients in (1.1), let us fix the following notation. Let

\[a = (a_{ij}) = (\partial_u A_{ij} : \mathbb{R}^N \times \mathbb{R} \rightarrow \mathbb{R}^{N \times M}),\]

\[b = (b_j) = (\text{div}_x A_j : \mathbb{R}^N \times \mathbb{R} \rightarrow \mathbb{R}^M),\]

and assume that \(a, b \in \text{Lip}^{\gamma+2}\) and \(g \in \text{Lip}^{\gamma+3}\) for some \(\gamma > p\). Here we adopt the notation of [17, Definition 10.2], namely, a mapping \(V : \mathbb{R}^k \rightarrow \mathbb{R}^d\) belongs to \(\text{Lip}^\beta\) provided it is bounded, \([\beta]\)-times continuously differentiable with bounded derivatives of all orders and its \([\beta]\)th derivative is \([\beta]\)-Hölder continuous.

Furthermore, we suppose that

\[b(x, 0) = 0 \quad g(x, 0) = 0 \quad \forall x \in \mathbb{R}^N,\]

and denote

\[C^2(x, \xi) = \sum_{k=1}^{K} |g_k(x, \xi)|^2, \quad \forall x \in \mathbb{R}^N, \xi \in \mathbb{R}.\]

2.1. Definitions. As the next step, let us introduce the kinetic formulation of (1.1) as well as the basic definitions concerning the notion of kinetic solution. The motivation behind this approach is given by the nonexistence of a strong solution and, on the other hand, the nonuniqueness of weak solutions, even in simple cases. The idea is to establish an additional criterion – the kinetic formulation – which is automatically satisfied by any strong solution to (1.1) and which permits to ensure the well-posedness.

We start with the definition of kinetic measure.

**Definition 2.1** (Kinetic measure). A mapping \(m\) from \(\Omega\) to \(\mathcal{M}_k^+(\mathbb{R}^N \times \mathbb{R})\), the set of nonnegative bounded measures over \([0, T] \times \mathbb{R}^N \times \mathbb{R}\), is said to be a kinetic measure provided

(i) \(m\) is measurable in the following sense: for each \(\rho \in C_0([0, T] \times \mathbb{R}^N \times \mathbb{R})\) the mapping \(m(\rho) : \Omega \rightarrow \mathbb{R}\) is measurable,

(ii) for any \(\phi \in C_0(\mathbb{R}^N \times \mathbb{R})\), \(t \mapsto m(1_{[0,t]}\phi)\) is progressively measurable.
The kinetic formulation corresponding to the conservation law at hand is given as follows

\begin{equation}
\begin{aligned}
dF + \nabla F \cdot a \, dz - \partial_t F \, b \, dz &= -\partial_t \langle G^2 \partial_z F \rangle \, dt + \partial_z m, \\
F(0) &= F_0,
\end{aligned}
\end{equation}

where $F = \mathbf{1}_{u > \xi}$ and $m$ is a kinetic measure. However, since the expected regularity of solutions is very low and consequently the rough path driven integrals are not well defined, it is necessary to define a suitable notion of weak solution to this problem. This leads us to the notion of kinetic solution to rough path driven conservation laws that we introduce in this work. Note that it is a consistent extension of the corresponding notion of kinetic solution for the case of smooth driving signal $z$, for further discussion on this subject we refer the reader to Subsection 2.2.

**Definition 2.2** (Kinetic solution). Let $u_0 \in L^1 \cap L^2(\Omega \times \mathbb{R}^N)$. Then a progressively measurable

\begin{equation}
u \in L^1(\Omega; L^\infty(0, T; L^1(\mathbb{R}^N))) \cap L^2(\Omega; L^2(0, T; L^2(\mathbb{R}^N)))
\end{equation}

is said to be a kinetic solution to (1.1) with initial datum $u_0$ provided there exists a kinetic measure $m$ such that the pair $(F = \mathbf{1}_{u > \xi}, m)$ satisfies, for all $\phi \in C^1_c(\mathbb{R}^N \times \mathbb{R})$ and $\alpha \in C^1_c((0, T))$,

\begin{equation}
\begin{aligned}
\int_0^T \langle F(t), \phi(\theta_t) \rangle \partial_t \alpha(t) \, dt + \langle F_0, \phi \rangle \alpha(0) \\
&= \int_0^T \langle \partial_t F(t) g, \phi(\theta_t) \rangle \alpha(t) \, dt - \frac{1}{2} \int_0^T \langle \partial_t (G^2 \partial_z F(t)), \phi(\theta_t) \rangle \alpha(t) \, dt + m(\alpha \partial_z \phi(\theta_t)),
\end{aligned}
\end{equation}

where $\theta = (\theta^t, \theta^x)$ is the inverse flow corresponding to

\begin{equation}
\begin{aligned}
d\pi^t &= -b(\pi_t) \, d\mathbf{z} \\
d\pi^x &= a(\pi_t) \, d\mathbf{z}.
\end{aligned}
\end{equation}

To be more precise, it follows from [17, Proposition 11.11] that under our assumptions (2.5) possesses a unique solution $\pi$ that defines a flow of $C^2$-diffeomorphisms. We denote by $\pi_{s,t}(x, \xi)$ the solution of (5.5) starting from $(x, \xi)$ at time $s$. To simplify the notation, we write $\pi_t$ instead of $\pi_{0,t}$ and we denote the corresponding inverse flow by $\theta$. Then $\theta_{t,s} = \pi_{t,s}^{-1}$ is the unique solution to the time-reversed problem

\begin{equation}
\begin{aligned}
d\theta^t_{s,s} &= -b(\theta_{t,s}) \, d\mathbf{z}^s \\
d\theta^x_{s,s} &= a(\theta_{t,s}) \, d\mathbf{z}^s,
\end{aligned}
\end{equation}

where $\mathbf{z}^s(\cdot) = \mathbf{z}(\cdot - s)$ is the time-reversed path to $\mathbf{z}$.

We proceed with a reminder of Young measures and the related definition of kinetic function that will eventually lead to the notion of generalized kinetic solution, see Definition 2.6. The concept of Young measures was developed in [39] as a technical tool for describing composite limits of nonlinear functions with weakly convergent sequences, for further reading we refer the reader e.g. to [33].

In what follows, we denote by $\mathcal{P}_1(\mathbb{R})$ the set of probability measures on $\mathbb{R}$.

**Definition 2.3** (Young measure). Let $(X, \lambda)$ be a finite measure space. A mapping $\nu : X \to \mathcal{P}_1(\mathbb{R})$ is called a Young measure provided it is weakly measurable, that is, for all $\phi \in C_b(\mathbb{R})$ the mapping $z \mapsto \nu_z(\phi)$ from $X$ to $\mathbb{R}$ is measurable. A Young measure $\nu$ is said to vanish at infinity if

\[ \int_X \int_{\mathbb{R}} |\xi| \, d\nu_z(\xi) \, d\lambda(z) < \infty. \]

**Definition 2.4** (Kinetic function). Let $(X, \lambda)$ be a finite measure space. A measurable function $f : X \times \mathbb{R} \to [0, 1]$ is called a kinetic function if there exists a Young measure $\nu$ on $X$ that vanishes at infinity such that for a.e. $z \in X$ and for all $\xi \in \mathbb{R}$

\[ f(z, \xi) = \nu_z(\xi, \infty). \]
Lemma 2.5 (Compactness of Young measures). Let $(X,\lambda)$ be a finite measure space such that $L^1(X)$ is separable. Let $(\nu^n)$ be a sequence of Young measures on $X$ such that for some $p \in [1, \infty)$

$$\sup_{n \in \mathbb{N}} \int_X \int_{\mathbb{R}} |\xi|^p \, d\nu^n(\xi) \, d\lambda(z) < \infty.$$ 

Then there exists a Young measure $\nu$ on $X$ and a subsequence still denoted by $(\nu^n)$ such that for all $h \in L^1(X)$ and all $\phi \in C_b(\mathbb{R})$

$$\lim_{n \to \infty} \int_X h(z) \int_{\mathbb{R}} \phi(\xi) \, d\nu^n(\xi) \, d\lambda(z) = \int_X h(z) \int_{\mathbb{R}} \phi(\xi) \, d\nu(\xi) \, d\lambda(z).$$

Moreover, if $f_n, \phi \in C(\mathbb{R})$, $n \in \mathbb{N}$, and $f$ are the kinetic functions corresponding to $\nu^n, n \in \mathbb{N}$, and $\nu$, respectively, then

$$f_n \xrightarrow{w^*} f \quad \text{in} \quad L^\infty(X \times \mathbb{R}).$$

Proof. The proof is classical and can be found for instance in [10, Theorem 5, Corollary 6].

Definition 2.6 (Generalized kinetic solution). Let $F_0 : \Omega \times \mathbb{R}^N \times \mathbb{R} \to [0,1]$ be a kinetic function. A progressively measurable function $F : \Omega \times [0,T] \times \mathbb{R}^N \times \mathbb{R} \to [0,1]$ is called a generalized solution to (1.1) with initial datum $F_0$ if $F(t)$ is a kinetic function for a.e. $t \in [0,T]$, there exists $C > 0$ such that

$$(2.7) \quad \mathbb{E} \, \text{ess sup}_{0 \leq t \leq T} \int_{\mathbb{R}^N} |\xi| \, d\nu(t)(\xi) \, dx + \mathbb{E} \int_0^T \int_{\mathbb{R}^N} |\xi|^2 \, d\nu(t)(\xi) \, dx \leq C,$$

where $\nu = -\partial_t F$, and if there exists a kinetic measure $m$ such that (2.4) holds true.

2.2. Kinetic solutions in the case of smooth driver $z$. For the sake of completeness, let us now present a formal derivation of (2.3) in case of sufficiently smooth solution to the conservation law (1.1) driven by a smooth path $z$ and a Wiener process $W$. The aim is to derive the equation satisfied by $1_{u>\xi}$. Towards this end, observe that for $\phi \in C^\infty(\mathbb{R})$

$$\langle 1_{u>\xi}, \phi' \rangle = \phi(u) - \phi(-\infty)$$

and therefore using the Itô formula, we obtain

$$d \langle 1_{u>\xi}, \phi' \rangle = -\phi'(u) \, d\langle A(x,u) \rangle + \phi'(u)g(x,u) \, dW + \frac{1}{2} \phi''(u)G^2(x,u) \, dt$$

$$= -\phi'(u)a(x,u) \cdot \nabla u \, dz - \phi'(u)b(x,u) \, dz + \phi'(u)g(x,u) \, dW + \frac{1}{2} \phi''(u)G^2(x,u) \, dt$$

$$= -\langle a(x,\xi) \nabla 1_{u>\xi}, \phi' \rangle \, dz - \langle b(x,\xi)\delta_{u=\xi}, \phi' \rangle \, dz + \langle g(x,\xi)\delta_{u=\xi}, \phi' \rangle \, dW + \frac{1}{2} \langle G^2(x,\xi)\delta_{u=\xi}, \phi'' \rangle \, dt$$

$$= -\langle a(x,\xi) \nabla 1_{u>\xi}, \phi' \rangle + \langle b(x,\xi)\delta_{1_{u>\xi}}, \phi' \rangle \, dz - \langle g(x,\xi)\partial_\xi 1_{u>\xi}, \phi' \rangle \, dW$$

$$+ \frac{1}{2} \langle \partial_\xi (G^2(x,\xi)\delta_{1_{u>\xi}}), \phi' \rangle \, dt.$$
2.3. The main result. Our main result reads as follows.

**Theorem 2.7.** Let \( u_0 \in L^1 \cap L^2(\Omega \times \mathbb{R}^N) \). Under the above assumptions, the following statements hold true:

(i) There exists a unique kinetic solution to (1.1).

(ii) Any generalized kinetic solution is actually a kinetic solution, that is, if \( F \) is a generalized kinetic solution to (1.1) with initial datum \( 1_{u_0 > \xi} \) then there exists a kinetic solution \( u \) to (1.1) with initial datum \( u_0 \) such that \( F = 1_{u_0 > \xi} \) for a.e. \((t,x,\xi)\).

(iii) If \( u_1, u_2 \) are kinetic solutions to (1.1) with initial data \( u_{1,0} \) and \( u_{2,0} \), respectively, then for a.e. \( t \in [0,T] \)

\[
\mathbb{E}\|u_1(t) - u_2(t)\|_{L^1} \leq \mathbb{E}\|u_{1,0} - u_{2,0}\|_{L^1}.
\]

2.4. Application to conservation laws with stochastic flux and forcing. It is immediately possible to apply our theory to conservation laws of the form (1.2), where \( B \) is a stochastic process that is independent on \( W \) and that can be enhanced to a stochastic process \( \bar{B} \) for which almost every realization is a geometric \( p \)-Hölder rough path. For instance, one may consider another Brownian motion but also more general Gaussian or Markov processes (the reader is referred to [17, Part 3] for further examples).

Let \( B \) be a Brownian motion defined on a stochastic basis \((\bar{\Omega}, \bar{\mathcal{F}}, (\bar{\mathcal{F}}_t), \bar{\mathbb{P}})\) and let \( B(\bar{\omega}) \) be a realization of its (Stratonovich) lift. For \( \bar{\omega} \) fixed, we set

\[
z = (z^1, z^2) := (B^1(\bar{\omega}), B^2(\bar{\omega}))
\]

and Theorem 2.7 yields the existence of a unique kinetic solution

\[
u(\bar{\omega}) \in L^1(\bar{\Omega}; L^\infty(0,T; L^1(\mathbb{R}^N))) \cap L^2(\bar{\Omega}; L^2(0,T; L^2(\mathbb{R}^N)))
\]

together with the corresponding \( L^1 \)-contraction property. Let us now set

\[
\bar{\Omega} = \Omega \times \Omega, \quad \bar{\mathcal{F}} = \mathcal{F} \otimes \mathcal{F}, \quad \bar{\mathcal{F}}_t = \mathcal{F}_t \otimes \mathcal{F}_t, \quad \bar{\mathbb{P}} = \mathbb{P} \otimes \mathbb{P}
\]

and

\[
\bar{B}(\bar{\omega}, \omega) := B(\bar{\omega}), \quad \bar{W}(\bar{\omega}, \omega) := W(\omega).
\]

Both these processes are Brownian motions on \((\bar{\Omega}, \bar{\mathcal{F}}, (\bar{\mathcal{F}}_t), \bar{\mathbb{P}})\) and if \( \bar{B} \) denotes the (Stratonovich) lift of \( \bar{B} \) then

\[
\bar{B}(\bar{\omega}, \omega) = B(\bar{\omega}).
\]

Moreover, standard results in rough path theory guarantee that the RDE solutions to (2.5) and (2.6) driven by \( z = \bar{B}(\bar{\omega}, \omega) \) coincide \( \bar{\mathbb{P}} \)-a.s. with the corresponding SDE solutions and in particular they are \((\bar{\mathcal{F}}_t)\)-adapted. It remains to verify that the Itô integral in the definition of kinetic solution (2.4), which is now constructed for a.e. \( \bar{\omega} \) as an Itô integral on \((\bar{\Omega}, \bar{\mathcal{F}}, (\bar{\mathcal{F}}_t), \bar{\mathbb{P}})\), can be extended to \((\bar{\Omega}, \bar{\mathcal{F}}, (\bar{\mathcal{F}}_t), \bar{\mathbb{P}})\). But that follows directly from its construction: the corresponding Riemann sums converge in \( \mathbb{P} \) for a.e. \( \bar{\omega} \), therefore by Fubini's theorem they also converge in \( \bar{\mathbb{P}} \).

As a consequence, (1.2) or more precisely its kinetic formulation (2.4) is solved in the natural (stochastic) sense and Theorem 2.7 applies pathwise in \( \bar{\omega} \).

3. Elements of rough path theory

In this section, we recall some basic notions and results from rough path theory which are used throughout the paper. For the general exposition we refer the reader to [17], [31], [32].

Let \( G^{[p]}(\mathbb{R}^d) \) denote the step-[\( p \)] nilpotent free group over \( \mathbb{R}^d \) and let us consider the following rough differential equation (RDE)

\[
dy = V(y)dx,
\]

\[
y(0) = y_0,
\]

(3.1)
where $V = (V_1, \ldots, V_d)$ is a family of sufficiently smooth vector fields on $\mathbb{R}^c$ and $x : [0, T] \to G^{[p]}(\mathbb{R}^d)$ is a geometric Hölder $p$-rough path, namely, it is a path with values in $G^{[p]}(\mathbb{R}^d)$ satisfying

$$ \|x\|_{\mathbb{H}^{\text{Höld}}[0,T]} = \sup_{0 \leq s < t \leq T} \frac{\|x_{s,t}\|}{|s - t|^{1/p}} < \infty. $$

We denote by $C^{0,\mathbb{H}^{\text{Höld}}}[0,T]; G^{[p]}(\mathbb{R}^d))$ the space of all geometric Hölder p-rough paths.

Definition of a solution to problems of the form (3.1) is based on Davie’s lemma (see [17, Lemma 10.7]) that gives uniform estimates for ODE solutions depending only on the rough path regularity (e.g. $p$-variation) of the canonical lift of a regular driving signal $x : [0, T] \to \mathbb{R}^d$. As a consequence, a careful limiting procedure yields a reasonable notion of solution to (3.1). To be more precise, we define a solution to (3.1) as follows.

**Definition 3.1.** Let $x \in C^{0,\mathbb{H}^{\text{Höld}}}[0,T]; G^{[p]}(\mathbb{R}^d))$ be a geometric Hölder $p$-rough path and suppose that $(x^n)$ is a sequence of Lipschitz paths with the corresponding step-$[p]$ lifts denoted by $x^n$, i.e.

$$ x^n_t := S_{[p]}(x^n)_t = \left(1, \int_{0 < v < t} dx^n_v, \ldots, \int_{0 < v_1 < \cdots < v_{[p]} < t} dx^n_{v_1} \otimes \cdots \otimes dx^n_{v_{[p]}} \right), $$

such that

$$ x^n \to x $$

uniformly on $[0, T]$ and $\sup_n \|x^n\|_{\mathbb{H}^{\text{Höld}}[0,T]} < \infty$. We say that $y \in C([0,T]; \mathbb{R}^c)$, also denoted by $\pi_{(V)}(0, y_0; x^n)$, is a solution to (3.1) provided it is a limit point (in uniform topology on $[0, T]$) of

$$ \{\pi_{(V)}(0, y_0; x^n); n \geq 1\} $$

where $\pi_{(V)}(0, y_0; x^n)$ denotes the solution to the ODE

$$ dy^n = V(y^n)dx^n, $$

$$ y^n(0) = y_0. $$

Under a sufficient regularity assumption upon the collection of vector fields $V$, there exists a unique solution to (3.1) which defines a flow of diffeomorphisms.

**Theorem 3.2.** Let $p \geq 1$. Let $x$ be a geometric Hölder $p$-rough path and suppose that the vector fields $V = (V_1, \ldots, V_d)$ are Lip$^{\gamma+k}$ for some $\gamma > p$. Then the following holds true.

(i) There exists a unique solution to (3.1), say $\pi_{(V)}(0, y_0; x) \in C([0,T]; \mathbb{R}^c)$, and the map

$$ \phi : (t, y) \in [0,T] \times \mathbb{R}^c \mapsto \pi_{(V)}(0, y; x)_t \in \mathbb{R}^c $$

is a flow of Lip$^k$-diffeomorphisms.

(ii) There exists $C > 0$ that depends on $|V|_{\text{Lip}^{\gamma+k}}$ and $\|x\|_{\mathbb{H}^{\text{Höld}}[0,T]}$ such that for every multiindex $\alpha$ with $1 \leq |\alpha| \leq k$ the following estimates hold true:

$$ \sup_{y \in \mathbb{R}^c} |\partial_\alpha \phi(y)|_{\mathbb{H}^{\text{Höld}}[0,T]} \leq C, \quad \sup_{y \in \mathbb{R}^c} |\partial_\alpha \phi^{-1}(y)|_{\mathbb{H}^{\text{Höld}}[0,T]} \leq C. $$

**Proof.** The proof of these results can be found in [17, Chapter 11] and in [7, Lemma 13].

---

1Here, $|\cdot|_{\mathbb{H}^{\text{Höld}}[0,T]}$ denotes the Hölder seminorm of a mapping taking values in $\mathbb{R}^e$. 
4. Contraction property

Let us start with the question of uniqueness. Our first result gives the existence of representatives of a generalized kinetic solution that possess certain left- and right-continuity properties. This builds the foundation for the key estimate of this section, Proposition 4.5, as it allows us to strengthen the sense in which (2.4) is satisfied, namely, we obtain a weak formulation that is only weak in $x, \xi$ (cf. Corollary 4.2) and therefore using time dependent test functions is no longer necessary.

**Proposition 4.1** (Left- and right-continuous representatives). Let $F$ be a generalized kinetic solution to (1.1). Then $F$ it admits representatives $F^-$ and $F^+$ which are left- and right-continuous, respectively, in the sense of $\mathcal{D}'(\mathbb{R}^N \times \mathbb{R})$. More precisely, for all $t^* \in [0, T]$ there exist kinetic functions $F^{\pm, \pm}$ on $\mathbb{R}^N \times \mathbb{R}$ such that setting $F^{\pm}(t^*) = F^{*, \pm}$ yields $F^{\pm} = F$ a.e. and

$$\langle F^{\pm}(t^* \pm \varepsilon), \phi(\theta_t, \pm \varepsilon) \rangle \to \langle F^{\pm}(t^*), \phi(\theta_t) \rangle \quad \varepsilon \downarrow 0 \quad \forall \phi \in C^1_c(\mathbb{R}^N \times \mathbb{R}).$$

**Proof.** Step 1: First, we show that $F$ it admits representatives $F^-$ and $F^+$ that are left- and right-continuous, respectively, in the required sense. The proof follows the ideas of [10, Proposition 8] and [21, Proposition 3.1], nevertheless, as the difficulties here are somewhat different we present the main steps. Let $\phi \in C^1_c(\mathbb{R}^N \times \mathbb{R})$ and $\alpha \in C^1([0, T])$. Integration by parts and the stochastic version of Fubini’s theorem applied to (2.4) yield

$$\int_0^T g_\phi(t)\alpha'(t)dt + \langle F_0, \phi(0) \rangle = \langle m, \partial_\xi \phi(\theta_t) \rangle(\alpha) \quad \mathbb{P}\text{-a.s.}$$

where

$$(4.1) \quad g_\phi(t) = \langle F(t), \phi(\theta_t) \rangle + \int_t^0 \langle \partial_\xi F\phi, \phi(\theta_s) \rangle dW - \frac{1}{2} \int_t^0 \langle \partial_\xi (G^2 \partial_\xi F)\phi, \phi(\theta_s) \rangle ds.$$

Hence $\partial_\xi g_\phi$ is a (pathwise) Radon measure on $[0, T]$ and by the Riesz representation theorem $g_\phi \in BV([0, T])$. Due to the properties of $BV$-functions, we obtain that $g_\phi$ admits left- and right-continuous representatives which coincide except for an at most countable set. Moreover, apart from the first one all terms in (4.1) are almost surely continuous in $t$. Hence, on a set of full measure, denoted by $\Omega_\phi$, $\langle F(t), \phi(\theta_t) \rangle$ also admits left- and right-continuous representatives which coincide except for an at most countable set. Let them be denoted by $\langle F, \phi(\theta_t) \rangle^{\pm}$ and set $\Omega_\phi = \cap_{\theta \in \mathbb{R}} \Omega_\phi, \Omega_\theta$. Note, that as $D_1$ is countable, $\Omega_\phi$ is also a set of full measure. Besides, for $\phi \in D_1$, $(t, \omega) \mapsto \langle F(t, \omega), \phi(\theta_t) \rangle^{+}$ has right continuous trajectories in time and is thus measurable with respect to $(t, \omega)$.

It is now straightforward to define $F^+(t), t \in [0, T], by

$$(4.2) \quad \langle F^+(t), \phi(\theta_t) \rangle := \langle F(t), \phi(\theta_t) \rangle^{+}, \quad \phi \in C^1_c(\mathbb{R}^N \times \mathbb{R})$$

and observe that $F^+(t, \pi_t)$ is right-continuous in the required sense. Note that $F^+(t)$ is a well-defined distribution since, due to the flow properties of $\pi$ and $\theta$, the mapping $\phi \mapsto \phi(\theta_t)$ is a bijection on $C^1_c(\mathbb{R}^N \times \mathbb{R})$ with the inverse mapping $\phi \mapsto \phi(\pi_t)$.

Next, we show that $F^+$ is a representative (in time) of $F$, i.e. for a.e. $t^* \in [0, T]$ it holds that $F^+(t^*) = F(t^*)$ where the equality is understood in the sense of $L^\infty(\mathbb{R}^N \times \mathbb{R})$. Indeed, due to the Lebesgue differentiation theorem,

$$\lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \int_{t^* - \varepsilon}^{t^* + \varepsilon} F(t, \pi_t(x, \xi)) dt = F(t^*, \pi_{t^*}(x, \xi)) \quad \text{a.e. } (t^*, x, \xi)$$

hence by the dominated convergence theorem

$$\lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \int_{t^* - \varepsilon}^{t^* + \varepsilon} \langle F(t, \pi_t), \phi \rangle dt = \langle F(t^*, \pi_{t^*}), \phi \rangle \quad \text{a.e. } t^*, \forall \phi \in C^1_c(\mathbb{R}^N \times \mathbb{R}).$$

Since the left hand side is equal to $\langle F^+(t^*, \pi_{t^*}), \phi \rangle$ for all $t^* \in [0, T]$, the conclusion follows.
Step 2: Second, we prove that for all $t^* \in [0, T]$ 
(4.3) 
$$\langle F^+(t^* \pm \varepsilon), \phi \rangle \to \langle F^+(t^*), \phi \rangle \quad \forall \phi \in C^1_c(\mathbb{R}^N \times \mathbb{R}).$$
Towards this end, we verify 
(4.4) 
$$\langle F^+(t^* + \varepsilon, \pi_{t^*}), \phi \rangle \to \langle F^+(t^*, \pi_{t^*}), \phi \rangle \quad \forall \phi \in C^1_c(\mathbb{R}^N \times \mathbb{R})$$
and observe that testing in (4.4) by $\phi(\pi_{t^*}) \in C^1_c(\mathbb{R}^N \times \mathbb{R})$ yields (4.3).

In order to prove (4.4), we write 
$$\left| \langle F^+(t^* + \varepsilon, \pi_{t^*}), \phi \rangle - \langle F^+(t^*, \pi_{t^*}), \phi \rangle \right|$$
$$\leq \left| \langle F^+(t^* + \varepsilon, \pi_{t^*}), \phi \rangle - \langle F^+(t^* + \varepsilon, \pi_{t^*+\varepsilon}), \phi \rangle \right| + \left| \langle F^+(t^* + \varepsilon, \pi_{t^*+\varepsilon}), \phi \rangle - \langle F^+(t^*, \pi_{t^*}), \phi \rangle \right|.$$ 

The second term on the right hand side vanishes according to Step 1 and regarding the first one, we have 
$$\left| \langle F^+(t^* + \varepsilon, \pi_{t^*}), \phi \rangle - \langle F^+(t^* + \varepsilon, \pi_{t^*+\varepsilon}), \phi \rangle \right| = \left| \langle F^+(t^*) + \phi(\pi_{t^*}) - \phi(\pi_{t^*+\varepsilon}) \rangle \right|$$
and we argue by the dominated convergence theorem: according to Theorem 3.2, if $\text{supp} \phi \subset B_R$,
where $B_R \subset \mathbb{R}^N \times \mathbb{R}$ is the ball of radius $R$ centred at 0, then $\phi(\pi_{t^*})$ and $\phi(\pi_{t^*+\varepsilon})$ remain compactly supported uniformly in $\varepsilon$ and their support is included in $B_{R+CT^{1/p}}$.

Besides, 
$$\phi(\pi_{t^*}(x, \xi)) - \phi(\pi_{t^*+\varepsilon}(x, \xi)) \to 0 \quad \forall (x, \xi) \in \mathbb{R}^N \times \mathbb{R}$$
and since $F^+$ takes values in $[0, 1]$, the claim follows.

Step 3: Now, it only remains to show that $F^+(t^*)$ is a kinetic function on $\mathbb{R}^N \times \mathbb{R}$ for all $t^* \in [0, T]$.
Towards this end, we observe that for all $t^* \in [0, T]$ 
$$F_n(t^*, x, \xi) := \frac{1}{\varepsilon_n} \int_{t^*}^{t^*+\varepsilon_n} F(t, x, \xi) \, dt$$
is a kinetic function on $\mathbb{R}^N$ and by (2.7) the assumptions of Lemma 2.5 are fulfilled. Accordingly, there exists a kinetic function $F^{*, +}$ and a subsequence $(n_k^*)$ such that, on the one hand, 
$$F_{n_k^*}(t^*) \to F^{*, +} \quad \text{in} \quad \mathcal{D}'(\mathbb{R}^N \times \mathbb{R}).$$

Since, on the other hand, we have due to Step 2 that 
$$F_{n_k^*}(t^*) \to F^+(t^*) \quad \text{in} \quad \mathcal{D}'(\mathbb{R}^N \times \mathbb{R}),$$
we deduce that $F^+(t^*) = F^{*, +}$ for all $t^* \in [0, T]$, which completes the proof.

The proof of existence of the left-continuous representative $F^-$ can be carried out similarly. \(\square\)

Remark that according to Proposition 4.1, a stronger version of (2.4) holds true provided $F$ is replaced by $F^+$ or $F^-$. The precise result is presented in the following corollary.

Corollary 4.2. Let $F$ be a generalized kinetic solution to (1.1) and let $F^-$ and $F^+$ be its left- and right-continuous representatives. Then couples $(F^+, m)$ and $(F^-, m)$, respectively, satisfy for all $t \in (0, T)$ and every $\phi \in C^1_c(\mathbb{R}^N \times \mathbb{R})$
$$\langle F^+(t), \phi(\theta_t) \rangle = \langle F_0, \phi \rangle - \int_0^t (\partial_{\xi} Fg, \phi(\theta_s)) \, dW + \frac{1}{2} \int_0^t (\partial_{\xi} (G^2 \partial_{\xi} F, \phi(\theta_s))) \, ds - m(1_{[0, t]} \partial_{\xi} \phi(\theta_s))$$
and
$$\langle F^-(t), \phi(\theta_t) \rangle = \langle F_0, \phi \rangle - \int_0^t (\partial_{\xi} Fg, \phi(\theta_s)) \, dW + \frac{1}{2} \int_0^t (\partial_{\xi} (G^2 \partial_{\xi} F, \phi(\theta_s))) \, ds - m(1_{[0, t]} \partial_{\xi} \phi(\theta_s)),$$

Furthermore, for every $\phi \in C^1_c(\mathbb{R}^N \times \mathbb{R})$ and $t^* \in (0, T)$ it holds true
(4.5) 
$$\langle F^+(t^*) - F^-(t^*), \phi(\theta_{t^*}) \rangle = -m(1_{(0, t^*)} \partial_{\xi} \phi(\theta_{t^*})).$$
In particular, there exists an at most countable set $N \subset [0,T]$ such that, for every $t^* \in [0,T] \setminus N$, $m$ does not have atom at $t^*$.

**Proof.** Consider (2.4) with a test function of the form $(s,x,\xi) \mapsto \phi(x,\xi)\alpha_\varepsilon(s)$ where $\phi \in C^1_c(\mathbb{R}^N \times \mathbb{R})$ and

$$
\alpha_\varepsilon(s) = \begin{cases} 
1, & s \leq t, \\
1 - \frac{s-t}{\varepsilon}, & t \leq s \leq t + \varepsilon, \\
0, & t + \varepsilon \leq s.
\end{cases}
$$

Due to Proposition 4.1 we obtain convergence of the left hand side of (2.4) as $\varepsilon \to 0$:

$$
\int_0^T \langle F(t), \phi(\theta_t) \rangle \, \partial_s \alpha_\varepsilon(s) \, ds = -\frac{1}{\varepsilon} \int_t^{t+\varepsilon} \langle F(t), \phi(\theta_t) \rangle \, ds \rightarrow -\langle F^+(t), \phi(\theta_t) \rangle.
$$

Then we have

$$
m(\alpha_\varepsilon(s) \partial_t \phi(\theta_s)) = \int_{[0,t] \times \mathbb{R}^N \times \mathbb{R}} \partial_t \phi(\theta_s) \, dm + \int_{(t,t+\varepsilon) \times \mathbb{R}^N \times \mathbb{R}} \left(1 - \frac{s-t}{\varepsilon}\right) \partial_t \phi(\theta_s) \, dm
$$

and the convergence of the remaining two terms on the right hand side of (2.4) is obvious because they are continuous in $t$. Therefore, we have justified the equation for $(F^+, m)$.

Concerning $(F^-, m)$, we apply a similar approach but use the function

$$
\alpha_\varepsilon(s) = \begin{cases} 
1, & s \leq t - \varepsilon, \\
\frac{t-s}{\varepsilon}, & t - \varepsilon \leq s \leq t, \\
0, & t \leq s
\end{cases}
$$

instead.

Finally, the formula (4.5) follows from (2.4) by testing by $(t, x, \xi) \mapsto \phi(x, \xi)\alpha_\varepsilon(t)$ where

$$
\alpha_\varepsilon(t) = \frac{1}{\varepsilon} \min \left\{(t - t^* + \varepsilon)^+, (t - t^* - \varepsilon)^-\right\}
$$

and sending $\varepsilon \to 0$. As a consequence, $m$ has an atom at $t^*$ if and only if $F^-(t^*) \neq F^+(t^*)$ and the proof is complete. \(\square\)

**Lemma 4.3.** Let $F$ be a generalized kinetic solution to (1.1) and let $F^+$ be its right-continuous representative. Assume that the initial condition $F_0$ is at equilibrium: there exists $u_0 \in L^1(\Omega \times \mathbb{R}^N)$ such that $F_0 = 1_{u_0 > \xi}$. Then $F^+(0) = F_0$ and in particular the corresponding kinetic measure $m$ has no atom at 0, i.e. the restriction of $m$ to $\{0\} \times \mathbb{R}^N \times \mathbb{R}$ vanishes.

**Proof.** Let $m_0$ be the restriction of $m$ to $\{0\} \times \mathbb{R}^N \times \mathbb{R}$. Then we deduce from (4.5) at $t^* = 0$ that

$$
\langle F^+(0) - 1_{u_0 > \xi}, \phi \rangle = -m_0(\partial_t \phi) \quad \forall \phi \in C^1_c(\mathbb{R}^N \times \mathbb{R}).
$$

Moreover, it follows from the integrability (in $\xi$) of

$$
F^+(0) - 1_{u_0 > \xi} = (F^+(0) - 1_{0 > \xi}) + (1_{u_0 > \xi} - 1_{0 > \xi})
$$

that one can test by $\phi(x, \xi) = \phi(x)$ to obtain

$$
\langle F^+(0) - 1_{0 > \xi}, \phi \rangle = \langle 1_{u_0 > \xi} - 1_{0 > \xi}, \phi \rangle \quad \forall \phi \in C^1_c(\mathbb{R}^N)
$$

which leads to

$$
\int_\mathbb{R} F^+(0, \xi) - 1_{0 > \xi} \, d\xi = u_0.
$$
Indeed, setting $\int_{-\infty}^{\xi} 1_{u_0 > \xi} - F^+(0, \xi) \, d\xi = \int_{-\infty}^{\xi} (1_{u_0 > \xi} - 1_{0 > \xi}) - (F^+(0, \xi) - 1_{0 > \xi}) \, d\xi \geq 0$.

The following result allows us to restart the evolution given by (1.1) at an arbitrary time $t^* \in (0, T)$.

**Lemma 4.4.** Let $F$ be a generalized kinetic solution to (1.1) on $[0, T]$ with the initial datum $F_0$. Then for every $t^* \in [0, T)$, $t \mapsto F(t^* + t)$ is a generalized kinetic solution to (1.1) on $[0, T - t^*]$ with the initial datum $F^-(t^*)$.

**Proof.** Let $\alpha \in C^1_c((0, T))$, $\phi \in C_c^1(\mathbb{R}^N \times \mathbb{R})$ and test (2.4) by $(t, x, \xi) \mapsto \phi(x, \xi)\alpha_\varepsilon(t)$ where

$$\alpha_\varepsilon(t) = \begin{cases} 
\alpha(t), & t \leq t^*, \\
\alpha(t)
\left(1 - \frac{t - t^*}{\varepsilon}\right), & t^* \leq t \leq t^* + \varepsilon, \\
0, & t^* + \varepsilon \leq t.
\end{cases}$$

In the limit $\varepsilon \to 0$ we infer

$$\int_0^{t^*} \langle F(t), \phi(\theta_t) \rangle \partial_t \alpha(t) \, dt - \langle F^+(t^*), \phi(\theta_{t^*}) \rangle \alpha(t^*) + \langle F_0, \phi \rangle \alpha(0)$$

$$\quad = \int_0^{t^*} \langle \partial_\xi Fg, \phi(\theta_t) \rangle \alpha(t) \, dW - \frac{1}{2} \int_0^{t^*} \langle \partial_\xi (G^2 \partial_\xi F), \phi(\theta_t) \rangle \alpha(t) \, dt + m \left(1_{[0, t^*]} \alpha \partial_\xi \phi(\theta_t)\right).$$

Subtracting (4.7) from (2.4) and using (4.5) leads to

$$\int_{t^*}^T \langle F(t), \phi(\theta_t) \rangle \partial_t \alpha(t) \, dt + \langle F^-(t^*), \phi(\theta_{t^*}) \rangle \alpha(t^*)$$

$$\quad = \int_{t^*}^T \langle \partial_\xi Fg, \phi(\theta_t) \rangle \alpha(t) \, dW - \frac{1}{2} \int_{t^*}^T \langle \partial_\xi (G^2 \partial_\xi F), \phi(\theta_t) \rangle \alpha(t) \, dt + m \left(1_{[t^*, T]} \alpha \partial_\xi \phi(\theta_{t^*})\right).$$

Finally we take $(t, x, \xi) \mapsto \phi(\pi_{t^*}(x, \xi))\alpha(t)$ as a test function and deduce that

$$\int_{t^*}^T \langle F(t), \phi(\theta_{t^*, t}) \rangle \partial_t \alpha(t) \, dt + \langle F^-(t^*), \phi \rangle \alpha(t^*)$$

$$\quad = \int_{t^*}^T \langle \partial_\xi Fg, \phi(\theta_{t^*, t}) \rangle \alpha(t) \, dW - \frac{1}{2} \int_{t^*}^T \langle \partial_\xi (G^2 \partial_\xi F), \phi(\theta_{t^*, t}) \rangle \alpha(t) \, dt + m \left(1_{[t^*, T]} \alpha \partial_\xi \phi(\theta_{t^*, t})\right).$$

As a consequence, $t \mapsto F(t^* + t)$ is a kinetic solution to (1.1) on $[0, T - t^*]$ with initial condition $F^-(t^*)$. □

We proceed with a result that plays the key role in our proof of reduction of generalized kinetic solution to a kinetic one, Theorem 4.6, and the $L^1$-contraction property, Corollary 4.7. To be more precise, the relevant estimate that we need to verify in case of two generalized kinetic solutions $F_1, F_2$ is the following:

$$\int_{\mathbb{R}^N \times \mathbb{R}} F_1(t)(1 - F_2(t)) \, dx \, d\xi \leq \int_{\mathbb{R}^N \times \mathbb{R}} F_{1,0}(1 - F_{2,0}) \, dx \, d\xi.$$ 

Indeed, setting $F := F_1 = F_2$ then leads to the Reduction Theorem and using the identity

$$\int_{\mathbb{R}} 1_{u_1(t) > \xi}(1 - 1_{u_2(t) > \xi}) \, d\xi = (u_1(t) - u_2(t))^+$$
gives the $L^1$-contraction property. However, it is not possible to multiply the two equations directly, i.e. the equation for $F_1$ and the one for $1 - F_2$ as it is necessary to mollify first. This is taken care of in Proposition 4.5. In order to simplify the notation, we denote $F = 1 - F$.

In the sequel, let $(\bar{\varrho}_t)$ be an approximation to the identity on $\mathbb{R}^N \times \mathbb{R}$ and $(\kappa_R)$ be a truncation on $\mathbb{R}^N$ such that $\kappa_R \equiv 1$ on $B_R$, supp$\kappa_R \subset B_{2R}$ and $|\kappa_R|, |\nabla_x \kappa_R| \leq 1$.

**Proposition 4.5.** Let $F_1, F_2$ be generalized kinetic solutions to (1.1). Let $s, t \in [0, T]$, $s \leq t$, be such that neither of the kinetic measures $m_1$ and $m_2$ has atom at $s$. Then it holds true

$$
\mathbb{E} \int \kappa_R(x) \langle F_1^-(t), \varrho_s((x, \xi) - \theta_{s,t}(\cdot, \cdot)) \rangle_{z,\sigma} \langle \bar{F}_2^-(t), \varrho_s((x, \xi) - \theta_{s,t}(\cdot, \cdot)) \rangle_{y,\zeta} dx \, d\xi \\
- \mathbb{E} \int \kappa_R(x) \langle F_1^-(s), \varrho_s((x, \xi) - \theta_{s,t}(\cdot, \cdot)) \rangle_{z,\sigma} \langle \bar{F}_2^-(s), \varrho_s((x, \xi) - \theta_{s,t}(\cdot, \cdot)) \rangle_{y,\zeta} dx \, d\xi
$$

(4.8)

where

$$
K_1 = C(t-s)^{1/p} \mathbb{E} \int_s^t \int_{1 \leq |z| \leq 2R} \varrho_s((x, \xi) - \theta_{s,r}(z, \sigma)) dm_1(r, z, \sigma) dx \, dz, \\
K_2 = C(t-s)^{1/p} \mathbb{E} \int_s^t \int_{1 \leq |z| \leq 2R} \varrho_s((x, \xi) - \theta_{s,r}(y, \zeta)) dm_2(r, y, \zeta) dx \, dz, \\
K_3 = C(t-s)^{1/p} \mathbb{E} \int_s^t \int_{1 \leq |z| \leq 2R} |\sigma|^2 \varrho_s((x, \xi) - \theta_{s,r}(z, \sigma)) dr \, dz \, d\xi \, dr, \\
K_4 = C(t-s)^{1/p} \mathbb{E} \int_s^t \int_{1 \leq |z| \leq 2R} |\sigma|^2 \varrho_s((x, \xi) - \theta_{s,r}(y, \zeta)) dr \, dz \, d\xi \, dr.
$$

**Proof.** Step 1: Applying Corollary 4.2 and Lemma 4.4 to the generalized kinetic solutions $F_1, F_2$ and to the mollifier $\varrho_s$ in place of a test function, we deduce that for all $(x, \xi) \in \mathbb{R}^N \times \mathbb{R}$

$$
\langle F_1^-(t), \varrho_s((x, \xi) - \theta_{s,t}(\cdot, \cdot)) \rangle_{z,\sigma}
$$

(4.9)

$$
= \langle F_1^-(s), \varrho_s((x, \xi) - \theta_{s,r}(\cdot, \cdot)) \rangle_{z,\sigma} - \int_s^t \langle \partial_\sigma F_1, \varrho_s((x, \xi) - \theta_{s,r}(\cdot, \cdot)) \rangle_{z,\sigma} \, dW
$$

and similarly

$$
\langle \bar{F}_2^-(t), \varrho_s((x, \xi) - \theta_{s,t}(\cdot, \cdot)) \rangle_{y,\zeta}
$$

(4.10)

$$
= \langle \bar{F}_2^-(s), \varrho_s((x, \xi) - \theta_{s,t}(\cdot, \cdot)) \rangle_{y,\zeta} + \int_s^t \langle \partial_\zeta \bar{F}_2, \varrho_s((x, \xi) - \theta_{s,r}(\cdot, \cdot)) \rangle_{y,\zeta} \, dW
$$

+ \int_s^t \langle \partial_\sigma \bar{F}_2, \varrho_s((x, \xi) - \theta_{s,r}(\cdot, \cdot)) \rangle_{y,\zeta} \, d\zeta.
$$

Our notation in the above is the following: the variable $(x, \xi)$ is fixed in both (4.9) and (4.10); in (4.9) we consider the test function $(z, \sigma) \mapsto \varrho_s((x, \xi) - \theta_{s,t}(z, \sigma))$ whereas in (4.10) we make use of $(y, \zeta) \mapsto \varrho_s((x, \xi) - \theta_{s,t}(y, \zeta))$. 

We denote
\[ I_1(t) := -\int_s^t \langle \partial_\sigma F_1, \theta_\sigma((x, \xi) - \theta_{s,r}(\cdot, \cdot)) \rangle_{z,\sigma} dW, \]
\[ I_2(t) := \int_s^t \langle \partial_\xi F_2, \theta_\xi((x, \xi) - \theta_{s,r}(\cdot, \cdot)) \rangle_{y,\xi} dW, \]
and observe that
\[ \mu_1 : C_b([s, T]) \to \mathbb{R}, \]
\[ \alpha \mapsto -m_1 \left( \alpha(\cdot) \partial_\sigma \theta_\sigma((x, \xi) - \theta_{s,r}(\cdot, \cdot)) \right) + \frac{1}{2} \int_s^T \alpha(r) \langle G^2 \partial_\sigma F_1, \partial_\sigma \theta_\sigma((x, \xi) - \theta_{s,r}(\cdot, \cdot)) \rangle_{z,\sigma} dr, \]
and
\[ \mu_2 : C_b([s, T]) \to \mathbb{R}, \]
\[ \alpha \mapsto m_2 \left( \alpha(\cdot) \partial_\xi \theta_\xi((x, \xi) - \theta_{s,r}(\cdot, \cdot)) \right) + \frac{1}{2} \int_s^T \alpha(r) \langle G^2 \partial_\xi F_2, \partial_\xi \theta_\xi((x, \xi) - \theta_{s,r}(\cdot, \cdot)) \rangle_{y,\xi} dr, \]
are a Radon measures hence
\[ t \mapsto \mu_1([s, t]), \quad t \mapsto \mu_2([s, t]) \]
are left-continuous functions of bounded variation.

Multiplying (4.9) and (4.10), we have
\[ E(\langle F_1^-(t), \theta_\delta((x, \xi) - \theta_{s,t}(\cdot, \cdot)) \rangle_{z,\sigma}) \langle F_2^-(t), \theta_\delta((x, \xi) - \theta_{s,t}(\cdot, \cdot)) \rangle_{y,\xi} \]
\[ - E(\langle F_1^-(s), \theta_\delta((x, \xi) - \theta_{s,t}(\cdot, \cdot)) \rangle_{z,\sigma}) \langle F_2^-(s), \theta_\delta((x, \xi) - \theta_{s,t}(\cdot, \cdot)) \rangle_{y,\xi} \]
\[ = E(\langle F_1^-(s), \theta_\delta((x, \xi) - \theta_{s,t}(\cdot, \cdot)) \rangle_{z,\sigma} I_2(t) + \langle F_1^-(s), \theta_\delta((x, \xi) - \theta_{s,t}(\cdot, \cdot)) \rangle_{z,\sigma} \mu_2([s, t]) \]
\[ + E I_1(t) \langle F_2^-(s), \theta_\delta((x, \xi) - \theta_{s,t}(\cdot, \cdot)) \rangle_{y,\xi} + E I_1(t) \mu_2([s, t]) \]
\[ + E \mu_1([s, t]) \langle F_2^-(s), \theta_\delta((x, \xi) - \theta_{s,t}(\cdot, \cdot)) \rangle_{y,\xi} + E \mu_1([s, t]) \mu_2([s, t]) \]
\[ = J_1 + \cdots + J_8. \]

For \( J_5 \) we use the following integration by parts formula
\[ I_1(t) \mu_2([s, t]) = \int_s^t \mu_2([s, r]) dI_1(r) + \int_{(s,t)} I_1(r) d\mu_2(r) \]
which implies
\[ J_5 = E \int_{(s,t)} I_1(r) d\mu_2(r) \]
and similarly we obtain
\[ J_7 = E \int_{(s,t)} I_2(r) d\mu_1(r). \]
The same reasoning yields
\[ J_1 = J_3 = 0. \]

\( J_4 \) can be rewritten as follows
\[ J_4 = E(\langle I_1(t), I_2(t) \rangle) \]
\[ = -E \int_s^t g(z, \sigma) \cdot g(y, \zeta) \theta_\delta((x, \xi) - \theta_{s,r}(z, \sigma)) \theta_\delta((x, \xi) - \theta_{s,r}(y, \zeta)) dr_{r,z}(\sigma) dr_{r,y}(\zeta) dz dy dr. \]
Next, we apply the integration by parts formula for functions of bounded variation (see e.g. [36, Chapter 0]) to $J_8$ and deduce

$$J_8 = \mathbb{E}\mu_1(\{s\})\mu_2(\{s\}) + \mathbb{E}\int_{(s,t)} \mu_1([s, r)) \, d\mu_2(r) + \mathbb{E}\int_{(s,t)} \mu_2([s, r)) \, d\mu_1(r)$$

$$= J_{80} + J_{81} + J_{82}.$$ 

Since the kinetic measures $m_1$ and $m_2$ do not have atoms at $s$ due to the assumptions, it follows that $J_{80} = 0$. Regarding $J_{81}$, we make use of the formula for $\mu_1([s, r))$ given by (4.9). Namely

$$J_{81} = \mathbb{E}\int_{(s,t)} \langle F_{1}^{-}(r), \varrho_\delta((x, \xi) - \theta_{s,r}(\cdot, \cdot)) \rangle_{z,\sigma} \, d\mu_2(r)$$

$$- \mathbb{E}\langle F_{1}^{-}(s), \varrho_\delta((x, \xi) - \theta_{s,r}(\cdot, \cdot)) \rangle_{z,\sigma} \mu_2([s, t)) - \mathbb{E}\int_{(s,t)} \mathcal{I}_1(r) \, d\mu_2(r) = J_{811} - J_2 - J_5.$$ 

Going back to the product of (4.9) and (4.10) we see that $J_2$ and $J_5$ cancel and for $J_{811}$ we write

$$J_{811} = \mathbb{E}\int_{(s,t)} \int F_{1}^{-}(r) \varrho_\delta((x, \xi) - \theta_{s,r}(z, \sigma)) \partial_z \varrho_\delta((x, \xi) - \theta_{s,r}(y, \zeta)) \, d\mu_2(r, y, \zeta) \, dz \, ds$$

$$- \frac{1}{2} \mathbb{E}\int_{s}^{t} \int F_{1}^{-}(r)G^2(y, \zeta) \varrho_\delta((x, \xi) - \theta_{s,r}(z, \sigma)) \partial_z \varrho_\delta((x, \xi) - \theta_{s,r}(y, \zeta)) \, d\nu^2_{y, \sigma}(\zeta) \, dy \, d\sigma \, dr.$$ 

Let us now continue with $J_{82}$. We apply the formula for $\mu_2([s, r))$ given by (4.10). It yields

$$J_{82} = \mathbb{E}\int_{(s,t)} \langle F_{2}^{-}(r), \varrho_\delta((x, \xi) - \theta_{s,r}(\cdot, \cdot)) \rangle_{y,\zeta} \, d\mu_1(r)$$

$$- \mathbb{E}\langle F_{2}^{-}(s), \varrho_\delta((x, \xi) - (\cdot, \cdot)) \rangle_{y,\zeta} \mu_1([s, t)) - \mathbb{E}\int_{(s,t)} \mathcal{I}_2(r) \, d\mu_1(r) = J_{821} - J_6 - J_7,$$

where again $J_6$ and $J_7$ cancel and for $J_{821}$ we have

$$J_{821} = -\mathbb{E}\int_{(s,t)} \int F_{2}^{-}(r) \varrho_\delta((x, \xi) - \theta_{s,r}(y, \zeta)) \partial_\sigma \varrho_\delta((x, \xi) - \theta_{s,r}(z, \sigma)) \, d\mu_1(r, z, \sigma) \, dy \, d\zeta$$

$$+ \frac{1}{2} \mathbb{E}\int_{s}^{t} \int F_{2}^{-}(r)G^2(y, \zeta) \varrho_\delta((x, \xi) - \theta_{s,r}(z, \sigma)) \partial_\sigma \varrho_\delta((x, \xi) - \theta_{s,r}(y, \zeta)) \, d\nu^2_{y, \sigma}(\zeta) \, dy \, d\sigma \, dr.$$ 

**Step 2:** Finally, we have all in hand to proceed with the proof. Integrating (4.11) with respect to $x, \xi$ we obtain

$$\mathbb{E}\int \kappa_R(x)F_{1}^{-}(t)F_{2}^{-}(t) \varrho_\delta((x, \xi) - \theta_{s,t}(\cdot, \cdot)) \, dx \, dy \, d\zeta \, dz \, d\sigma$$

$$- \mathbb{E}\int \kappa_R(x)F_{1}^{-}(s)F_{2}^{-}(s) \varrho_\delta((x, \xi) - (\cdot, \cdot)) \, dx \, dy \, d\zeta \, dz \, d\sigma = I_1 + \cdots + I_5,$$

where

$$I_1 = \mathbb{E}\int_{(s,t)} \kappa_R(x)F_{1}^{-}(r) \varrho_\delta((x, \xi) - \theta_{s,r}(\cdot, \cdot)) \, dx \, dy \, d\zeta \, dz \, d\sigma,$$

$$I_2 = -\mathbb{E}\int_{(s,t)} \kappa_R(x)F_{2}^{-}(r) \varrho_\delta((x, \xi) - \theta_{s,r}(\cdot, \cdot)) \, dx \, dy \, d\zeta \, dz \, d\sigma,$$

$$I_3 = -\mathbb{E}\int_{s}^{t} \kappa_R(x)g(z, \sigma) \cdot g(y, \zeta) \varrho_\delta((x, \xi) - \theta_{s,r}(\cdot, \cdot)) \, dx \, dy \, d\zeta \, dz \, d\sigma,$$

$$I_4 = -\frac{1}{2} \mathbb{E}\int_{s}^{t} \kappa_R(x)F_{1}^{-}(r)G^2(y, \zeta) \varrho_\delta((x, \xi) - \theta_{s,r}(\cdot, \cdot)) \, dx \, dy \, d\zeta \, dz \, d\sigma,$$

$$I_5 = -\frac{1}{2} \mathbb{E}\int_{s}^{t} \kappa_R(x)F_{2}^{-}(r)G^2(y, \zeta) \varrho_\delta((x, \xi) - \theta_{s,r}(\cdot, \cdot)) \, dx \, dy \, d\zeta \, dz \, d\sigma.$$
\[ I_5 = \frac{1}{2} E \int_s^t \int \kappa_R(x) \bar{F}_2(r) G^2(z, \sigma) \varrho_\delta((x, \xi) - \theta_{s,r}(y, \zeta)) \partial_\sigma \varrho_\delta((x, \xi) - \theta_{s,r}(z, \sigma)) d\nu_{1,z}(\sigma). \]

Above as well as in the sequel, when we integrate with respect to all variables \( r, x, \xi, z, \sigma, y, \zeta \), we only specify the kinetic measures \( m_1, m_2 \) and the Young measures \( \nu^1, \nu^2 \), the Lebesgue measure being omitted.

Since the technique of estimation of \( I_1 \) and \( I_2 \), let us just focus on \( I_1 \). It holds that
\[
I_1 = E \int_{(s,t)} \int \kappa_R(x) \partial_r \bar{F}_1(r) \varrho_\delta((x, \xi) - \theta_{s,r}(z, \sigma)) \varrho_\delta((x, \xi) - \theta_{s,r}(y, \zeta)) dm_2(r, y, \zeta)
+ E \int_{(s,t)} \int \kappa_R(x) \partial_\sigma \varrho_\delta((x, \xi) - \theta_{s,r}(z, \sigma)) \varrho_\delta((x, \xi) - \theta_{s,r}(y, \zeta)) dm_2(r, y, \zeta)
+ E \int_{(s,t)} \int \kappa_R(x) \bar{F}_1(r) \varrho_\delta((x, \xi) - \theta_{s,r}(z, \sigma)) \partial_\xi \varrho_\delta((x, \xi) - \theta_{s,r}(y, \zeta)) dm_2(r, y, \zeta)
= I_{11} + I_{12} + I_{13},
\]
where the first equality holds true because the sum of the first two terms is equal to zero using integration by parts. Since \( \bar{F}_1(r) \) is a kinetic function for every \( r \in [s, t] \), it can be seen that \( I_{11} \leq 0 \).

We will show that \( I_{12} + I_{13} \) is small if \( t - s \) is small. Towards this end, we observe that
\[
\partial_\sigma \varrho_\delta((x, \xi) - \theta_{s,r}(z, \sigma)) = -\nabla_{x, \xi} \varrho_\delta((x, \xi) - \theta_{s,r}(z, \sigma)) \cdot \partial_\sigma \theta_{s,r}(z, \sigma)
\]
and
\[
\int \kappa_R(x) \nabla_{x, \xi} \varrho_\delta((x, \xi) - \theta_{s,r}(z, \sigma)) \varrho_\delta((x, \xi) - \theta_{s,r}(y, \zeta)) dx d\xi
+ \int \kappa_R(x) \varrho_\delta((x, \xi) - \theta_{s,r}(z, \sigma)) \nabla_{x, \xi} \varrho_\delta((x, \xi) - \theta_{s,r}(y, \zeta)) dx d\xi
= -\int \nabla_{x, \xi} \kappa_R(x) \varrho_\delta((x, \xi) - \theta_{s,r}(z, \sigma)) \varrho_\delta((x, \xi) - \theta_{s,r}(y, \zeta)) dx d\xi
\]
it follows that
\[
I_{12} + I_{13} = E \int_{(s,t)} \int \bar{F}_1(r) \kappa_R(x) [\partial_\sigma \theta_{s,r}(z, \sigma) - \partial_\xi \theta_{s,r}(y, \zeta)]
\cdot \nabla_{x, \xi} \varrho_\delta((x, \xi) - \theta_{s,r}(z, \sigma)) \varrho_\delta((x, \xi) - \theta_{s,r}(y, \zeta)) dm_2(r, y, \zeta)
+ E \int_{(s,t)} \int \bar{F}_1(r) \nabla_{x, \xi} \kappa_R(x) \cdot \partial_\sigma \theta_{s,r}(z, \sigma)
\times \varrho_\delta((x, \xi) - \theta_{s,r}(z, \sigma)) \varrho_\delta((x, \xi) - \theta_{s,r}(y, \zeta)) dm_2(r, y, \zeta)
= I_{121} + I_{131}
\]
To proceed, we use the ideas from [19]. Namely, we observe that for every \( r \in [s, t] \), \( \partial_\sigma \theta_{s,r}(\cdot) \) is Lipschitz continuous and its Lipschitz constant can be estimated by \( C(t-s)^{1/p} \). Indeed, according to Theorem 3.2

\[
\sup_{x, \xi \in \mathbb{R}^N \times \mathbb{R}}|\partial_\sigma \theta_{s,r}(x, \xi)| = \sup_{x, \xi \in \mathbb{R}^N \times \mathbb{R}}|\partial_\sigma \theta_{s,r}(x, \xi) - \partial_\sigma \theta_{s,s}(x, \xi)| \leq C(r-s)^{1/p} \leq C(t-s)^{1/p}.
\]

Besides, we only have to consider the case of
\[
|\theta_{s,s}(z, \sigma) - \theta_{s,s}(y, \zeta)| < \delta
\]
hence
\[
|\theta_{s,s}(z, \sigma) - \theta_{s,s}(y, \zeta)| < 2\delta
\]
and consequently
\[
|\theta_{s,s}(z, \sigma) - \theta_{s,s}(y, \zeta)| < C\delta
\]
which implies
\begin{equation}
(4.16) \quad |\partial_s \theta_{s,r}(z, \sigma) - \partial_s \theta_{s,r}(y, \zeta)| \leq C(t - s)^{1/p} \delta.
\end{equation}

Let us now consider the remaining terms in the integral with respect to \((x, \xi, z, \sigma)\) and use the fact that \(\delta |\nabla x, \xi \theta_{s}\| (\cdot) \leq C \delta_{2}\theta_{s}(\cdot)\). It yields
\begin{align}
&\int \kappa_R(x)|\nabla x, \xi \theta_{s}\|((x, \xi) - \theta_{s,r}(z, \sigma))| \theta_{s}(x, \xi) - \theta_{s,r}(y, \zeta)\|dx\|d\xi\|dz\|d\sigma \\
&\leq \int |\nabla x, \xi \theta_{s}\|((x, \xi) - (z, \sigma))| \theta_{s}(x, \xi) - \theta_{s,r}(y, \zeta)\|dx\|d\xi\|dz\|d\sigma \\
&\leq \frac{C}{\delta} \int \theta_{s}(x, \xi) - \theta_{s,r}(y, \zeta)\|dx\|d\xi\|dz\|d\sigma = \frac{C}{\delta}.
\end{align}

Plugging (4.16) and (4.17) into \(I_{121}\) we conclude that
\begin{equation}
I_{121} \leq C(t - s)^{1/p} \mathbb{E} \int_s^t \int dm_2(s, y, \zeta).
\end{equation}

Since due to Theorem 3.2
\begin{equation}
\sup_{(x, \xi) \in \mathbb{R}^N \times \mathbb{R}} |\partial_s \theta_{s,r}(x, \xi)| = \sup_{(x, \xi) \in \mathbb{R}^N \times \mathbb{R}} |\partial_s \theta_{s,r}(x, \xi) - \partial_s \theta_{s,r}(x, \xi)| \leq C(t - s)^{1/p},
\end{equation}

we deduce
\begin{equation}
I_{131} \leq C(t - s)^{1/p} \mathbb{E} \int_s^t \int 1_{R \leq |y| \leq 2R} \theta_{s}(x, \xi) - \theta_{s,r}(y, \zeta)\|dx\|d\xi\|dz\|d\sigma.
\end{equation}

Let us proceed with \(I_4\) and \(I_5\). Using the ideas from (4.12) and (4.13) we deduce
\begin{align}
I_4 &= -\frac{1}{2} \mathbb{E} \int_s^t \int F_1(r)G^2(y, \zeta)\kappa_R(x) \left[\partial_s \theta_{s,r}(y, \zeta) - \partial_s \theta_{s,r}(z, \sigma)\right] \\
&\quad \cdot \nabla x, \xi \theta_{s}\|((x, \xi) - \theta_{s,r}(z, \sigma))| \theta_{s}(x, \xi) - \theta_{s,r}(y, \zeta)\| \|dx\| \|d\sigma\|
\end{align}

and similarly
\begin{align}
I_5 &= -\frac{1}{2} \mathbb{E} \int_s^t \int F_2(r)G^2(z, \sigma)\kappa_R(x) \left[\partial_s \theta_{s,r}(y, \zeta) - \partial_s \theta_{s,r}(z, \sigma)\right] \\
&\quad \cdot \nabla x, \xi \theta_{s}\|((x, \xi) - \theta_{s,r}(z, \sigma))| \theta_{s}(x, \xi) - \theta_{s,r}(y, \zeta)\| \|dx\| \|d\sigma\|
\end{align}
Using similar approach as for $I_{121}$ together with the fact that $G^2(y, \zeta) \leq C|\zeta|^2$ which follows from the assumptions on $g$, we deduce that

$$I_{41} + I_{51} \leq C(t-s)^{1/p} \left( E \int_s^t \int G^2(y, \zeta) d\nu_{r,y}^2(\zeta) dydr + E \int_s^t \int G^2(z, \sigma) d\nu_{r,z}^1(\sigma) dzdr \right) \leq C(t-s)^{1/p} \left( E \int_s^t \int |\zeta|^2 d\nu_{r,y}^2(\zeta) dydr + E \int_s^t \int |\sigma|^2 d\nu_{r,z}^1(\sigma) dzdr \right)$$

and similarly to $I_{131}$ we get

$$I_{42} + I_{52} \leq C(t-s)^{1/p} E \int_s^t \int 1_{R \leq |x| \leq 2R} |\zeta|^2 \rho_\delta((x, \xi) - \theta_{s,r}(y, \zeta)) d\nu_{r,y}^2(\zeta) dydzd\xi dr + C(t-s)^{1/p} E \int_s^t \int 1_{R \leq |x| \leq 2R} |\sigma|^2 \rho_\delta((x, \xi) - \theta_{s,r}(z, \sigma)) d\nu_{r,z}^1(\sigma) dzdzd\xi dr.$$

Besides,

$$I_3 + I_{43} + I_{53} = \frac{1}{2} E \int_s^t \int \kappa_R(x) |g(z, \sigma) - g(y, \zeta)|^2 \rho_\delta((x, \xi) - \theta_{s,r}(z, \sigma))\rho_\delta((x, \xi) - \theta_{s,r}(y, \zeta)) d\nu_{r,z}^1(\sigma) d\nu_{r,y}^2(\zeta) \leq C \delta^2 E \int_s^t \int_{Q_R} \rho_\delta((x, \xi) - \theta_{s,r}(z, \sigma))\rho_\delta((x, \xi) - \theta_{s,r}(y, \zeta)) d\nu_{r,z}^1(\sigma) d\nu_{r,y}^2(\zeta)$$

where $Q_R = B_{2R} \times \mathbb{R}^k \times B_{2R+CT^{1/p},z} \times \mathbb{R}^\sigma \times B_{2R+CT^{1/p},y} \times \mathbb{R}^\zeta$ and we used (4.15) and the Lipschitz continuity of $g$. Next, we write

$$I_3 + I_{43} + I_{53} = C \delta^2 E \int_s^t \int_{Q_R} \rho_\delta((x, \xi) - \theta_{s,r}(z, \sigma)) - \rho_\delta((x, \xi) - \theta_{s,r}(y, \zeta)) d\nu_{r,z}^1(\sigma) d\nu_{r,y}^2(\zeta) + C \delta^2 E \int_s^t \int_{Q_R} \rho_\delta((x, \xi) - \theta_{s,r}(y, \zeta)) - \rho_\delta((x, \xi) - \theta_{s,r}(y, \zeta)) d\nu_{r,z}^1(\sigma) d\nu_{r,y}^2(\zeta) + C \delta^2 E \int_s^t \int_{Q_R} \rho_\delta((x, \xi) - \theta_{s,r}(y, \zeta)) - \rho_\delta((x, \xi) - \theta_{s,r}(y, \zeta)) d\nu_{r,z}^1(\sigma) d\nu_{r,y}^2(\zeta) = I_{61} + I_{62} + I_{63}$$

and estimate each of the terms separately. We have

$$I_{61} \leq C \delta E \int_s^t \int_{Q_R} |\theta_{s,r}(z, \sigma) - (z, \sigma)| \rho_\delta((x, \xi) - \theta_{s,r}(y, \zeta)) d\nu_{r,z}^1(\sigma) d\nu_{r,y}^2(\zeta) \leq C \delta(t-s)^{1/p} E \int_s^t \int_{Q_R} \rho_\delta((x, \xi) - \theta_{s,r}(y, \zeta)) d\nu_{r,z}^1(\sigma) d\nu_{r,y}^2(\zeta) \leq CR^2 \delta(t-s)\leq(t-s)^{1+1/p}.$$ 

The same approach yields

$$I_{62} \leq CR^2 \delta(t-s)^{1+1/p}$$

and similarly

$$I_{63} \leq C \delta E \int_s^t \int_{Q_R} \rho_\delta((x, \xi) - \theta_{s,r}(y, \zeta)) d\nu_{r,z}^1(\sigma) d\nu_{r,y}^2(\zeta) \leq CR^2 \delta(t-s),$$

which completes the proof of (4.8).

Finally, we have all in hand to prove the Reduction Theorem and the $L^1$-contraction property. \qed
Theorem 4.6 (Reduction Theorem). Let $F$ be a generalized kinetic solution to (1.1) with initial datum $F_0 = 1_{u_0 > \xi}$. Then it is a kinetic solution to (1.1), that is, there exists
\[
F(t) \in L^1(\Omega; L^\infty(0,T; L^1(\mathbb{R}^N))) \cap L^2(\Omega, L^2(0,T; L^2(\mathbb{R}^N)))
\]
such that $F(t,x,\xi) = 1_{u(t,x) > \xi}$ a.e.

Proof. Let $t \in (0,T)$, $h > 0$ and $\{0 = t_0 < t_1 < \cdots < t_n = t\}$ be a partition of $[0,t]$ with step size $h$. We shall assume (without loss of generality) that the kinetic measure $m$ has no atom at $t_i$, $i = 1, \ldots, n-1$. Moreover, due to Lemma 4.3, $m$ has no atom at 0. Therefore application of Proposition 4.5 yields
\[
\mathbb{E} \int \kappa_R(x)(F^-(t_i), \delta ((x, \xi - \theta_{t_i-1}, t_i, (\cdot, \cdot)))_{z,\sigma} \langle F^- (t_i), \delta ((x, \xi - \theta_{t_i-1}, t_i, (\cdot, \cdot)))_{y,\sigma} \rangle_{y,\xi} \, dx \, d\xi \\
- \mathbb{E} \int \kappa_R(x)(F^-(t_i-1), \delta ((x, \xi - \theta_{t_i-1}, t_i-1, (\cdot, \cdot)))_{z,\sigma} \langle F^- (t_i-1), \delta ((x, \xi - \theta_{t_i-1}, t_i-1, (\cdot, \cdot)))_{y,\sigma} \rangle_{y,\xi} \, dx \, d\xi \\
\leq Ch^{1/p} \mathbb{E} \int_{t_{i-1}}^{t_i} \int \text{dm}(r, y, \zeta) + Ch^{1/p} \mathbb{E} \int_{t_{i-1}}^{t_i} \int \langle \sigma \rangle \, d\nu_{r,y}(\zeta) \, dy \, d\zeta.
\]

Consequently,
\[
\mathbb{E} \int F^-(t, x, \xi) \bar{F}^-(t, x, \xi) \, dx \, d\xi - \mathbb{E} \int F^-(t_{i-1}, x, \xi) \bar{F}^-(t_{i-1}, x, \xi) \, dx \, d\xi \\
= \sum_{i=1}^{n} \mathbb{E} \int F^-(t_i, x, \xi) \bar{F}^-(t_{i-1}, x, \xi) \, dx \, d\xi - \mathbb{E} \int F^-(t_{i-1}, x, \xi) \bar{F}^-(t_{i-1}, x, \xi) \, dx \, d\xi \\
\leq Ch^{1/p} \mathbb{E} \int_0^t \int \text{dm}(r, y, \zeta) + Ch^{1/p} \mathbb{E} \int_0^t \int \langle \sigma \rangle \, d\nu_{r,y}(\zeta) \, dy \, d\zeta
\]
and sending $h \to 0$ yields
\[
\mathbb{E} \int F^-(t, x, \xi) \bar{F}^-(t, x, \xi) \, dx \, d\xi \leq \mathbb{E} \int F_0 \bar{F}_0 \, dx \, d\xi = \int 1_{u_0(x) > \xi}(1 - 1_{u_0(x) > \xi}) \, dx \, d\xi = 0.
\]

Now, the fact that $F^-$ is a kinetic function for all $t \in [0,T]$ implies that $F^-(t, x, \xi) \in \{0,1\}$ for a.e. $(x, \xi)$ and all $t \in [0,T]$ and therefore there exists $u : \Omega \times [0,T] \times \mathbb{R}^N \times \mathbb{R} \to \mathbb{R}$ such that $F^-(t, x, \xi) = 1_{u(t,x) > \xi}$ for a.e. $(\omega, x, \xi)$ and all $t \in [0,T]$. Moreover, according to (2.7), it holds
\[
\mathbb{E} \text{ess sup}_{0 \leq t \leq T} \int_{\mathbb{R}^N} |u(t,x)| \, dx = \mathbb{E} \text{ess sup}_{0 \leq t \leq T} \int_{\mathbb{R}^N} \int_{\mathbb{R}} |\xi| \, d\nu_{t,x}(\xi) \, dx \, d\xi \leq C
\]
and
\[
\mathbb{E} \int_0^T \int_{\mathbb{R}^N} |u(t,x)|^2 \, dx dt = \mathbb{E} \int_0^T \int_{\mathbb{R}^N} \int_{\mathbb{R}} |\xi|^2 \, d\nu_{t,x}(\xi) \, dx dt \leq C
\]
hence $u \in L^1(\Omega; L^\infty(0,T; L^1(\mathbb{R}^N))) \cap L^2(\Omega; L^2(0,T; L^2(\mathbb{R}^N)))$ is a kinetic solution. \qed

Corollary 4.7 ($L^1$-contraction property). Let $u_1$ and $u_2$ be kinetic solutions to (1.1) with initial data $u_{1,0}$ and $u_{2,0}$, respectively. Then there exists $C > 0$ such that for a.e. $t \in [0,T]$
\[
\mathbb{E} \|u_1(t) - u_2(t)\|_{L^1} \leq \mathbb{E} \|u_{1,0} - u_{2,0}\|_{L^1}.
\]
Proof. First, we observe that the following identity holds true

\[(u_1 - u_2)^+ = \int_{\mathbb{R}} 1_{u_1 > \xi} 1_{u_2 > \xi} d\xi.\]

We proceed similarly to Theorem 4.6, apply Proposition 4.5 and obtain

\[E\| (u_1(t) - u_2(t))^+\|_{L^1_x} = E \int F_1^-(t, x, \xi) \bar{F}_2^- (t, x, \xi) d\xi dx \leq E \int F_1,0(x, \xi) \bar{F}_2,0(x, \xi) d\xi dx = E\| u_{1,0} - u_{2,0}\|_{L^1_x}\]

which completes the proof. \(\square\)

5. Existence

In the existence part of the proof of Theorem 2.7 we make use of the so-called Bhatnagar-Gross-Krook approximation (BGK), which is a standard tool in the deterministic setting. Its stochastic counterpart was established in [20] and could be understood as an alternative proof of existence to [10].

Let us now briefly describe this method. Towards the proof of existence of a kinetic solution to (1.1), which is (formally speaking) a distributional solution to the kinetic formulation (2.3), we consider the following BGK model

\[
\begin{aligned}
dF^\varepsilon + \nabla F^\varepsilon \cdot a d z - \partial_\xi F^\varepsilon b d z &= -\partial_\xi F^\varepsilon g dW + \frac{1}{2} \partial_\xi (G^2 \partial_\xi F^\varepsilon) dt + \frac{1}{\varepsilon} u^\varepsilon - F^\varepsilon dt, \\
F^\varepsilon(0) &= F_0^\varepsilon,
\end{aligned}
\]

where the local density of particles is defined by

\[u^\varepsilon(t, x) = \int_{\mathbb{R}} (F^\varepsilon(t, x, \xi) - 1_{0 > \xi}) d\xi.\]

In other words, the unknown kinetic measure is replaced by a right hand side written in terms of \(F^\varepsilon\). Heuristically, solving (5.1) is significantly easier and can be reduced to solving the homogeneous problem

\[
\begin{aligned}
dX + \nabla X \cdot a d z - \partial_\xi X b d z &= -\partial_\xi X g dW + \frac{1}{2} \partial_\xi (G^2 \partial_\xi F) dt, \\
X(s) &= X_0,
\end{aligned}
\]

establishing the properties of its solution operator and employing Duhamel’s principle. The final idea is that, as the microscopic scale \(\varepsilon\) vanishes, \(F^\varepsilon\) converges to \(F\) which is a generalized kinetic solution to (1.1). The Reduction theorem, Theorem 4.6, then applies and as a consequence there exists \(u\) such that \(F = 1_{u > \xi}\) is the unique kinetic solution to (1.1).

We point out that the expected regularity of solutions to (5.4), (5.1) is low, namely \(L^\infty\), and therefore the rigorous treatment of the above outlined technique requires a suitable notion of weak solution in the context of rough paths. Here the usual definition of distributional solutions encounters several obstacles due to rough regularity of the driving signals. Motivated by the discussion in Section 2, we introduce weak formulations which correspond to solving (5.4) and (5.1) in the sense of distributions but do not involve any rough path driven terms.

In the sequel, we will use the following notation: for \(\alpha \in \mathbb{R}\) we denote by \(\chi_\alpha : \mathbb{R} \to \mathbb{R}\) the so-called equilibrium function which is defined as

\[\chi_\alpha(\xi) = 1_{0 < \xi < \alpha} - 1_{\alpha < \xi < 0}.\]
5.1. **Rough transport equation.** This subsection is devoted to the study of the auxiliary equation (5.2). Our approach here as well as in Subsection 5.2 is rough-pathwise and therefore we work with one fixed realization \( \omega \) and the probability space \((\Omega, \mathcal{F}, P)\) remains hidden. The stochastic integral will then re-appear in Section 5.3 for the final passage to the limit.

It can be seen that the Stratonovich form of (5.2) reads as follows
\[
dX + \nabla X \cdot a \, dz - \partial_t X \, b \, dz = -\partial_t X \, g \circ dW + \frac{1}{4} \partial_t X \partial_t G^2 \, dt.
\]
Now we recall that (the stochastic process) \( \Lambda \) is the joint lift of \( z \) and \( W \) constructed in (2.1). Let us fix its one realization \( \Lambda(\omega) \). This leads us us to the study of the following rough transport equation
\[
dX = \left( \frac{\partial_t X}{\nabla X} \right) \cdot \left( \begin{array}{cc} \frac{1}{2} \partial_t G^2 & b \\ 0 & -a \end{array} \right) \, d(t, \Lambda(\omega)),
\]
However, for notational simplicity (and with a slight abuse of notation) we may rather write
\[
dX + \nabla X \cdot a \, dz - \partial_t X \, b \, dz = -\partial_t X \, g \, dw + \frac{1}{4} \partial_t X \partial_t G^2 \, dt,
\]
where \( w \) denotes the corresponding realization of the Stratonovich lift of \( W \) and the cross-iterated integrals between \( t, z \) and \( W \) are not explicitly mentioned.

We introduce two notions of solution to (5.4). The first definition follows the usual approach from rough path theory which is based on the approximation of the driving signals (see [4, Definition 3]).

**Definition 5.1.** Let \((z, w)\) be a geometric Hölder \( p \)-rough path and suppose that \((z^n, w^n)\) is a sequence of Lipschitz paths such that
\[
S_p(z^n, w^n) \equiv (z^n, w^n) \longrightarrow (z, w)
\]
uniformly on \([0,T]\) and
\[
\sup_n \| (z^n, w^n) \|_{p-\text{Hölder}} < \infty.
\]
Assume that for each \( n \in \mathbb{N} \)
\[
dX^n + \nabla X^n \cdot a \, dz^n - \partial_t X^n \, b \, dz^n = -\partial_t X^n \, g \, dw^n + \frac{1}{4} \partial_t X^n \partial_t G^2 \, dt
\]
has a unique solution \( X^n \) which belongs to \( C^1_b([0,T] \times \mathbb{R}^N \times \mathbb{R}) \). Then any limit point (in the uniform topology) of \((X^n)\) is called a solution for the rough PDE, denoted formally by (5.4).

Such a solution can be constructed by making use of the method of characteristics established in [4] provided the initial datum \( X_0 \) is sufficiently regular. To be more precise, the associated characteristic system is given by
\[
d\varphi^0_t = -b(\varphi_t) \, dz + g(\varphi_t) \, dw - \frac{1}{4} \partial_t G^2(\varphi_t) \, dt,
\]
\[
d\varphi^c_t = a(\varphi_t) \, dz.
\]
where \( \varphi^0_t \) and \( \varphi^c_t \) describe the evolution of the \( \xi \)-coordinate and \( x \)-coordinate, respectively, of the characteristic curve. Let us denote by \( \varphi_{s,t}(x, \xi) \) the solution of (5.5) starting from \((x, \xi)\) at time \( s \). It follows from [17, Proposition 11.11] that under our assumptions \( \varphi \) defines a flow of \( C^2 \)-diffeomorphisms and we denote by \( \psi \) the corresponding inverse flow.

Our first existence and uniqueness result for (5.4) is taken from [4, Theorem 4] and reads as follows.

**Proposition 5.2.** Let \( X_0 \in C^1_b([\mathbb{R}^N \times \mathbb{R}) \). Then there exists a unique solution to (5.4) given explicitly by
\[
X(t, x, \xi; s) = X_0(\psi_{s,t}(x, \xi)).
\]
We conclude that the solution operator \( S(t,s)X_0 = X_0(\psi_{s,t}(x,\xi)) \) is well defined on \( C^1_b(\mathbb{R}^N \times \mathbb{R}) \). Nevertheless, as the right hand side makes sense even for more general initial conditions \( X_0 \) which do not necessarily fulfill the assumptions of Proposition 5.2, the domain of definition of the operator \( S \) can be extended. In particular, since diffeomorphisms preserve sets of measure zero the above is well defined also if \( X_0 \) is only defined almost everywhere. In this case, we define consistently
\[
S(t,s)X_0 = X_0(\psi_{s,t}(x,\xi)), \quad 0 \leq s \leq t \leq T.
\]

The aim is to show that this extension is the solution operator to (5.4) in a weak sense. Towards this end, it is necessary to weaken the assumption upon the initial condition and therefore a suitable notion of weak solution is required.

In the following proposition we establish basic properties of the operator \( S \).

**Proposition 5.3.** Let \( \mathcal{S} = \{ S(t,s), 0 \leq s \leq t \leq T \} \) be defined as above. Then

(i) for any \( p \in [1,\infty] \), \( S \) is a family of linear operators on \( L^p(\mathbb{R}^N \times \mathbb{R}) \) which is uniformly bounded in the operator norm, i.e., there exists \( C > 0 \) such that for any \( X_0 \in L^p(\mathbb{R}^N \times \mathbb{R}) \),
\[
0 \leq s \leq t \leq T,
\]
\[
\| S(t,s)X_0 \|_{L^p_{x,\xi}} \leq C \| X_0 \|_{L^p_{x,\xi}},
\]
\[
(5.6)
\]
(ii) \( S \) verifies the semigroup law
\[
S(t,s) = S(t,r) \circ S(r,s), \quad 0 \leq s \leq r \leq t \leq T,
\]
\[
S(s,s) = \text{Id}, \quad 0 \leq s \leq T.
\]

**Proof.** The proof of (ii) as well as (i) in the case of \( p = \infty \) follows immediately from the definition of the operator \( S \) and the flow property of \( \psi \). In order to show (i) for \( p \in [1,\infty) \), we observe that due to [17, Proposition 11.11] the map
\[
(s,t,x,\xi) \mapsto |J\psi_{s,t}(x,\xi)|,
\]
where \( J \) denotes the Jacobian, is bounded from above and below by a positive constant. Consequently, using the notation of Proposition 5.2, we have
\[
\| X_0 \|_{L^p_{x,\xi}}^p = \int_{\mathbb{R}^N \times \mathbb{R}} |X(t,\varphi_{s,t}(x,\xi);s)|^p \, dx \, d\xi
\]
\[
= \int_{\mathbb{R}^N \times \mathbb{R}} |X(t,x,\xi;s)|^p |J\psi_{s,t}(x,\xi)| \, dx \, d\xi \geq C \| X \|_{L^p_{x,\xi}}^p.
\]
\[
\]
\[
\]
\[
\]
\[
\]

**Definition 5.4.** \( X : [s,T] \times \mathbb{R}^N \times \mathbb{R} \rightarrow \mathbb{R} \) is called a weak solution to (5.4) provided
\[
dX(t,\varphi_{s,t}) = 0
\]
\[
X(s) = X_0
\]
holds true in the sense of \( D'(\mathbb{R}^N \times \mathbb{R}) \).

**Corollary 5.5.** Let \( X_0 \in L^\infty(\mathbb{R}^N \times \mathbb{R}) \). Then there exists a unique \( X \in L^\infty([s,T] \times \mathbb{R}^N \times \mathbb{R}) \) that is a weak solution to (5.4). Moreover, it is represented by
\[
X = S(t,s)X_0.
\]

**Proof.** In order to prove existence, observe that if \( X \) is given by the above representation then clearly for all \( 0 \leq s \leq t \leq T \)
\[
X(t,\varphi_{s,t}(x,\xi);s) = X_0(\psi_{s,t} \circ \varphi_{s,t}(x,\xi)) = X_0(x,\xi) \quad \text{for a.e.} \ (x,\xi)
\]
and hence \( X \) solves (5.7). However, let us also give another proof of existence which in addition justifies Definition 5.4 as the appropriate notion of weak solution to (5.4). To this end, let us consider
smooth approximations of $X_0$, namely, let $(\delta_k)$ be an approximation to the identity on $\mathbb{R}^N \times \mathbb{R}$ and set $X^\delta_0 = X_0 * \delta_k$. According to Proposition 5.2, there exists a unique $X^\delta$ which solves (5.4) with the initial condition $X^\delta_0$ and is given by $X^\delta(t; s) = X^\delta_0(\psi_{s,t})$ hence
\begin{equation}
X^\delta(t, \varphi_{s,t}(x, \xi); s) = X^\delta_0(x, \xi).
\end{equation}
Moreover,
\[ X^\delta \xrightarrow{w^*} X_0 \quad \text{in} \quad L^\infty(\mathbb{R}^N \times \mathbb{R}) \]
so
\[ X^\delta(t; s) \xrightarrow{w^*} X_0(\psi_{s,t}) \quad \text{in} \quad L^\infty(\mathbb{R}^N \times \mathbb{R}) \quad \forall t \in [s, T] \]
which justifies the passage to the limit in (5.8) and completes the proof.

Regarding uniqueness, due to linearity it is enough to show that any weak solution with $X_0 = 0$ vanishes identically. Let $X$ be such a weak solution, i.e. it holds
\[ \langle X(t, \varphi_{s,t}; s), \phi \rangle = 0 \quad \forall \phi \in C^1_c(\mathbb{R}^N \times \mathbb{R}). \]
Testing by the mollifier $\delta_\varepsilon$ we deduce that
\[ 0 = [X(t, \varphi_{s,t}; s)]^\delta(\psi_{s,t}(x, \xi)) = \int_{\mathbb{R}^N \times \mathbb{R}} X(t, \varphi_{s,t}(y, \zeta); s) \delta_\varepsilon(\psi_{s,t}(x, \xi) - (y, \zeta)) \, dy \, d\zeta \]
and since for a.e. $(x, \xi)$ the right hand side converges to $X(t, \varphi_{s,t} \circ \psi_{s,t}(x, \xi); s) = X(t, x, \xi; s)$, the claim follows. \hfill $\square$

**Lemma 5.6.** For all $0 \leq s \leq t \leq T$ it holds true that
\[ S(s, t)1_{0 > \xi} - 1_{0 \geq \xi} = 0. \]

**Proof.** Let us only prove the result for $0 \leq s \leq t \leq T$, the other case being handled similarly. It follows from (2.2) that for all $x \in \mathbb{R}^N$ the function $\varphi^0_{s,t}(x, 0) \equiv 0$ is a solution to the first equation in (5.5). Moreover, since the solution to (5.5) is unique, we deduce that
\[ \varphi^0_{s,t}(x, \xi) \begin{cases} \geq 0, & \text{if } \xi \geq 0, \\ \leq 0, & \text{if } \xi \leq 0, \end{cases} \]
which yields the claim. \hfill $\square$

### 5.2. Solution to the BGK model.
Throughout this section we continue with our rough-pathwise analysis, that is, we consider one fixed realization of the driving path $\Lambda(\omega)$ and therefore the underlying probability space $(\Omega, \mathcal{F}, \mathbb{P})$ remains hidden.

We will apply the auxiliary results for the rough transport equation (5.4) and establish existence and uniqueness for the BGK model (5.1). First of all, it is necessary to specify in which sense (5.1) is to be solved. As it can be seen in Definition 5.1, the solution given by Proposition 5.2 satisfies the equation (5.4) only on a formal level. This obstacle was overcome by the definition of weak solution (see Definition 5.4) which also permitted generalization to less regular initial data, namely $X_0 \in L^\infty(\mathbb{R}^N \times \mathbb{R})$. We continue in this fashion and define the notion of weak solution to the BGK model similarly.

Concerning the initial data for the BGK model (5.1), let us simply set $F^\varepsilon_0 = 1_{u_0 > \xi}$ which is sufficient for our purposes. Nevertheless, our proof of existence and convergence of (5.1) would remain valid if we considered $F^\varepsilon_0 = 1_{u^\varepsilon_0 > \xi}$ with $(u^\varepsilon_0)$ being a suitable approximation of $u_0$.

**Definition 5.7.** Let $\varepsilon > 0$. Then $F^\varepsilon \in L^\infty([0, T] \times \mathbb{R}^N \times \mathbb{R})$ satisfying $F^\varepsilon - 1_{0 > \xi} \in L^1([0, T] \times \mathbb{R}^N \times \mathbb{R})$ is called a weak solution to the BGK model (5.1) with initial condition $F^\varepsilon_0$ provided
\begin{equation}
\begin{aligned}
dF^\varepsilon(t, \varphi_{0,t}) &= \frac{1_{u^\varepsilon(t) > \xi} \circ \varphi_{0,t} - F^\varepsilon(t, \varphi_{0,t})}{\varepsilon} \, dt \\
F^\varepsilon(0) &= F^\varepsilon_0
\end{aligned}
\end{equation}
holds true in the sense of \( \mathcal{D}'(\mathbb{R}^N \times \mathbb{R}) \).

The result reads as follows.

**Theorem 5.8.** For any \( \varepsilon > 0 \), there exists a unique weak solution of the BGK model (5.1) and is represented by

\[
(5.10) \quad F^\varepsilon(t) = e^{-\frac{t}{\varepsilon}}S(t,0)F_0^\varepsilon + \frac{1}{\varepsilon} \int_0^t e^{-\frac{t-s}{\varepsilon}} S(t,s) 1_{u^\varepsilon(s) > \xi} \, ds.
\]

**Proof.** By Duhamel’s principle, the problem (5.9) admits an equivalent integral representation

\[
F^\varepsilon(t, \varphi_{0,t}) = e^{-\frac{t}{\varepsilon}}F_0^\varepsilon + \frac{1}{\varepsilon} \int_0^t e^{-\frac{t-s}{\varepsilon}} 1_{u^\varepsilon(s) > \xi}(\varphi_{0,s}) \, ds
\]

which can be rewritten as (5.10). Recall, that the local densities are defined as follows

\[
(5.11) \quad u^\varepsilon(t,x) = \int_{\mathbb{R}} (F^\varepsilon(t,x,\xi) - 1_{0 > \xi}) \, d\xi
\]

hence the function \( F^\varepsilon \) is not integrable with respect to \( \xi \). For the purpose of the proof it is therefore more convenient to consider \( f^\varepsilon(t) = F^\varepsilon(t) - 1_{0 > \xi} \) instead and prove that it exists and is given by a corresponding integral representation. Due to Lemma 5.6 and Corollary 5.5, \( 1_{0 > \xi} \) is the unique weak solution to (5.4) hence \( h^\varepsilon \) solves

\[
(5.12) \quad df^\varepsilon(t,\varphi_{0,t}) = \frac{\chi_{u^\varepsilon(t)} \circ \varphi_{0,t} - f^\varepsilon(t,\varphi_{0,t})}{\varepsilon} \, dt \quad \text{and } f^\varepsilon(0) = \chi_{u_0}
\]

in the sense of distributions. By a similar reasoning as above, (5.12) has the integral representation

\[
(5.13) \quad f^\varepsilon(t) = e^{-\frac{t}{\varepsilon}}S(t,0)\chi_{u_0} + \frac{1}{\varepsilon} \int_0^t e^{-\frac{t-s}{\varepsilon}} S(t,s) \chi_{u^\varepsilon(s)} \, ds
\]

and thus can be solved by a fixed point method. According to the identity

\[
\int_{\mathbb{R}} |\chi_{\alpha} - \chi_{\beta}| \, d\xi = |\alpha - \beta|, \quad \alpha, \beta \in \mathbb{R}
\]

some space of \( \xi \)-integrable functions would be well suited to deal with the nonlinearity term \( \chi_{u^\varepsilon} \). Let us denote \( \mathcal{H} = L^\infty(0,T; L^1(\mathbb{R}^N \times \mathbb{R})) \) and show that the mapping

\[
(\mathcal{H}g)(t) = e^{-\frac{t}{\varepsilon}}S(t,0)\chi_{u_0} + \frac{1}{\varepsilon} \int_0^t e^{-\frac{t-s}{\varepsilon}} S(t,s) \chi_{v(s)} \, ds,
\]

where the local density \( v(s) = \int_{\mathbb{R}} g(s,\xi) \, d\xi \) is defined consistently with (5.11), is a contraction on \( \mathcal{H} \). Let \( g, g_1, g_2 \in \mathcal{H} \) with corresponding local densities \( v, v_1, v_2 \). By Proposition 5.3 and the assumptions on initial data, we arrive at

\[
\|(\mathcal{H}g)(t)\|_{L^1_{v \xi}} \leq e^{-\frac{t}{\varepsilon}}\|S(t,0)\chi_{u_0}\|_{L^1_{\xi}} + \frac{1}{\varepsilon} \int_0^t e^{-\frac{t-s}{\varepsilon}} \|S(t,s)\chi_{v(s)}\|_{L^1_{\xi}} \, ds
\]

\[
\leq C \left( 1 + \sup_{0 \leq s \leq t} \|v(s)\|_{L^1} \right)
\]

with a constant independent on \( t \), hence

\[
\|\mathcal{H}g\|_{\mathcal{H}} \leq C \left( 1 + \|g\|_{\mathcal{H}} \right) < \infty.
\]

2By \( 1_{u^\varepsilon(t) > \xi} \circ \varphi_{0,t} \) we denote the composition of \( (x, \xi) \mapsto 1_{u^\varepsilon(t,x) > \xi} \) with \( (x, \xi) \mapsto \varphi_{0,t}(x, \xi) \).
Next, we estimate
\[
\| (\mathcal{H} g_1)(t) - (\mathcal{H} g_2)(t) \|_{L^1_{\tau,\xi}} \leq \frac{1}{\varepsilon} \int_0^t e^{-\frac{t-s}{\varepsilon}} \| \mathcal{S}(t, s)(\chi_{v_1(s)} - \chi_{v_2(s)}) \|_{L^1_{\tau,\xi}} \, ds
\]
\[
\leq C \varepsilon \int_0^t e^{-\frac{t-s}{\varepsilon}} \| \mathcal{S}(t, s)(\chi_{v_1(s)} - \chi_{v_2(s)}) \|_{L^1_{\tau,\xi}} \, ds
\]
\[
\leq C \int_0^t e^{-\frac{t-s}{\varepsilon}} \| v_1(s) - v_2(s) \|_{L^1_{\tau,\xi}} \, ds.
\]
so
\[
\| \mathcal{H} g_1 - \mathcal{H} g_2 \|_{\mathcal{H}} \leq C (1 - e^{-\frac{T}{\varepsilon}}) \| g_1 - g_2 \|_{\mathcal{H}}
\]
and according to the Banach fixed point theorem, the mapping \( \mathcal{H} \) has a unique fixed point in \( \mathcal{H} \) provided \( T \) was small enough. Nevertheless, this condition on \( T \) can be easily removed by considering the equation on smaller intervals \([0, \tilde{T}], [\tilde{T}, 2\tilde{T}], \) etc. As a consequence, we obtain the existence of a unique weak solution to (5.1) that is given by (5.10) and the proof is complete. \( \square \)

5.3. Convergence of the BGK model. In this final section we establish another weak formulation of the BGK model (5.1) that actually includes the stochastic integral and is therefore better suited for proving the convergence to (2.4). To be more precise, we prove the following result, which will complete the proof of Theorem 2.7.

**Theorem 5.9.** Let \( f^\varepsilon = F^\varepsilon - 1_{0>\xi}. \) Then \( (f^\varepsilon) \) converges weak-star in \( L^\infty(\Omega \times [0, T] \times \mathbb{R}^N \times \mathbb{R}) \) to the equilibrium function \( \chi_u \), where \( u \) is the unique kinetic solution to the conservation law (1.1).

We start with an auxiliary result.

**Lemma 5.10.** \( F^\varepsilon \in L^\infty(\Omega \times [0, T] \times \mathbb{R}^N \times \mathbb{R}) \) satisfies the following weak formulation of (5.1): let \( \phi \in C^1_c(\mathbb{R}^N \times \mathbb{R}) \), then
\[
\langle F^\varepsilon(t), \phi(\theta_{0,t}) \rangle = -\langle \partial_\xi F^\varepsilon g, \phi(\theta_{0,t}) \rangle \, dW + \frac{1}{2} \langle \partial_\xi (G^2 \partial_\xi F^\varepsilon), \phi(\theta_{0,t}) \rangle \, dt + \frac{1}{\varepsilon} \langle 1_{u^\varepsilon > \xi} - F^\varepsilon, \phi(\theta_{0,t}) \rangle.
\]
Moreover, it is progressively measurable.

**Proof.** In order to verify (5.14), we test (5.10) by \( \phi(\theta_{0,t}) \) and obtain
\[
\langle F^\varepsilon(t), \phi(\theta_{0,t}) \rangle = e^{-\frac{t}{\varepsilon}} \langle S(t, 0) F^\varepsilon_0, \phi(\theta_{0,t}) \rangle + \frac{1}{\varepsilon} \int_0^t e^{-\frac{t-s}{\varepsilon}} \langle S(t, s) 1_{u^\varepsilon > \xi, \phi(\theta_{0,s}(\theta_{s,t}))} \rangle \, ds.
\]
According to Duhamel’s principle, it is enough to verify that if \( X_0 \in L^\infty(\mathbb{R}^N \times \mathbb{R}) \) then
\[
\langle X(t) := S(t, s) X_0, \phi(\theta_{0,s}(\theta_{s,t})) \rangle
\]
solves
\[
\langle X(t), \phi(\theta_{0,s}(\theta_{s,t})) \rangle = -\langle \partial_\xi X g, \phi(\theta_{0,s}(\theta_{s,t})) \rangle \, dW + \frac{1}{2} \langle \partial_\xi (G^2 \partial_\xi X), \phi(\theta_{0,s}(\theta_{s,t})) \rangle \, dt,
\]
\[
\langle X(s), \phi(\theta_{0,s}) \rangle = \langle X_0, \phi(\theta_{0,s}) \rangle.
\]
Let us now justify (5.15). It was shown in [12, Theorem 8] that the RDE solution to the characteristic system (5.5) corresponding to the joint lift \( \Lambda \) constructed in (2.1) can be obtained as limit of SDE solutions to
\[
\begin{align*}
    d\phi_t^{n,0} &= g(\phi_t^n) \circ dW - \frac{1}{4} \partial_\xi G^2(\phi_t^n) \, dt - b(\phi_t^n) \, d\zeta^n, \\
    d\phi_t^{n,x} &= a(\phi_t^n) \, d\zeta^n,
\end{align*}
\]
where the corresponding lifts \((z^n)\) approximate \(z\) as in Definition 3.1. Let \(\psi^n\) and \(\theta^n\), respectively, denote the inverse flows corresponding to \((5.16)\) and
\[
\begin{align*}
\left. \begin{array}{c}
\pi_{n,t}^0 = -b(x^n) \, dx^n, \\
\pi_{n,t}^x = a(x^n) \, dz^n,
\end{array} \right\}
\end{align*}
\]
respectively. If \(X_0 \in C^0_c(\mathbb{R}^N \times \mathbb{R})\) and \(\phi \in C^1_c(\mathbb{R}^N \times \mathbb{R})\), then setting \(X^n(t) = X_0(\psi^n_{s,t})\) yields by the Itô formula
\[
\begin{align*}
\frac{d\langle X^n(t), \phi(\theta^n_{0,s}(\theta^n_{s,t})) \rangle}{dt} &= -\langle \partial_x X^n g, \phi(\theta^n_{0,s}(\theta^n_{s,t})) \rangle \, dW + \frac{1}{2} \langle \partial_x (\partial^2_x X^n), \phi(\theta^n_{0,s}(\theta^n_{s,t})) \rangle dt, \\
\langle X(s), \phi(\theta^n_{0,s}) \rangle &= \langle X_0, \phi(\theta^n_{0,s}) \rangle.
\end{align*}
\]
As mentioned above, due to [12, Theorem 8], \(\psi^n(x, \xi) \to \psi(x, \xi)\) uniformly in \(t \in [0, T]\) in probability, where \(\psi\) is the RDE inverse flow to \((5.5)\). Due to invariance of the Lip\(^1\)-norm under translation, we deduce (for a subsequence) that \(\psi^n \to \psi\) uniformly in \(t \in [0, T]\) and \((x, \xi) \in \mathbb{R}^N \times \mathbb{R}\) a.s. Moreover, \(\theta^n \to \theta\) uniformly in \(s, t \in [0, T]\), \((x, \xi) \in \mathbb{R}^N \times \mathbb{R}\) (see [4, Theorem 4]) and the same holds true for their first and second order derivatives with respect to \(\xi\). Therefore, we may pass to the limit, apply dominated convergence theorem (for both Lebesgue and stochastic integral, see [35, Theorem 32] for the latter one) and \((5.15)\) follows. If \(X_0 \in L^\infty(\mathbb{R}^N \times \mathbb{R})\) then we consider its smooth approximation \(X^\delta_0\), apply the previous result and pass to the limit.

The progressive measurability follows immediately from the construction. \(\square\)

As the next step we prove a stochastic version of Proposition 5.3(i) for \(p = 1\).

**Lemma 5.11.** The family \(\mathcal{S} = \{S(t, s), 0 \leq s \leq t \leq T\}\) consists of bounded linear operators on \(L^1(\Omega \times \mathbb{R}^N \times \mathbb{R})\). In particular, if \(X_0 \in L^1(\Omega \times \mathbb{R}^N \times \mathbb{R})\) then
\[
\sup_{0 \leq s \leq T} \mathbb{E} \sup_{s \leq t \leq T} \|S(t, s)X_0\|_{L^1_{x, \xi}} \leq C \mathbb{E}\|X_0\|_{L^1_{x, \xi}}.
\]

**Proof.** Assume in addition that \(X_0\) is nonnegative, bounded and compactly supported in \((x, \xi)\). In general, \((5.15)\) holds true for all \(\phi \in C^1_c(\mathbb{R}^N \times \mathbb{R})\), nevertheless, since \(X \in L^1(\mathbb{R}^N \times \mathbb{R})\) a.s., which follows from Proposition 5.3, the assumption on the test function \(\phi\) can be relaxed and we may take \(\phi \equiv 1\). Taking expectation, the stochastic integral vanishes due to the additional assumption and we obtain
\[
\mathbb{E}\|S(t, s)X_0\|_{L^1_{x, \xi}} \leq \mathbb{E}\|X_0\|_{L^1_{x, \xi}}.
\]

Besides, the nonnegativity assumption can be immediately omitted by splitting \(X_0\) into negative and positive part. In the general case of \(X_0 \in L^1(\Omega \times \mathbb{R}^N \times \mathbb{R})\), we approximate \(X_0\) in \(L^1(\Omega \times \mathbb{R}^N \times \mathbb{R})\) by \(X^\delta_0\) bounded and compactly supported such that
\[
\mathbb{E}\|X^\delta_0\|_{L^1_{x, \xi}} \leq \mathbb{E}\|X_0\|_{L^1_{x, \xi}}.
\]

Apply \((5.18)\) to \(X^\delta_0\). Due to linearity of \(S(t, s)\) it implies that \(S(t, s)X^\delta_0\) is Cauchy in \(L^1(\Omega \times \mathbb{R}^N \times \mathbb{R})\), the limit is necessarily \(S(t, s)X_0\) so Fatou’s lemma yields \((5.18)\) and the first claim follows.

To obtain, \((5.17)\), we test \((5.15)\) again by \(\phi \equiv 1\), take supremum and expectation. Applying Burkholder-Davis-Gundy’s and weighted Young’s inequalities and \((5.18)\) we obtain
\[
\mathbb{E}\sup_{s \leq t \leq T}\|S(t, s)X_0\|_{L^1_{x, \xi}} \leq \mathbb{E}\|X_0\|_{L^1_{x, \xi}} + \mathbb{E}\sup_{s \leq t \leq T}\left| \int_s^t (X, \partial_t g) \, dW \right| \leq \mathbb{E}\|X_0\|_{L^1_{x, \xi}} + C\mathbb{E}\left( \int_s^T \|X\|_{L^1_{x, \xi}}^2 \, dt \right)^{1/2} \leq C\mathbb{E}\|X_0\|_{L^1_{x, \xi}} + \frac{1}{2}\mathbb{E}\sup_{s \leq t \leq T}\|S(t, s)X_0\|_{L^1_{x, \xi}}.
\]
and the proof is complete. □

Proof of Theorem 5.9. In view of Lemma 5.10 we remark that $F^\varepsilon$ also satisfies the following weak formulation of (5.9): let $\phi \in C_c^\infty(\mathbb{R}^N \times \mathbb{R})$ and $\alpha \in C^1_c([0, T))$ then

$$
\int_0^T \left\langle F^\varepsilon(t), \phi(\theta_t) \right\rangle \partial_t \alpha(t) \, dt + \left\langle F_0^\varepsilon, \phi \right\rangle \alpha(0)
$$

$$
= \int_0^T \left\langle \partial_t F^\varepsilon(t) g, \phi(\theta_t) \right\rangle \alpha(t) \, dW(t) - \frac{1}{2} \int_0^T \left\langle \partial_\xi (G^2 \partial_\xi F^\varepsilon(t)), \phi(\theta_t) \right\rangle \alpha(t) \, dt
$$

$$
- \frac{1}{\varepsilon} \int_0^T \left\langle 1_{u^\varepsilon(t) > \xi} - F^\varepsilon(t), \phi(\theta_t) \right\rangle \alpha(t) \, dt
$$

(5.19)

and a similar expression holds true for $F^\varepsilon$, namely, it satisfies the weak formulation of (5.12). Taking the limit on the left hand side of (5.19) is quite straightforward. Indeed, according to the representation formula (5.10) it holds that the set ($F^\varepsilon$) is bounded in $L^\infty(\Omega \times [0, T] \times \mathbb{R}^N \times \mathbb{R})$, more precisely, $F^\varepsilon \in [0, 1]$, $\varepsilon \in (0, 1)$. Therefore, by the Banach-Alaoglu theorem, there exists $F \in L^\infty(\Omega \times [0, T] \times \mathbb{R}^N \times \mathbb{R})$ such that, up to subsequences,

$$
F^\varepsilon \xrightarrow{w^*} F \text{ in } L^\infty(\Omega \times [0, T] \times \mathbb{R}^N \times \mathbb{R}).
$$

As a consequence, a.s.

$$
\int_0^T \left\langle F^\varepsilon(t), \phi(\theta_t) \right\rangle \partial_t \alpha(t) \, dt \longrightarrow \int_0^T \left\langle F(t), \phi(\theta_t) \right\rangle \partial_t \alpha(t) \, dt,
$$

$$
\frac{1}{2} \int_0^T \left\langle \partial_\xi (G^2 \partial_\xi F^\varepsilon(t)), \phi(\theta_t) \right\rangle \alpha(t) \, dt \longrightarrow \frac{1}{2} \int_0^T \left\langle \partial_\xi (G^2 \partial_\xi F(t)), \phi(\theta_t) \right\rangle \alpha(t) \, dt,
$$

and due to dominated convergence for stochastic integrals

$$
\int_0^T \left\langle \partial_t F^\varepsilon(t) g, \phi(\theta_t) \right\rangle \alpha(t) \, dW(t) \longrightarrow \int_0^T \left\langle \partial_t F(t) g, \phi(\theta_t) \right\rangle \alpha(t) \, dW(t).
$$

According to the hypotheses on the initial data,

$$
\left\langle F_0^\varepsilon, \phi \right\rangle \alpha(0) = \left\langle 1_{u_0 > \xi}, \phi \right\rangle \alpha(0).
$$

Furthermore, multiplying (5.19) by $\varepsilon$ yields a.s.

$$
\int_0^T \left\langle 1_{u^\varepsilon(t) > \xi} - F^\varepsilon(t), \phi(\theta_t) \right\rangle \alpha(t) \, dt \longrightarrow 0.
$$

(5.21)

However, due to the fact that $1_{u^\varepsilon > \xi} - F^\varepsilon \in [0, 1]$, the condition on test functions can be relaxed so that if $\phi \in L^1([0, T] \times \mathbb{R}^N \times \mathbb{R})$ we may test by $\phi(t, \pi_t)$ to deduce

$$
1_{u^\varepsilon > \xi} - F^\varepsilon \xrightarrow{w^*} 0 \text{ in } L^\infty([0, T] \times \mathbb{R}^N \times \mathbb{R})
$$

and, in particular,

$$
\partial_\xi 1_{u^\varepsilon > \xi} - \partial_\xi F^\varepsilon \xrightarrow{w^*} 0 \text{ in } \mathcal{D}'((0, T) \times \mathbb{R}^N \times \mathbb{R}).
$$

(5.22)

In order to obtain the convergence in the remaining term of (5.19) and in view of the kinetic formulation (2.4), we need to show that the term $\frac{1}{\varepsilon} \left(1_{u^\varepsilon > \xi} - F^\varepsilon\right)$ can be written as $\partial_\xi m^\varepsilon$ where $m^\varepsilon$ is a nonnegative measure over $[0, T] \times \mathbb{R}^N \times \mathbb{R}$ bounded uniformly in $\varepsilon$. If we define

$$
m^\varepsilon(\xi) = \frac{1}{\varepsilon} \int_{-\infty}^\xi \left(1_{u^\varepsilon > \xi} - F^\varepsilon(\zeta)\right) \, d\zeta = \frac{1}{\varepsilon} \int_{-\infty}^\xi \left(\chi_{u^\varepsilon}(\zeta) - f^\varepsilon(\zeta)\right) \, d\zeta,
$$

(5.23)

it is easy to check that $m^\varepsilon \geq 0$ since $F^\varepsilon \in [0, 1]$. Indeed, $m^\varepsilon(-\infty) = m^\varepsilon(\infty) = 0$ and $m^\varepsilon(t, x, \cdot)$ is increasing if $\xi \in (-\infty, u^\varepsilon(t, x))$ and decreasing if $\xi \in (u^\varepsilon(t, x), \infty)$.
Due to the convergence in (5.19) it can be seen that there exists a distribution $m$ such that

$$
(5.24) \quad \int_0^T \langle m^\varepsilon, \phi(t) \rangle \, dt \longrightarrow \int_0^T \langle m, \phi(t) \rangle \, dt,
$$

for any $\phi \in C^\infty_c([0, T) \times \mathbb{R}^N \times \mathbb{R})$. Furthermore, since all $m^\varepsilon$ are nonnegative, the same remains valid also for $m$.

We proceed with a uniform estimate for the $L^1$-norm of $u^\varepsilon$.

**Proposition 5.12.** The set of local densities $(u^\varepsilon)$ is bounded in $L^1(\Omega; L^\infty(0, T; L^1(\mathbb{R}^N)))$.

**Proof.** It follows from the definition of $u^\varepsilon$ in (5.11) and from (5.13) that

$$
(5.25) \quad u^\varepsilon(t) = e^{-t} \int_\mathbb{R} S(t, 0) \chi_{u_0} \, d\xi + \frac{1}{\varepsilon} \int_0^t e^{-\frac{t-s}{\varepsilon}} \int_\mathbb{R} S(t, s) \chi_{u^\varepsilon(s)} \, d\xi \, ds.
$$

Let us now define the following auxiliary function

$$
H(s) = \left| \int_\mathbb{R} S(t, s) \chi_{u^\varepsilon(s)} \, d\xi \right|.
$$

Then

$$
H(t) \leq e^{-\frac{t}{\varepsilon}} H(0) + (1 - e^{-\frac{t}{\varepsilon}}) \max_{0 \leq s \leq t} H(s)
$$

and we conclude that $H(t) \leq H(0)$, $t \in [0, T]$. In order to estimate $H(0)$, we apply Lemma 5.11 and obtain

$$
E \sup_{0 \leq t \leq T} \|u^\varepsilon(t)\|_{L^1_\varepsilon} \leq E \sup_{0 \leq t \leq T} \|S(t, 0) \chi_{u_0}\|_{L^1_{\varepsilon,x}} \leq C E \|u_0\|_{L^1_\varepsilon}.
$$

\[ \square \]

**Proposition 5.13.** For all $t \in [0, T]$ it holds true that

$$
(5.25) \quad \mathbb{E} \int_{[0,t] \times \mathbb{R}^N \times \mathbb{R}} dm^\varepsilon(t, x, \xi) + \mathbb{E} \langle f^\varepsilon(t), \xi \rangle \leq C (\mathbb{E} \|u_0\|_{L^1_\varepsilon}^2 + \mathbb{E} \|u_0\|_{L^1_\varepsilon}^2).
$$

**Proof.** Here we follow the ideas of [19]. Let $\{0 = t_0 < t_1 < \cdots < t_n = T\}$ be a partition of $[0, T]$ with step size $h$, to be chosen below. It follows immediately from the formula (5.13) that for every $t_i \in [0, T)$, $t \mapsto f^\varepsilon(t_i + t)$ is a solution to (5.12) on $[0, T - t_i]$ with the initial condition $f^\varepsilon(t_i)$ and therefore the corresponding version of (5.14) holds true, namely, for all $\phi \in C^1_c(\mathbb{R}^N \times \mathbb{R})$ and all $t \in [s, T]$

$$
d(f^\varepsilon(t), \phi(\theta_{t_i, s})) = -\langle \partial \phi f^\varepsilon g, \phi(\theta_{t_i, s}) \rangle \, dW + \frac{1}{2} \langle \partial \phi (\partial_\xi^2 f^\varepsilon), \phi(\theta_{t_i, s}) \rangle \, dt + \langle \partial_\phi m^\varepsilon, \phi(\theta_{t_i, s}) \rangle.
$$

Now, we need to test by $\phi(\xi) = \xi$. Since $f^\varepsilon \in L^\infty([0, T]; L^1(\mathbb{R}^N \times \mathbb{R}))$ a.s., we can test by constants. Therefore, let $\phi_R \in C^1_c(\mathbb{R})$ be a truncation of $\phi$ which preserves the sign, i.e. $\xi \text{ sgn} \phi_R(\xi) \geq 0$, and which satisfies $|\partial_\xi \phi_R|, |\partial^2_\xi \phi_R| \leq C$ uniformly in $R$. Using the weak formulation for $f^\varepsilon$ above we deduce

$$
\int_{t_i}^{t_{i+1}} \langle m^\varepsilon, \partial_\xi \phi_R(\theta^0_{t_i, s}) \rangle \, dt + \langle f^\varepsilon(t_{i+1}), \phi_R(\theta^0_{t_i, t_{i+1}}) \rangle = \langle f^\varepsilon(t_i), \phi_R \rangle - \int_{t_i}^{t_{i+1}} \langle \partial_\xi f^\varepsilon g, \phi_R(\theta^0_{t_i, s}) \rangle \, dW + \frac{1}{2} \int_{t_i}^{t_{i+1}} \langle \partial_\xi (\partial_\xi^2 f^\varepsilon), \phi_R(\theta^0_{t_i, s}) \rangle \, dt.
$$
Recall that $0 \leq \xi \operatorname{sgn} f^\varepsilon(t) \leq 1$ as a consequence of (5.13). Since also $\xi \operatorname{sgn} \theta_{t_0, t_{i+1}}(\xi) \geq 0$ due to the assumption (2.2), the second term on the left hand side is nonnegative and

$$
\mathbb{E} \int_{t_i}^{t_{i+1}} \langle m^\varepsilon, (\partial_t \phi_R) (\theta^0_{t_i, i}) \partial_t \theta^0_{t_i, i} \rangle \, dt + \mathbb{E} (f^\varepsilon(t_{i+1}), \phi_R (\theta^0_{t_0, i+1})) \\
\leq \mathbb{E} (f^\varepsilon(t_i), \phi_R) + \frac{1}{2} \mathbb{E} \int_{t_i}^{t_{i+1}} \langle \partial_x (G \partial_x f^\varepsilon), \phi_R (\theta^0_{t_i, i}) \rangle \, dt.
$$

(5.26)

Moreover, it follows from (5.13), Lemma 5.11 and Proposition 5.12 that

$$
\mathbb{E} \| f^\varepsilon(t) \|_{L^2_2} \leq \sup_{0 \leq s \leq t \leq T} \mathbb{E} \| S(t, s) \chi_{u^\varepsilon} \|_{L^2_2} \leq C \mathbb{E} \| u_0 \|_{L^1_2}
$$

and therefore the second term on the right hand side can be estimated as follows

$$
\frac{1}{2} \mathbb{E} \int_{t_i}^{t_{i+1}} \langle \partial_x (G^2 \partial_x f^\varepsilon), \phi_R (\theta^0_{t_i, i}) \rangle \, dr \leq Ch \mathbb{E} \| u_0 \|_{L^1_2}.
$$

Moreover, on a small time interval $\partial_x \theta^0_{t_0, t}$ remains close to its initial value, that is, choosing $h$ sufficiently small we may assume that for every $i$

$$
\inf_{t_i \leq t \leq t_{i+1}} \partial_x \theta^0_{t_i, t} \geq \frac{1}{2}.
$$

Therefore, if $i = 0$ then we estimate the first term on the right hand side of (5.26) by

$$
\mathbb{E} \langle f^\varepsilon_0, \phi_R \rangle = \mathbb{E} \langle \chi_{u_0}, \xi \rangle = \frac{1}{2} \mathbb{E} \| u_0 \|_{L^2_2}^2
$$

and obtain by Fatou’s lemma

$$
\frac{1}{2} \mathbb{E} \int_0^{t_1} \langle m^\varepsilon, 1 \rangle \, dt + \mathbb{E} (f^\varepsilon(t_1), \theta^0_{0, t_1}) \leq \frac{1}{2} \mathbb{E} \| u_0 \|_{L^2_2}^2 + Ch \mathbb{E} \| u_0 \|_{L^1_2}.
$$

(5.29)

In order to get a similar estimate on $[t_1, t_2]$ we go back to (5.26) and assume without loss of generality that $h$ was small enough so that for every $i$

$$
\sup_{t_i \leq t \leq t_{i+1}} |\xi - \theta^0_{t_i, t}| \leq 1.
$$

Consequently, by (5.27) and (5.29)

$$
\mathbb{E} (f^\varepsilon(t_1), \xi) = \mathbb{E} \langle f^\varepsilon(t_1), \theta^0_{0, t_1} \rangle + \mathbb{E} \langle f^\varepsilon(t_1), (\xi - \theta^0_{0, t_1}) \rangle \leq C (\mathbb{E} \| u_0 \|_{L^2_2}^2 + \mathbb{E} \| u_0 \|_{L^1_2}).
$$

Iterating the above technique finitely many times, the claim follows. 

As a consequence of Proposition 5.12, the assumptions of Lemma 2.5 are satisfied for $\nu^\varepsilon = \delta_{u^\varepsilon(t,x)=\xi}$ and hence there exists a Young measure $\nu_{t,x}$ vanishing at infinity such that $\nu^\varepsilon \rightharpoonup \nu$ in the sense given by this Lemma. We deduce from (5.22) that $\partial_x F = -\nu$ hence $F$ is a kinetic function. In order to show that $F$ is a generalized kinetic solution to (1.1), it remains to verify the second estimate from (2.7) and show that $m$ is a kinetic measure.

Due to the definition of $m^\varepsilon$ in (5.23), it follows from (5.25) that

$$
\frac{1}{\varepsilon} \left| \mathbb{E} \int_0^T \langle \chi_{u^\varepsilon(t)} - f^\varepsilon(t), \xi \rangle \, dt \right| + \mathbb{E} \langle f^\varepsilon(t), \xi \rangle \leq C (\mathbb{E} \| u_0 \|_{L^2_2}^2 + \mathbb{E} \| u_0 \|_{L^1_2})
$$

which implies

$$
\mathbb{E} \int_0^T \langle \chi_{u^\varepsilon(t)}, \xi \rangle \, dt \leq C (\mathbb{E} \| u_0 \|_{L^2_2}^2 + \mathbb{E} \| u_0 \|_{L^1_2}).
$$

(5.30)
Rewriting the left hand side by the same argument as in (5.28) we deduce
\[ \mathbb{E} \int_0^T \| u^\varepsilon(t) \|^2_{L^2_x} dt \leq C(\mathbb{E}\| u_0 \|^2_{L^2_x} + \mathbb{E}\| u_0 \|_{L^1_x}) \]
and as a consequence the limit Young measure \( \nu \) fulfills (2.7).

Now, we regard \( m^\varepsilon \) as a random variable with values in \( \mathcal{M}_b([0, T] \times \mathbb{R}^N \times \mathbb{R}) \) whose norm is given by the total variation of measures. In view of Proposition 5.13 we infer that the set of laws \( \{ \mathbb{P} \circ m^\varepsilon^{-1} ; \varepsilon \in (0, 1) \} \) is tight and therefore any sequence has a weakly convergent subsequence due to the Prokhorov theorem. Consequently, the law of \( m \) is supported in \( \mathcal{M}_b([0, T] \times \mathbb{R}^N \times \mathbb{R}) \). The remaining points of Definition 2.1 follow directly from the fact that a pointwise limit of measurable functions is measurable. Therefore, we have proved that \( m \) is a kinetic measure and \( F \) is a generalized kinetic solution to (1.1). Since any generalized kinetic solution is actually a kinetic one, due to Reduction Theorem 4.6, it follows that \( F = 1_{u > \xi} \) and \( \nu = \delta_u \), where \( u \in L^1(\Omega; L^\infty(0, T; L^1(\mathbb{R}^N))) \cap L^2(\Omega; L^2(0, T; L^2(\mathbb{R}^N))) \) is the unique kinetic solution to (1.1).

The weak-star convergence of \( f^\varepsilon \) to \( \chi_u \) follows immediately from (5.20) and therefore the proof is complete. \( \square \)

References

[1] C. Bauzet, G. Vallet, P. Wittbold, The Cauchy problem for conservation laws with a multiplicative noise, Journal of Hyp. Diff. Eq. 9 (4) (2012) 661-709.
[2] F. Berthelin, J. Vovelle, A BGK approximation to scalar conservation laws with discontinuous flux, Proc. of the Royal Society of Edinburgh A 140 (5) (2010) 953-972.
[3] J. Carrillo, Entropy solutions for nonlinear degenerate problems, Arch. Rational Mech. Anal. 147 (1999) 269-361.
[4] M. Caruana, P. Friz, Partial differential equations driven by rough paths, Proc. Lond. Math. Soc. (3) 100 (2010), no. 1, 177-215.
[5] C.Q. Chen, Q. Ding, K.H. Karlsen, On nonlinear stochastic balance laws, Arch. Rational Mech. Anal. 204 (2012) 707-743.
[6] G.Q. Chen, B. Perthame, Well-posedness for non-isotropic degenerate parabolic-hyperbolic equations, Ann. Inst. H. Poincaré Anal. Non Linéaire 20 (4) (2003) 645-668.
[7] D. Crisan, J. Diehl, P.K. Friz, and H. Oberhauser, Robust filtering: correlated noise and multidimensional observation, Ann. Appl. Probab. 23 (2013) no. 5, 2139-2160.
[8] A. Debussche, M. Hofmanová, J. Vovelle, Degenerate parabolic stochastic partial differential equations: Quasilinear case, to appear in Ann. Probab., arXiv:1309.5817.
[9] A. Debussche, J. Vovelle, Invariant measure of scalar first-order conservation laws with stochastic forcing, to appear in Probability Theory and Related Fields, arXiv:1310.3779.
[10] A. Debussche, J. Vovelle, Scalar conservation laws with stochastic forcing, J. Funct. Anal. 259 (2010) 1014-1042.
[11] A. Debussche, J. Vovelle, Scalar conservation laws with stochastic forcing, revised version, http://math.univ-lyon1.fr/~vovelle/DebusscheVovelleRevised.pdf.
[12] J. Diehl, H. Oberhauser, S. Riedel, A Lévy area between Brownian motion and rough paths with applications to robust nonlinear filtering and rough partial differential equations, Stoch. Pr. and their Appl. 125 (2015) 161-181.
[13] R.J. DiPerna, P.L. Lions, Ordinary differential equations, transport theory and Sobolev spaces, Invent. Math. 98 (1989) 511-547.
[14] Weinan E, K. Khanin, A. Mazel, Ya. Sinai, Invariant measures for Burgers equation with stochastic forcing, Annals of Mathematics 151 (2000) 877-960.
[15] J. Feng, D. Nualart, Stochastic scalar conservation laws, J. Funct. Anal. 255 (2) (2008) 313-373.
[16] P.K. Friz, B. Gess, Stochastic scalar conservation laws driven by rough paths, to appear in Ann. Inst. H. Poincaré Analyse Non Linéaire, arXiv:1403.6785.
[17] P. Friz, N.B. Victoir, Multidimensional Stochastic Processes as Rough Paths, Theory and Applications, Cambridge Studies in Advanced Mathematics, vol. 120, Cambridge University Press, Cambridge, 2010.
[18] B. Gess, P.E. Souganidis, Long-time behavior, invariant measures and regularizing effects for stochastic scalar conservation laws, arXiv:1411.3939.
[19] B. Gess, P.E. Souganidis, Scalar conservation laws with multiple rough fluxes, to appear in Comm. in Mathematical Sciences, arXiv:1406.2978.
[20] M. Hofmanová, A Bhatnagar-Gross-Krook approximation to stochastic scalar conservation laws, arXiv:1305.6450, to appear in Ann. Inst. H. Poincaré Probab. Statist.
