Stochastic non-isotropic degenerate parabolic–hyperbolic equations

Benjamin Gess [a,*], Panagiotis E. Souganidis [b]

[a] *Max-Planck Institute for Mathematics in the Sciences, 04103 Leipzig, Germany*
[b] *Department of Mathematics, University of Chicago, Chicago, IL 60637, USA*

Abstract

We introduce the notion of pathwise entropy solutions for a class of degenerate parabolic–hyperbolic equations with non-isotropic nonlinearity and fluxes with rough time dependence and prove their well-posedness. In the case of Brownian noise and periodic boundary conditions, we prove that the pathwise entropy solutions converge to their spatial average and provide an estimate on the rate of convergence. The third main result of the paper is a new regularization result in the spirit of averaging lemmata. This work extends both the framework of pathwise entropy solutions for stochastic scalar conservation laws introduced by Lions, Perthame and Souganidis and the analysis of the long time behavior of stochastic scalar conservation laws by the authors to a new class of equations.

* Corresponding author.
E-mail addresses: bgess@mis.mpg.de (B. Gess), souganidis@math.uchicago.edu (P.E. Souganidis).

1. Introduction

We are interested in "stochastic", scalar, degenerate parabolic–hyperbolic equations of the form

$$\begin{cases} du + \sum_{i=1}^{N} \partial_{x_i} F^i(u) \circ dz^i = \mathrm{div}(A(u)Du)dt & \text{in } \mathbb{R}^N \times (0, T), \\ u = u^0 & \text{on } \mathbb{R}^N \times \{0\}, \end{cases} \tag{1.1}$$

where

$$u_0 \in (BV \cap L^\infty)(\mathbb{R}^N), \qquad z \in C_0([0, T]; \mathbb{R}^N) \quad \text{and} \quad A \in L^\infty_{loc}(\mathbb{R}; \mathbb{R}^{N \times N}). \tag{1.2}$$

A particular example of the signal z is the N-dimensional (fractional) Brownian motion. The Stratonovich notation $\circ$ in (1.1) is justified by showing that the pathwise entropy solution constructed in this paper is the limit of solutions $u^{(l)}$ to (1.1) with z replaced by smooth approximations $z^{(l)}$ converging to z.

Equations of the type (1.1) arise, for example, in the study of mean field games, see [24,31] for details. In this case one typically has $A(u) = \mathrm{Id}$. More generally, (1.1) describes the evolution of a density subject to nonlinear diffusion and convected by a nonlinear, stochastic flux function. For example, this models the diffusion of the temperature in an impure, turbulent fluid, loosely related to the so-called Kraichnan models of turbulent mixing of passive scalars. For simplicity, a possible additional semilinear noise term is neglected in (1.1), see [22,27] for progress in this direction.

We assume that

$$A \in L^\infty_{loc}(\mathbb{R}; \mathcal{S}^N) \text{ is nonnegative}, \tag{1.3}$$

where $\mathcal{S}^N$ is the space of symmetric $N \times N$ matrices, and

$$F \in C^2_{loc}(\mathbb{R}; \mathbb{R}^N) \quad \text{with } F'(0) = 0. \tag{1.4}$$

In this paper we introduce the notion of pathwise entropy solutions for (1.1) and prove that it is well-posed. In the setting of stochastic scalar conservation laws, that is when $A \equiv 0$ in (1.1), the notion of pathwise entropy solutions was introduced by Lions, Perthame and Souganidis [31], while Chen and Perthame [10] studied the entropy solutions of the "deterministic" version of (1.1), that is for $z(t) = (t, \ldots, t)$. We also refer to Perthame and Souganidis [42] who introduced an equivalent solution in the deterministic case called dissipative solution and to Lions, Perthame and Souganidis [33], Gess and Souganidis [24], Friz and Gess [20], Hofmanova [27] for stochastic scalar conservation laws with $A \equiv 0$ allowing for spatially inhomogeneous flux and semilinear noise. We note that, as compared to [24,27,31,33], the inclusion of the non-isotropic parabolic part A in (1.1) leads to additional complications and additional approximation and localization steps in the proof of well-posedness are needed.

As particular examples, (1.1) includes stochastic porous media equations, that is

$$du + \sum_{i=1}^{N} \partial_{x_i} F_i(u) \circ dz^i = \Delta u^{[m]}dt \quad \text{in } \mathbb{R}^N \times (0, T) \tag{1.5}$$

where $m > 1$ and $u^{[m]} := |u|^{m-1}u$.

In this case our results extend previous work on stochastic porous media equations, based on the variational approach to SPDE, which did not allow treatment of nonlinear transport noise as

in (1.5). See, for example, Prévot and Röckner [43], Krylov and Rozovskiĭ [29], Pardoux [40], Barbu, Da Prato and Röckner [1–3] and the references therein. In addition, in contrast to previous work, the pathwise approach to (1.1) developed in this paper immediately implies the existence of a corresponding random dynamical system. This is of particular interest, since (1.1) contains nonlinear multiplicative noise for which the existence of a random dynamical system is known to be a difficult problem (see, for example, Mohammed, Zhang and Zhao [39], Flandoli [19] and Gess [21]).

In the second part of the paper we study the long-time behavior and regularity properties of (1.1) set on the torus $\mathbb{T}^N$ and driven by Brownian motion, that is the initial value problem

$$\begin{cases} du + \sum_{i=1}^{N} \partial_{x_i} F^i(u) \circ d\beta_t^i = \operatorname{div}(A(u)Du)dt & \text{in } \mathbb{T}^N \times (0, \infty), \\ u = u_0 & \text{on } \mathbb{T}^N \times \{0\}, \end{cases} \tag{1.6}$$

with

$$\begin{aligned} & A \in C(\mathbb{R}; \mathcal{S}^N) \cap C^1(\mathbb{R} \setminus \{0\}; \mathcal{S}^N), \\ & |F''(\xi)| + |A'(\xi)| \le C(1 + |\xi|^{p_1} + |\xi|^{p_2}) \end{aligned} \tag{1.7}$$

for some $C > 0$, $p_1, p_2 \in (-1, \infty)$ and all $\xi \in \mathbb{R} \setminus \{0\}$,

$$\beta = (\beta^1, \ldots, \beta^N) \text{ is a standard two-sided Brownian motion,} \tag{1.8}$$

and

$$u_0 \in L^1(\mathbb{T}^N). \tag{1.9}$$

We prove that, under a genuine nonlinearity condition for F, A, the solutions to (1.6) converge to their spatial average and we provide a rate of convergence. In the deterministic setting, that is when $z(t) = (t, \ldots, t)$, Chen and Perthame [11] proved the convergence to the spatial average by different methods and without an estimate on the rate of convergence. In the hyperbolic case, that is when $A \equiv 0$ and restricted to one and two space dimensions respectively, Lax [30], E and Engquist [17] provided estimates on the rate of convergence. Convergence to the spatial average under weak conditions on the flux F has been shown by Dafermos [12]. A rate of convergence in any dimension for both deterministic and stochastic scalar conservation laws ($A \equiv 0$) was established by the authors in [25].

The third result presented here is a regularity estimate based on averaging techniques. This extends the regularity results developed in [25] for the hyperbolic case and provides new regularity estimates for stochastic porous media equations. Averaging lemmata for deterministic parabolic–hyperbolic equations have been established by Tadmor and Tao [44].

Stochastic scalar conservation laws driven by multiplicative semilinear noise, that is,

$$du + \operatorname{div} F(u)dt = g(u)dW_t \tag{1.10}$$

have been intensively studied in recent years; we refer to Holden and Risebro [28], Feng and Nualart [18], Debussche and Vovelle [14], Chen, Ding and Karlsen [7], Bauzet, Vallet and Wittbold [4], Hofmanova [26] and the references therein. In particular, Debussche, Hofmanova and Vovelle [13] and Bauzet, Vallet and Wittbold [5] studied SPDE of the type (1.10) additionally including a second order quasilinear term. Part of the results obtained in this paper is related to questions of continuous dependence of solutions to quasilinear degenerate PDE with respect to

the coefficients. For related results we refer to [8,9]. The long-time behavior of solutions to (1.10) with additive noise has been analyzed by E, Khanin, Mazel and Sinai [16] and Debussche and Vovelle [15].

Organization of the paper. In Section 2, we introduce the notion of pathwise entropy solutions and state the results about its well-posedness for general continuous paths and the long time behavior with rates as well as regularity in the stochastic case. The proofs are given in Sections 3, 4, 5, 6 respectively. Since some of the results are technical, we have chosen to present in the first two appendices the proofs of the most technical estimates. Finally, for the convenience of the reader we recall in Appendix D the definition [10, Definition 2.2] of the kinetic solution to (1.1) for smooth driving signals.

Notation. We work in $\mathbb{R}^N$ or on the torus $\mathbb{T}^N$, S^{N-1} is the unit sphere in $\mathbb{R}^N$, $\mathbb{R}_+ := (0, \infty)$ and δ is the "Dirac" mass at the origin in $\mathbb{R}$. For notational convenience, we write $A \lesssim B$ if $A \leq CB$ for some $C > 0$, and $A \sim B$, if $A \lesssim B$ and $B \lesssim A$. For a matrix $A \in \mathbb{R}^{N \times N}$ we write $A = (a_{ij})_{i,j=1}^N$ and, given two matrices A, B, we set $A : B := \sum_{i,j=1}^N a_{ij} b_{ij}$. The subspace of L^1-functions with bounded total variation is BV and its norm is $\| \cdot \|_{BV}$. If $f \in BV$, $BV(f)$ is its total variation. We write δ_c for the "Dirac mass" measure in $L^1(\mathbb{T}^N)$ charging the constant function $c \in \mathbb{R}$ and we set $L_c^1(\mathbb{T}^N) := \{u \in L^1(\mathbb{T}^N) : u \text{ has average } c\}$. We set $C_0([0, T]; \mathbb{R}^N) := \{z \in C([0, T]; \mathbb{R}^N) : z(0) = 0\}$. For an open set $\mathcal{O} \subseteq \mathbb{R}^N$, $C_c^1(\mathcal{O})$ is the space of all continuously differentiable functions with compact support in $\mathcal{O}$. If $F : [0, T] \to \mathbb{R}$, $F|_s^t := F(t) - F(s)$ for all $s, t \in [0, T]$. In the second part of the paper, we will omit, when it does not cause any confusion, the dependence in ω and we occasionally write $m(x, \xi, s)dxd\xi ds$ instead of $dm(x, \xi, s)$. The space of homogeneous Bessel potentials, $W^{\lambda, p}$ for $\lambda > 0$, $p \in [1, \infty)$, is

$$W^{\lambda, p} := \{f \in L^p(\mathbb{T}^N) : (|n|^\lambda \hat{f}(n))^\vee \in L^p(\mathbb{T}^N)\},$$

where $\hat{f}$ is the discrete Fourier transform of f on $\mathbb{T}^N$ and $f^\vee$ is its inverse. The homogeneous Bessel potential spaces coincide with the domains of the fractional Laplace operators $(-\Delta)^{\frac{\lambda}{2}}$ on $L^p(\mathbb{T}^N)$. For notational simplicity we set $H^\lambda := W^{\lambda, 2}$.

2. The main results

Pathwise entropy solutions. We introduce here the notion of pathwise entropy solutions to (1.1), which is based on the kinetic formulation of (1.1) and on choosing test-functions transported along the characteristics. For (hyperbolic) scalar conservation laws this was introduced in [31] and it was motivated from the theory of stochastic viscosity solutions for fully nonlinear first- and second-order PDE including stochastic Hamilton–Jacobi equations developed by Lions and Souganidis (see [34–38]). The theory was extended to the spatially inhomogeneous case by Lions, Perthame and Souganidis [33], Gess and Souganidis [25] and Hofmanova [27]. In view of the parabolic part in (1.1), the kinetic formulation has to be appropriately adapted; in the deterministic case this was done in [10].

For smooth driving paths $z \in C^1([0, T]; \mathbb{R}^N)$ the well-established theories of entropy solutions and kinetic solutions apply to (1.1). In order to fix the ideas and the notation, it is necessary to introduce the functions

$$\chi(u, \xi) := \begin{cases} +1 & \text{if } 0 \leq \xi \leq u, \\ -1 & \text{if } u \leq \xi \leq 0, \\ 0 & \text{otherwise}, \end{cases} \tag{2.1}$$

and, given $u : \mathbb{R}^N \times [0, \infty) \to \mathbb{R}$,

$$\chi(x, \xi, t) := \chi(u(x, t), \xi) \tag{2.2}$$

and to observe that

$$\partial_\xi \chi(x, \xi) = \delta(\xi) - \delta(\xi - u).$$

Finally, to simplify the notation, we use the functions

$$f_i(\xi, t) := (F')^i(\xi)\dot{z}^i(t).$$

The kinetic form of (1.1) then is

$$\begin{cases} \partial_t \chi + f(\xi, t) \cdot D_x \chi - \displaystyle\sum_{i,j=1}^d a_{ij}(\xi)\partial^2_{x_i x_j} \chi = \partial_\xi(m + n) & \text{in } \mathbb{R}^N \times \mathbb{R} \times [0, \infty), \\ \chi(x, \xi, 0) = \chi(u_0(x), \xi), \end{cases} \tag{2.3}$$

where the entropy defect measure m and the parabolic dissipation measure n are nonnegative, bounded measures in $\mathbb{R}^N \times \mathbb{R} \times [0, T]$ for each $T > 0$. In the sequel, when we do not need to differentiate between m and n, we set $q := m + n$.

The notion of kinetic solutions is not well defined for rough driving signals z that are merely continuous, since the coefficients $f(\xi, t)$ blow up with $\dot{z}$. On the other hand, following [25,31, 33] we observe that the linearity of (2.3) in χ suggests that we may use the characteristics of (2.3) to derive a stable notion of solutions, which will be the pathwise entropy solutions.

We continue assuming $z \in C^1([0, T]; \mathbb{R}^N)$ and consider the transport equation

$$\begin{cases} \partial_t \varrho + f(\xi, t) \cdot D_x \varrho = 0 & \text{in } \mathbb{R}^N \times \mathbb{R} \times [0, T], \\ \varrho_0 = \varrho^0 & \text{on } \mathbb{R}^N \times \mathbb{R} \times \{0\}. \end{cases} \tag{2.4}$$

For each $(y, \eta) \in \mathbb{R}^{N+1}$ and $\varrho^0 \in C_c^\infty(\mathbb{R}^N \times \mathbb{R})$, $\varrho = \varrho(x, y, \xi, \eta, t)$ is the solution to (2.4) with

$$\varrho(\cdot, y, \cdot, \eta, 0) = \varrho^0(\cdot - y, \cdot - \eta) = \varrho^{s,0}(\cdot - y)\varrho^{v,0}(\cdot - \eta) \tag{2.5}$$

and $\varrho^{s,0} \in C_c^\infty(\mathbb{R}^N)$, $\varrho^{v,0} \in C_c^\infty(\mathbb{R})$. Here, we use the superscripts s and v to emphasize that the functions act on the space and velocity variables respectively.

The convolution along characteristics is then given by

$$\varrho * \chi(y, \eta, t) := \int \varrho(x, y, \xi, \eta, t)\chi(x, \xi, t)dxd\xi,$$

and it follows from Lemma 2.2 that, for all $(y, \eta) \in \mathbb{R}^{N+1}$ and in the sense of distributions in $t \in [0, T]$,

$$\partial_t \varrho * \chi(y, \eta, t) = \sum_{i,j=1}^N \int a_{ij}(\xi)\chi(x, \xi, t)\partial^2_{x_i x_j}\varrho(x, y, \xi, \eta, t)dxd\xi$$

$$- \int \partial_\xi \varrho(x, y, \xi, \eta, t)q(x, \xi, t)dxd\xi \quad \text{in } \mathbb{R}^N \times \mathbb{R} \times (0, T).$$

The solution of (2.4) can be expressed in terms of the associated (backward) characteristics. Due to the spatial homogeneity of the flux f, the characteristics starting at (y, ξ) and the solution

$\varrho = \varrho(x, y, \xi, \eta, t)$ to (2.4) are given respectively by the explicit form

$$Y^i_{(y,\eta)}(t) = y + f^i(\xi)z^i(t),$$

and

$$\varrho(x, y, \xi, \eta, t) = \varrho^s(x, y, \xi, t)\varrho^{v,0}(\xi - \eta) := \varrho^{s,0}(x - f(\xi)z(t) - y)\varrho^{v,0}(\xi - \eta). \quad (2.6)$$

Note that, in contrast to (2.4), the characteristic equations are trivially well defined also for rough driving signals $z \in C_0([0, T]; \mathbb{R}^N)$.

Since $A(u)$ is a symmetric, nonnegative matrix, it has a square root, that is there exists $\sigma(u) = (\sigma_{ij}(u))^N_{i,j=1} := A(u)^{\frac{1}{2}} \in L^\infty_{loc}(\mathbb{R}; \mathbb{R}^{N \times N})$, such that

$$a_{ij}(u) = \sum_{k=1}^{N} \sigma_{ik}(u)\sigma_{jk}(u).$$

For notational convenience we write

$$\beta'_{ik}(u) = \sigma_{ik}(u).$$

Next we give the definition of pathwise entropy solutions.

Definition 2.1. Let $u_0 \in (L^1 \cap L^\infty)(\mathbb{R}^N)$ and $T > 0$. A function $u \in C([0, T]; L^1(\mathbb{R}^N)) \cap L^\infty(\mathbb{R}^N \times [0, T])$ is a pathwise entropy solution to (1.1), if

(i) for all $k \in \{1, \ldots, N\}$

$$\sum_{i=1}^{N} \partial_{x_i}\beta_{ik}(u) \in L^2(\mathbb{R}^N \times [0, T]), \quad (2.7)$$

(ii) for all $(y, \eta) \in \mathbb{R}^{N+1}$ and all $k \in \{1, \ldots, N\}$

$$\sum_{i=1}^{N} \int_{\mathbb{R}^{N+1}} \chi(x, \xi, t)\sigma_{ik}(\xi)\partial_{x_i}\varrho(x, y, \xi, \eta, t)dxd\xi$$

$$= -\sum_{i=1}^{N} \int_{\mathbb{R}^N} \partial_{x_i}\beta_{ik}(u(x, t))\varrho(x, y, u(x, t), \eta, t)dx, \quad (2.8)$$

(iii) there exists a non-negative bounded measure m on $\mathbb{R}^N \times \mathbb{R} \times [0, T]$ such that, for all test functions ϱ given by (2.4) with $\varrho^0 \in C^\infty_c(\mathbb{R}^{N+1})$ as in (2.5), for all $(y, \eta) \in \mathbb{R}^{N+1}$ and in the sense of distributions in $t \in [0, T]$,

$$\frac{d}{dt}(\varrho * \chi)(y, \eta, t) = \sum_{i,j=1}^{N} \int \chi(x, \xi, t)a_{ij}(\xi)\partial^2_{x_i x_j}\varrho(x, y, \xi, \eta, t)dxd\xi$$

$$- \int \partial_\xi\varrho(x, y, \xi, \eta, t)(m + n)(x, \xi, t)dxd\xi, \quad (2.9)$$

where n is the non-negative measure on $\mathbb{R}^N \times \mathbb{R} \times [0, T]$ given by

$$n(x, \xi, t) := \delta(\xi - u(t, x))\sum_{k=1}^{N}\left(\sum_{i=1}^{N}\partial_{x_i}\beta_{ik}(u(t, x))\right)^2. \quad (2.10)$$

In the sequel, we refer to (2.7) and (2.8) as the regularity and chain rule properties respectively.

When $A \equiv 0$, that is for the hyperbolic problem, (2.9) alone was used as the definition of a pathwise entropy solution in [31]. In the presence of A it is, however, necessary to also include the chain-rule (2.8) in the definition, a fact already observed in the deterministic case in [10].

The definition can be reformulated as follows: If u_0, m and n are as in Definition 2.1, then $u \in C([0, T]; L^1(\mathbb{R}^N)) \cap L^\infty([0, T] \times \mathbb{R}^N)$ is a pathwise entropy solution to (1.1) if and only if, instead of (2.9), for all $(y, \eta) \in \mathbb{R}^{N+1}$, $s \leq t$ and all test functions ϱ given by (2.4) with $\varrho^0 \in C_c^\infty(\mathbb{R}^{N+1})$,

$$
\begin{cases}
\varrho * \chi(y, \eta, t) - \varrho * \chi(y, \eta, s) = \displaystyle\sum_{i,j=1}^{N} \int_s^t \int \chi(x, \xi, r) a_{ij}(\xi) \partial^2_{x_i x_j} \varrho(x, y, \xi, \eta, s) d\xi dx dr \\
\qquad\qquad - \displaystyle\int_s^t \int \partial_\xi \varrho(x, y, \xi, \eta, r)(m + n)(x, \xi, r) dx d\xi dr.
\end{cases}
\tag{2.11}
$$

Next we show that, for smooth paths, the notions of kinetic and pathwise entropy solutions are equivalent. For the convenience of the reader we recall from [10] the definition of the kinetic solution in Appendix D.

Lemma 2.2. *Assume that $z \in C_0^1([0, T]; \mathbb{R}^N)$ and $u \in C([0, T]; L^1(\mathbb{R}^N)) \cap L^\infty([0, T] \times \mathbb{R}^N)$. Then u is a pathwise entropy solution to (1.1) if and only if it is a kinetic solution.*

Proof. Let u be a kinetic solution to (1.1). Since $u \in L^\infty([0, T] \times \mathbb{R}^N)$, we may take $\psi \equiv 1$ in (D.1) and $\psi_2 \equiv 1$ in (D.2).

Then, for all $\varphi \in C_c^1(\mathbb{R}^N \times \mathbb{R} \times (0, T))$,

$$
\int_0^T \int \chi\left(\partial_t \varphi + f(\xi, r) \cdot D_x \varphi + \sum_{i,j=1}^N a_{ij}(\xi) \partial^2_{x_i x_j} \varphi\right) d\xi dx dr = \int_0^T \int q \partial_\xi \varphi d\xi dx dr.
$$

Since $u \in C([0, T]; L^1(\mathbb{R}^N))$ it follows that, for all $s \leq t$ and all $\varphi \in C_c^1(\mathbb{R}^N \times \mathbb{R} \times [0, T])$,

$$
\int_s^t \int \chi(x, \xi, r)\left(\partial_t \varphi(x, \xi, r) + f(\xi, r) \cdot D_x \varphi(x, \xi, r) + \sum_{i,j=1}^N a_{ij}(\xi) \partial^2_{x_i x_j} \varphi(x, \xi, r)\right) d\xi dx dr
$$

$$
- \int \chi(x, \xi, t)\varphi(x, \xi, t) d\xi dx + \int \chi(x, \xi, s)\varphi(x, \xi, s) d\xi dx = \int_s^t \int q \partial_\xi \varphi d\xi dx dr.
$$

Choosing $\varphi = \varrho$ yields

$$
- \int \chi(x, \xi, \cdot)\varrho(x, y, \xi, \eta, \cdot) d\xi dx \Big|_s^t
$$

$$
+ \sum_{i,j=1}^N \int_s^t \int \chi(x, \xi, r) a_{ij}(\xi) \partial^2_{x_i x_j} \varrho(x, y, \xi, \eta, s) d\xi dx dr
$$

$$
= \int_s^t \int q \partial_\xi \varrho d\xi dx dr,
$$

which in view of (2.11) implies that χ is pathwise entropy solution to (1.1).

Next we derive the chain-rule. It follows from (D.2) that, for all $\psi \in C_c^\infty$,

$$
\sqrt{\psi(u(x, t))} \sum_{i=1}^N \partial_{x_i} \beta_{ik}(u(x, t)) = \sum_{i=1}^N \partial_{x_i} \beta_{ik}^\psi(u(x, t)) \quad \text{a.e. in } (t, x)
\tag{2.12}
$$

where

$$(\beta_{ik}^{\psi})'(u) = \sigma_{ik}(u)\sqrt{\psi(u)}, \qquad (2.13)$$

and thus

$$\beta_{ik}^{\psi}(u(x,t)) = \int \chi(x,\xi,t)\sigma_{ik}(\xi)\sqrt{\psi(\xi)}d\xi.$$

Choose $\sqrt{\psi(\cdot)} := \varrho^{v,0}(\cdot - \eta)\varrho_{\delta}(\cdot - \xi)$ in (2.12), where ϱ_{δ} is a standard Dirac sequence, multiply the resulting equation by $\varrho^{s}(x,y,\xi,t) = \varrho^{s,0}(x - y + f(\xi)z(t))$ and integrate in x and ξ to get

$$\sum_{i=1}^{N}\int \varrho^{v}(u(x,t) - \eta)\varrho_{\delta}(u(x,t) - \xi)\partial_{x_i}\left(\int \sigma_{ik}(\tau)\chi(x,\tau,t)d\tau\right)\varrho^{s}(x,y,\xi,t)dxd\xi$$

$$= \sum_{i=1}^{N}\int \partial_{x_i}\left(\int \sigma_{ik}(\tau)\varrho^{v}(\tau - \eta)\varrho_{\delta}(\tau - \xi)\chi(x,\tau,t)d\tau\right)\varrho^{s}(x,y,\xi,t)dxd\xi$$

$$= -\sum_{i=1}^{N}\int \left(\int \sigma_{ik}(\tau)\varrho^{v}(\tau - \eta)\varrho_{\delta}(\tau - \xi)\chi(x,\tau,t)d\tau\right)\partial_{x_i}\varrho^{s}(x,y,\xi,t)dxd\xi.$$

Letting $\delta \to 0$ we obtain

$$\sum_{i=1}^{N}\int \varrho^{v}(u(x,t) - \eta)\partial_{x_i}\left(\int \sigma_{ik}(\tau)\chi(x,\tau,t)d\tau\right)\varrho^{s}(x,y,u(x,t),t)dx$$

$$= -\sum_{i=1}^{N}\int \sigma_{ik}(\xi)\varrho^{v}(\xi - \eta)\chi(x,\xi,t)\partial_{x_i}\varrho^{s}(x,y,\xi,t)dxd\xi,$$

and, thus,

$$\sum_{i=1}^{N}\int \partial_{x_i}\beta_{ik}(u(x,t))\varrho(x,y,u(x,t),\eta,t)dx$$

$$= -\sum_{i=1}^{N}\iint \sigma_{ik}(\xi)\chi(x,\xi,t)\partial_{x_i}\varrho(x,y,\xi,\eta,t)dxd\xi.$$

Let now u be a pathwise entropy solution to (1.1). It follows from [10] that there exists a kinetic solution $\tilde{u}$ to (1.1), which, in view of the above, is also a pathwise entropy solution. Since, by Theorem 2.3, pathwise entropy solutions are unique, it follows that $u = \tilde{u}$ and hence u is a kinetic solution. □

The well-posedness of pathwise entropy solutions. The first result here is the uniqueness and stability of pathwise entropy solutions with respect to the initial condition and driving paths.

Theorem 2.3. *Assume* (1.3) *and* (1.4) *and let* $u^{(1)}, u^{(2)} \in L^{\infty}([0, T]; BV(\mathbb{R}^{N}))$ *be two pathwise entropy solutions to* (1.1) *with driving signals* $z^{(1)}, z^{(2)} \in C_{0}([0, T]; \mathbb{R}^{N})$ *and initial values* $u_{0}^{1}, u_{0}^{2} \in (L^{\infty} \cap BV)(\mathbb{R}^{N})$. *Then, for all* $s \leq t \in [0, T]$, *there exists* $C > 0$, *which may depend on* $\|u^{(1)}\|_{L^{\infty}([s,t];(L^{\infty} \cap BV)(\mathbb{R}^{N}))}$ *and* $\|u^{(2)}\|_{L^{\infty}([s,t];(L^{\infty} \cap BV)(\mathbb{R}^{N}))}$, *such that*

$$\begin{cases} \|u^{(1)}(t) - u^{(2)}(t)\|_{L^{1}(\mathbb{R}^{N})} \leq \|u^{(1)}(s) - u^{(2)}(s)\|_{L^{1}(\mathbb{R}^{N})} + C\|z^{(1)} - z^{(2)}\|_{C([s,t];\mathbb{R}^{N})}^{1/2} \\ \qquad\qquad + C\|z^{(1)} - z^{(2)}\|_{C([s,t];\mathbb{R}^{N})}. \end{cases} \quad (2.14)$$

The next result is about the existence of pathwise entropy solutions, which, in view of Theorem 2.3, are unique.

Theorem 2.4. (i) *Assume* (1.3) *and* (1.4) *and let* $u_0 \in (L^\infty \cap BV)(\mathbb{R}^N)$ *and* $z \in C_0([0, T]; \mathbb{R}^N)$. *Then there exists a pathwise entropy solution* u *to* (1.1) *satisfying, for all* $p \in [1, \infty]$,

$$\|u\|_{L^\infty([0,T];L^p(\mathbb{R}^N))} \leq \|u_0\|_p, \tag{2.15}$$

$$\|u\|_{L^\infty([0,T];BV(\mathbb{R}^N))} \leq \|u_0\|_{BV}, \tag{2.16}$$

and

$$\int_{\mathbb{R}^N \times \mathbb{R} \times \mathbb{R}_+} m(x, \xi, t) dx d\xi dt + \int_{\mathbb{R}^N \times \mathbb{R}_+} \sum_{k=1}^{N} \left(\sum_{i=1}^{N} \partial_{x_i} \beta_{ik}(u) \right)^2 dx dt \leq \frac{1}{2} \|u_0\|_2^2. \tag{2.17}$$

Moreover, for $p \in (-1, \infty)$ *and* $t \geq 0$,

$$\|u(\cdot, t)\|_{p+2}^{p+2} + (p+2)(p+1) \int_0^t \int_{\mathbb{R}^N \times \mathbb{R}} |\xi|^p d(m+n)(x, \xi, s) = \|u_0\|_{p+2}^{p+2}. \tag{2.18}$$

(ii) *Assume* (1.3) *and* (1.4) *and let* $u_0 \in (L^\infty \cap L^1)(\mathbb{R}^N)$ *and* $z \in C_0([0, T]; \mathbb{R}^N)$. *Then there exists a pathwise entropy solution* u *to* (1.1) *satisfying* (2.15) *for all* $p \in [1, \infty]$, (2.17) *and* (2.18) *for* $p \in (-1, \infty)$ *and* $t \geq 0$.

It follows easily from the arguments leading to Theorem 2.4 that the well-posedness theory for (1.1) extends to (1.1) set on the torus $\mathbb{T}^N$.

Proposition 2.5. *Assume* (1.3) *and* (1.4) *and let* $u_0 \in (L^\infty \cap BV)(\mathbb{T}^N)$ *and* $z \in C_0([0, T]; \mathbb{R}^N)$. *Then, there exists a unique pathwise entropy solution* $u \in C([0, \infty); L^1(\mathbb{T}^N)) \cap L^\infty([0, T]; (L^\infty \cap BV)(\mathbb{T}^N))$, *for all* $T > 0$, *and the solution operator is an* L^1-*contraction and, hence, is defined on* $L^1(\mathbb{T}^N)$.

The long time behavior and regularity. Here we assume that z is a Brownian motion. We start by considering the long time behavior of (1.6). The claim is that, under a genuine nonlinearity assumption for F and A (see (2.20)), as $t \to \infty$ and almost surely (a.s. for short) in ω,

$$u(\cdot, t; \omega, u_0) \to \bar{u}_0 := \int_{\mathbb{T}^N} u_0(x) dx.$$

Moreover, we provide an estimate for the rate of convergence, we show that $\delta_{\bar{u}_0}$ is the unique, strongly mixing invariant measure of the associated Markov semigroup and that $\bar{u}_0$ is the random attractor of the associated random dynamical system (see Theorem 2.6).

Without loss of generality we consider the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in \mathbb{R}_+}, \mathbb{P})$ with the canonical realization of the two-sided Brownian motion on $\Omega = C_0(\mathbb{R}; \mathbb{R}^N) := \{b \in C(\mathbb{R}; \mathbb{R}^N)$ and $b(0) = 0\}$, and $\mathbb{P}, \mathbb{E}, \mathcal{F}_t$ and $\bar{\mathcal{F}}_t$ respectively the two-sided standard Gaussian measure on Ω, the expectation with respect to $\mathbb{P}$, the canonical, uncompleted filtration and its completion.

A measurable map $m : \Omega \to \mathcal{M}$, where $\mathcal{M}$ is the space of nonnegative bounded measures on $\mathbb{T}^N \times \mathbb{R} \times [0, T]$, is a kinetic measure, if the process $t \mapsto m(., [0, t])$ with values in the space of nonnegative bounded measures on $\mathbb{T}^N \times \mathbb{R}$ is $\mathcal{F}_t$-adapted.

A map $u : \mathbb{T}^N \times [0, \infty) \times \Omega \to \mathbb{R}$ is called a (stochastic) pathwise entropy solution to (1.6) if, for all $\omega \in \Omega$, $u(\cdot, \omega)$ is a pathwise entropy solution to (1.6), n, m are kinetic measures and $t \mapsto u(\cdot, t)$ is an $\mathcal{F}_t$-adapted process in $L^1(\mathbb{T}^N)$.

Since the entropy solution to (1.6) is constructed in a pathwise manner, for each $u_0 \in L^1(\mathbb{T}^N)$ and $t \geq 0$, the map

$$\varphi(t, \omega)u_0 := u(\cdot, t; \omega, u_0) \tag{2.19}$$

defines a continuous random dynamical system (RDS) on $L^1(\mathbb{T}^N)$. For some background on RDS we refer to [25, Appendix A].

The associated Markovian semigroup $(P_t)_{t \geq 0}$ is given, for each bounded measurable function f on $L^1(\mathbb{T}^N)$, $u_0 \in L^1(\mathbb{T}^N)$ and $t \geq 0$, by

$$P_t f(u_0) := \mathbb{E} f(u(\cdot, t; \cdot, u_0)) = \mathbb{E} f(\varphi(t, \cdot)u_0).$$

As usual, by duality we may consider the action of $(P_t)_{t \geq 0}$ on the space $\mathcal{M}_1$ of probability measures on $L^1(\mathbb{T}^N)$, that is, for each $\mu \in \mathcal{M}_1$, we define

$$P_t^* \mu(f) := \int_{L^1} P_t f(x) d\mu(x).$$

A probability measure μ is an invariant measure for $(P_t)_{t \geq 0}$ if, for all $t \geq 0$,

$$P_t^* \mu = \mu,$$

and μ is said to be strongly mixing if, for each $\nu \in \mathcal{M}_1$ and as $t \to \infty$,

$$P_t^* \nu \rightharpoonup \mu \text{ weak } \star \text{ in } \mathcal{M}_1.$$

The study of the long time behavior of the stochastic entropy solutions is based on a new stochastic averaging-type lemma (Theorem 2.7) which holds true under the following genuine nonlinearity condition:

$$\begin{cases} \text{there exist } \theta \in (0, 1] \text{ and } C > 0 \text{ such that, for all } \sigma \in S^{N-1}, z \in \mathbb{R}^N \text{ and } \varepsilon > 0, \\ |\{\xi \in \mathbb{R} : |f(\xi)\sigma - z|^2 + \sigma A(\xi)\sigma \leq \varepsilon\}| \leq C\varepsilon^{\frac{\theta}{2}}, \end{cases} \tag{2.20}$$

where the product of the two vectors $f(\xi)\sigma$ is defined by $(f(\xi)\sigma)^i := f^i(\xi)\sigma^i$ and $\sigma A(\xi)\sigma := \sum_{i,j} a_{ij}(\xi)\sigma_i\sigma_j$. We choose the scaling $\frac{\theta}{2}$ on the right hand side in order to be consistent with our previous work [25].

Also note that in the corresponding deterministic case, that is (1.6) with $z(t) = (t, \ldots, t)$, the expectation is that the non-degeneracy condition should be (see Tadmor and Tao [44])

$$|\{\xi \in \mathbb{R} : |f(\xi) \cdot \sigma - z| + \sigma A(\xi)\sigma \leq \varepsilon\}| \lesssim \varepsilon^{\frac{\theta}{2}}. \tag{2.21}$$

In (2.21) the hyperbolic and parabolic parts $|f(\xi) \cdot \sigma - z|$ and $\sigma A(\xi)\sigma$ scale respectively linearly and quadratically in σ. In contrast, in the stochastic case both the hyperbolic and parabolic parts scale quadratically in σ. This change in the scaling of the hyperbolic part in (2.20) is due to the quadratic scaling of Brownian motion. The fact that both the hyperbolic and the parabolic part of (1.6) are scaled quadratically is crucial for the proof of the averaging lemma and allows a significantly simpler treatment than in [44].

For notational simplicity we set

$$p_0 := 0 \vee p_1 \vee p_2.$$

Theorem 2.6. *Assume* (1.7)–(1.9) *and* (2.20). *Then, as* $t \to \infty$

$$u(\cdot, t; \cdot, u_0) \to \bar{u}_0 \quad \text{in } L^1(\Omega; L^1(\mathbb{T}^N)) \text{ and } \mathbb{P}\text{-a.s. in } L^1(\mathbb{T}^N). \tag{2.22}$$

Moreover, for $t \geq 1$ *and* $u_0 \in L^{2+p_0}(\mathbb{T}^N)$,

$$\mathbb{E}\|u(\cdot, t; \cdot, u_0) - \bar{u}_0\|_1 \leq t^{-\frac{\theta}{4+\theta}} \left(\|u_0\|_{2+p_0}^{2+p_0} + 1 \right).$$

In particular, $\delta_{\bar{u}_0}$ *is the unique invariant measure for* $(P_t)_{t \geq 0}$ *on* $L^1_{\bar{u}_0}(\mathbb{T}^N)$ *and is strongly mixing. Restricted to* $L^1_{\bar{u}_0}(\mathbb{T}^N)$ *the random dynamical system* φ *has* $\bar{u}_0$ *as a forward and pullback random attractor.*

Note that, if $u_0 \in L^\infty(\mathbb{T}^N)$, then the convergence in (2.22) holds in $L^p(\Omega; L^p(\mathbb{T}^N))$ and $\mathbb{P}$-a.s. in $L^p(\mathbb{T}^N)$ for all $p \in [1, \infty)$, since $\|u(\cdot, t; \cdot, u_0)\|_\infty \leq \|u_0\|_\infty$ (see Theorem 2.4).

Using essentially the same estimates leading to Theorem 2.6 we also prove a new regularity result, which extends the estimates obtained in [25,32] for the hyperbolic case, that is, when $A \equiv 0$ and in [44] for the deterministic problem.

Theorem 2.7. *Assume* (1.7), (1.8), (1.9), (2.20) *and let* $u_0 \in L^{2+p_0}(\mathbb{T}^N)$. *Then, for all* $\lambda \in (0, \frac{2\theta}{\theta+4})$, $\delta > 0$ *and* $T > 0$,

$$\mathbb{E} \int_0^T \|u(t)\|_{W^{\lambda,1}} dt \lesssim (1 + \|u_0\|_{2+p_0}^{2+p_0}) \quad \text{and} \quad \sup_{t \geq \delta} \mathbb{E}\|u(t)\|_{W^{\lambda,1}} < \infty.$$

3. The stability of pathwise entropy solutions

We present the proof of Theorem 2.3. Since it is rather long, we have divided it into several parts and we have chosen to present some technical arguments in Appendix A. To help the reader, in the first part of this section we present an informal overview of the arguments and the main ideas and then we continue with the actual proof.

Informal arguments and main ideas. We present an informal overview of the proof of uniqueness of pathwise entropy solutions. While several of the arguments presented next need further justification, all the main steps of the rigorous proof are included. In order to make everything rigorous, it is necessary to introduce appropriate smoothing/mollifications and doubling of variables. All these create significant technical difficulties since it is necessary to estimate the errors, etc.

***Step* 1**: *Transformation*. The definition of a pathwise entropy solution corresponds to the transformation

$$\tilde{\chi}(x, \xi, t) := \chi(x + f(\xi)z_t, \xi, t),$$

with $\tilde{\chi}$ solving, informally,

$$\begin{aligned} \partial_t \tilde{\chi}(x, \xi, t) &= \partial_t(\chi(x + f(\xi)z_t, \xi, t)) \\ &= D_x\chi(x + f(\xi)z_t, \xi, t) \cdot f(\xi)\dot{z}_t + \partial_t\chi(x + f(\xi)z_t, \xi, t) \\ &= \sum_{i,j} a_{ij}(\xi)\partial_{ij}^2 \tilde{\chi}(x, \xi, t) + (\partial_\xi q)(x + f(\xi)z_t, \xi, t). \end{aligned} \tag{3.1}$$

As noted earlier, the point of this transformation is that the derivative of z does not appear in (3.1). As compared to the usual proof of the stability of kinetic solutions in the deterministic

case (see, for example, [41]) additional complications appear since the right hand side is not anymore the derivative of a measure.

Indeed

$$(\partial_\xi q)(x + f(\xi)z_t, \xi, t) = \partial_\xi (q(x + f(\xi)z_t, \xi, t)) - D_x q(x + f(\xi)z_t, \xi, t) \cdot \dot{f}(\xi)z_t \quad (3.2)$$

and the second term causes additional difficulties.

Making use of the special properties of the function χ, we find that

$$
\begin{cases}
\|u^{(1)}(t) - u^{(2)}(t)\|_{L^1(\mathbb{R}^N)} = \int (\chi^{(1)}(x, \xi, t) - \chi^{(2)}(x, \xi, t))^2 dx d\xi \\
= \int |\chi^{(1)}(x, \xi, t)| + |\chi^{(2)}(x, \xi, t)| - 2\chi^{(1)}(x, \xi, t)\chi^{(2)}(x, \xi, t) dx d\xi.
\end{cases}
\quad (3.3)
$$

The first step is to rewrite (3.3) in terms of $\tilde{\chi}$. Obviously,

$$\int |\chi^{(1)}(x, \xi, t)| + |\chi^{(2)}(x, \xi, t)| dx d\xi = \int |\tilde{\chi}^{(1)}(x, \xi, t)| + |\tilde{\chi}^{(2)}(x, \xi, t)| dx d\xi.$$

For the third term in (3.3), it is easy to see that

$$\left| \int \chi^{(1)}(x, \xi, t)\chi^{(2)}(x, \xi, t) dx d\xi - \int \tilde{\chi}^{(1)}(x, \xi, t)\tilde{\chi}^{(2)}(x, \xi, t) dx d\xi \right|$$

$$\lesssim \|u^{(2)}\|_{L^\infty([0,T];(L^\infty \cap BV)(\mathbb{R}^N))} |z_t^{(1)} - z_t^{(2)}|,$$

and, hence, for some $C > 0$,

$$\|u^{(1)}(\cdot) - u^{(2)}(\cdot)\|_{L^1(\mathbb{R}^N)} \Big|_s^t \leq \int (\tilde{\chi}^{(1)}(x, \xi, \cdot) - \tilde{\chi}^{(2)}(x, \xi, \cdot))^2 dx d\xi \Big|_s^t$$

$$+ C\|u^{(2)}\|_{L^\infty([0,T];(L^\infty \cap BV)(\mathbb{R}^N))} \|z^{(1)} - z^{(2)}\|_{C([s,t];\mathbb{R}^N)}.$$

Therefore, it is enough to estimate

$$\partial_t \int (\tilde{\chi}^{(1)}(x, \xi, t) - \tilde{\chi}^{(2)}(x, \xi, t))^2 dx d\xi$$

$$= \partial_t \int \left(|\tilde{\chi}^{(1)}(x, \xi, t)| + |\tilde{\chi}^{(2)}(x, \xi, t)| - 2\tilde{\chi}^{(1)}(x, \xi, t)\tilde{\chi}^{(2)}(x, \xi, t) \right) dx d\xi,$$

the advantage being that here we may use (3.1) which makes sense also for non-smooth z.

Step 2: *The product term*. Next we note that

$$-2\frac{d}{dt} \int \tilde{\chi}^{(1)}(x, \xi, t)\tilde{\chi}^{(2)}(x, \xi, t) dx d\xi$$

$$= -2 \int [\tilde{\chi}^{(1)}(x, \xi, t)\partial_t \tilde{\chi}^{(2)}(x, \xi, t) + \partial_t \tilde{\chi}^{(1)}(x, \xi, t)\tilde{\chi}^{(2)}(x, \xi, t)] dx d\xi$$

$$= -2 \sum_{i,j} \int [\tilde{\chi}^{(1)}(x, \xi, t)a_{ij}(\xi)\partial_{ij}^2 \tilde{\chi}^{(2)}(x, \xi, t)$$

$$+ a_{ij}(\xi)\partial_{ij}^2 \tilde{\chi}^{(1)}(x, \xi, t)\tilde{\chi}^{(2)}(x, \xi, t)] dx d\xi$$

$$+ 2 \int [\tilde{\chi}^{(1)}(x, \xi, t)(\partial_\xi q^{(2)})(x + f(\xi)z_t^{(2)}, \xi, t)$$

$$+ (\partial_\xi q^{(1)})(x + f(\xi)z_t^{(1)}, \xi, t)\tilde{\chi}^{(2)}(x, \xi, t)] dx d\xi \quad (3.4)$$

and

$$2 \int [\tilde{\chi}^{(1)}(x, \xi, t)(\partial_\xi q^{(2)})(x + f(\xi)z_t^{(2)}, \xi, t)$$
$$+ (\partial_\xi q^{(1)})(x + f(\xi)z_t^{(1)}, \xi, t)\tilde{\chi}^{(2)}(x, \xi, t)]dxd\xi$$
$$= 2 \int [\chi^{(1)}(x + f(\xi)(z_t^{(1)} - z_t^{(2)}), \xi, t)\partial_\xi q^{(2)}(x, \xi, t)$$
$$+ \partial_\xi q^{(1)}(x, \xi, t)\chi^{(2)}(x + f(\xi)(z_t^{(2)} - z_t^{(1)}), \xi, t)]dxd\xi$$
$$= 2 \int [(\partial_\xi \chi^{(1)})(x + f(\xi)(z_t^{(1)} - z_t^{(2)}), \xi, t)q^{(2)}(x, \xi, t)$$
$$+ q^{(1)}(x, \xi, t)(\partial_\xi \chi^{(2)})(x + f(\xi)(z_t^{(2)} - z_t^{(1)}), \xi, t)]dxd\xi + Err^{(1,2)}.$$

The error term $Err^{(1,2)}$, which is a consequence of the defect identified in (3.2), is given by

$$Err^{(1,2)}(t) := 2 \int \left(f'(\xi)(z_t^{(1)} - z_t^{(2)})\left(D_x\chi^{(1)}(x + f(\xi)(z_t^{(1)} - z_t^{(2)}), \xi, t)q^{(2)}(x, \xi, t)\right.\right.$$
$$\left.\left. - q^{(1)}(x, \xi, t)D_x\chi^{(2)}(x + f(\xi)(z_t^{(2)} - z_t^{(1)}), \xi, t)\right)\right)dxd\xi.$$

We set

$$\tilde{q}^{(i)}(x, \xi, t) := q^{(i)}(x + f(\xi)z_t, \xi, t) \quad \text{and} \quad \tilde{u}^{(i)}(x, \xi, t) := u^{(i)}(x + f(\xi)z_t^{(i)}, t),$$

and observe that

$$2 \int \left((\partial_\xi \chi^{(1)})(x + f(\xi)(z_t^{(1)} - z_t^{(2)}), \xi, t)q^{(2)}(x, \xi, t)\right.$$
$$\left. + q^{(1)}(x, \xi, t)(\partial_\xi \chi^{(2)})(x + f(\xi)(z_t^{(2)} - z_t^{(1)}), \xi, t)\right)dxd\xi$$
$$= 2 \int \left((\partial_\xi \chi^{(1)})(x + f(\xi)z_t^{(1)}, \xi, t)\tilde{q}^{(2)}(x, \xi, t)\right.$$
$$\left. + \tilde{q}^{(1)}(x, \xi, t)(\partial_\xi \chi^{(2)})(x + f(\xi)z_t^{(2)}, \xi, t)\right)dxd\xi$$
$$= 2 \int \left((\delta(\xi) - \delta(\xi - \tilde{u}^{(1)}(x, \xi, t)))\tilde{q}^{(2)}(x, \xi, t)\right.$$
$$\left. + \tilde{q}^{(1)}(x, \xi, t)(\delta(\xi) - \delta(\xi - \tilde{u}^{(2)}(x, \xi, t)))\right)dxd\xi$$
$$\leq -2 \int \left(\delta(\xi - \tilde{u}^{(1)}(x, \xi, t))\tilde{n}^{(2)}(x, \xi, t) + \tilde{n}^{(1)}(x, \xi, t)\delta(\xi - \tilde{u}^{(2)}(x, \xi, t))\right)dxd\xi$$
$$+ 2 \int (q^{(2)}(x, 0, t) + q^{(1)}(x, 0, t))dx.$$

The key difference with the purely hyperbolic case is that here we do not drop the terms containing the parabolic dissipation measures $\tilde{n}^{(1)}$ and $\tilde{n}^{(2)}$. Instead, we will use them to compensate the additional parabolic terms appearing in (3.4).

In conclusion, we have

$$-2\frac{d}{dt} \int \tilde{\chi}^{(1)}(x, \xi, t)\tilde{\chi}^{(2)}(x, \xi, t)dxd\xi \leq I^{par}(t) + I^{hyp}(t) + Err^{(1,2)}(t),$$

where

$$I^{par}(t) := -2\sum_{i,j}\int \big(\tilde{\chi}^{(1)}(x,\xi,t)a_{ij}(\xi)\partial^2_{ij}\tilde{\chi}^{(2)}(x,\xi,t)$$

$$+ a_{ij}(\xi)\partial^2_{ij}\tilde{\chi}^{(1)}(x,\xi,t)\tilde{\chi}^{(2)}(x,\xi,t)\big)dxd\xi$$

$$- 2\int \big(\tilde{n}^{(1)}(x,\xi,t)\delta(\xi - \tilde{u}^{(2)}(x,\xi,t)) + \delta(\xi - \tilde{u}^{(1)}(x,\xi,t))\tilde{n}^{(2)}(x,\xi,t)\big)dxd\xi$$

and

$$I^{hyp}(t) := 2\int \big(q^{(2)}(x,0,t) + q^{(1)}(x,0,t)\big)dx.$$

Step 3: *The hyperbolic term I^{hyp}.*
Multiplying (3.1) by $\mathrm{sgn}(\xi)$ and integrating yields

$$\frac{d}{dt}\int |\tilde{\chi}(x,\xi,t)|dxd\xi = \frac{d}{dt}\int \mathrm{sgn}(\xi)\tilde{\chi}(x,\xi,t)dxd\xi$$

$$= \sum_{i,j}\int \mathrm{sgn}(\xi)a_{ij}(\xi)\partial^2_{ij}\tilde{\chi}(x,\xi,t)dxd\xi$$

$$+ \int \mathrm{sgn}(\xi)(\partial_\xi q)(x + f(\xi)z_t,\xi,t)dxd\xi$$

$$= -2\int q(x,0,t)dx,$$

and thus

$$I^{hyp} = -\frac{d}{dt}\int |\tilde{\chi}^1|dxd\xi - \frac{d}{dt}\int |\tilde{\chi}^2|dxd\xi.$$

Step 4: *The parabolic term I^{par}.* Using integration by parts, we observe

$$-2\sum_{i,j}\int \big(\tilde{\chi}^{(1)}(x,\xi,t)a_{ij}(\xi)\partial^2_{ij}\tilde{\chi}^{(2)}(x,\xi,t) + a_{ij}(\xi)\partial^2_{ij}\tilde{\chi}^{(1)}(x,\xi,t)\tilde{\chi}^{(2)}(x,\xi,t)\big)dxd\xi$$

$$= 4\sum_{i,j}\int a_{ij}(\xi)\partial_j\tilde{\chi}^{(1)}(x,\xi,t)\partial_i\tilde{\chi}^{(2)}(x,\xi,t)dxd\xi$$

$$= 4\sum_{i,j}\sum_k \int \sigma_{ik}(\xi)\sigma_{kj}(\xi)\partial_j\tilde{\chi}^{(1)}(x,\xi,t)\partial_i\tilde{\chi}^{(2)}(x,\xi,t)dxd\xi$$

and informally (this is where the chain-rule will be required in the rigorous proof)

$$\sum_{i,j}\int \sigma_{ik}(\xi)\sigma_{kj}(\xi)\partial_j\tilde{\chi}^{(1)}(x,\xi,t)\partial_i\tilde{\chi}^{(2)}(x,\xi,t)dxd\xi$$

$$= \sum_{i,j}\int \sigma_{kj}(\xi)\partial_j\tilde{u}^{(1)}(x,\xi,t)\delta(\xi - \tilde{u}^{(1)}(x,\xi,t))$$

$$\times \sigma_{ik}(\xi)\partial_i\tilde{u}^{(2)}(x,\xi,t)\delta(\xi - \tilde{u}^{(2)}(x,\xi,t))dxd\xi$$

$$= \sum_{i,j}\int \partial_j\beta_{kj}(\tilde{u}^{(1)}(x,\xi,t))\delta(\xi - \tilde{u}^{(1)}(x,\xi,t))\partial_i\beta_{ik}(\tilde{u}^{(2)}(x,\xi,t))$$

$$\times \delta(\xi - \tilde{u}^{(2)}(x,\xi,t))dxd\xi.$$

Hence,

$$-2\sum_{i,j}\int\big(\tilde{\chi}^{(1)}(x,\xi,t)a_{ij}(\xi)\partial_{ij}^2\tilde{\chi}^{(2)}(x,\xi,t)+a_{ij}(\xi)\partial_{ij}^2\tilde{\chi}^{(1)}(x,\xi,t)\tilde{\chi}^{(2)}(x,\xi,t)\big)dxd\xi$$

$$=4\sum_{i,j}\sum_{k}\int\partial_j\beta_{kj}(\tilde{u}^{(1)}(x,\xi,t))\delta(\xi-\tilde{u}^{(1)}(x,\xi,t))$$

$$\times\,\partial_i\beta_{ik}(\tilde{u}^{(2)}(x,\xi,t))\delta(\xi-\tilde{u}^{(2)}(x,\xi,t))dxd\xi.$$

Since

$$\left(\sum_i\partial_i\beta_{ik}(\tilde{u}^{(1)}(x,t))\right)^2+\left(\sum_j\partial_j\beta_{jk}(\tilde{u}^{(2)}(x,t))\right)^2$$

$$\geq 2\sum_{i,j}\partial_i\beta_{ik}(\tilde{u}^{(1)}(x,t))^2\partial_j\beta_{jk}(\tilde{u}^{(2)}(x,t)),$$

using (2.10) we find

$$-\int\tilde{n}^{(1)}(x,\xi,t)\delta(\xi-\tilde{u}^{(2)}(x,t))+\delta(\xi-\tilde{u}^{(1)}(x,t))\tilde{n}^{(2)}(x,\xi,t)dxd\xi$$

$$\leq -2\sum_{k}\sum_{i,j}\int\delta(\xi-\tilde{u}^{(1)}(x,\xi,t))\delta(\xi-\tilde{u}^{(2)}(x,\xi,t))$$

$$\times\,\partial_i\beta_{ik}(\tilde{u}^{(1)}(x,\xi,t))\partial_j\beta_{jk}(\tilde{u}^{(2)}(x,\xi,t))dxd\xi,$$

and, in conclusion,

$$I^{par}\leq 0.$$

Step 5: Combining the previous steps we find

$$\frac{d}{dt}\int[|\tilde{\chi}^1|-2\tilde{\chi}^{(1)}(x,\xi,t)\tilde{\chi}^{(2)}(x,\xi,t)+|\tilde{\chi}^2|]dxd\xi\leq Err^{(1,2)}(t).$$

Thus, it remains to estimate $Err^{(1,2)}$. For this we approximate $\chi^{(i)}$ by convolution with parameter ε, that is, roughly speaking, we replace $\chi^{(i)}$ by a smooth $\chi^{(i),\varepsilon}$. Since $\chi^{(i)}$ has bounded variation, this introduces an error of order ε, while the gradient $D_x\chi^{(i),\varepsilon}$ blows up like $\frac{1}{\varepsilon}$ as $\varepsilon\to 0$. This leads to the estimate

$$Err^{(1,2)}(t)\lesssim\frac{|z_t^{(1)}-z_t^{(2)}|}{\varepsilon}+\varepsilon.$$

Thus, choosing $\varepsilon=|z_t^{(1)}-z_t^{(2)}|^{1/2}$, we obtain the bound

$$Err^{(1,2)}(t)\leq|z_t^{(1)}-z_t^{(2)}|^{1/2},$$

which allows to conclude the proof.

The rigorous proof of Theorem 2.3. We start with the observation

$$\|u^{(1)}(t)-u^{(2)}(t)\|_{L^1(\mathbb{R}^N)}=\int(\chi^{(1)}(x,\xi,t)-\chi^{(2)}(x,\xi,t))^2dxd\xi$$

$$=\int\big(|\chi^{(1)}(x,\xi,t)|+|\chi^{(2)}(x,\xi,t)|-2\chi^{(1)}(x,\xi,t)\chi^{(2)}(x,\xi,t)\big)dxd\xi,$$

where, for $i = 1, 2$, $\chi^{(i)}$ is related to $u^{(i)}$ by (2.2). To prove (2.14) we want to estimate the time derivative of the right hand side of the above equality. In order to be able to use the solution property, following [24,31,33], we replace $\chi^{(1)}$, $\chi^{(2)}$ by $\varrho * \chi^{(1)}$, $\varrho * \chi^{(2)}$ for suitable choices of ϱ. Then it is possible to differentiate with respect to t at the expense of creating several additional terms that need to be estimated.

***Step* 1**: *Approximation.* We perform three approximations. Firstly we consider the convolutions along characteristics in the space variable, secondly we localize the space and velocity variable and thirdly we regularize in the velocity variable. All these are explained next.

Given $t \in [0, T]$ define

$$G(t) := \int (\chi^{(1)}(y, \eta, t) - \chi^{(2)}(y, \eta, t))^2 dy d\eta,$$

and let $\varrho_\varepsilon^s, \varrho_\delta^v$ be standard smooth mollifiers, that is approximations of Dirac masses, and for $i = 1, 2$ we consider

$$\varrho_\varepsilon^{s,(i)}(x, y, \xi, t) := \varrho_\varepsilon^s(x - f(\xi)z^{(i)}(t) - y)$$

and

$$\varrho_{\varepsilon,\delta}^{(i)}(x, y, \xi, \eta, t) := \varrho_\varepsilon^{s,(i)}(x, y, \xi, t)\varrho_\delta^v(\xi - \eta),$$

that is $\varrho_\varepsilon^{s,(i)}$ is the solution to (2.4) with $z = z^{(i)}$ and initial condition $\varrho_\varepsilon^{s,(i)}(\cdot, y, \xi, 0) = \varrho_\varepsilon^s(\cdot - y)$. Although we can start from the beginning with the $\varrho_{\varepsilon,\delta}^{(i)}$'s, in order to keep the presentation simpler and, in view of the different role played by ε and δ, we work first with the $\varrho_\varepsilon^{s,(i)}$'s.

We then define

$$G_\varepsilon(t) := -2 \int \left(\chi^{(1)} * \varrho_\varepsilon^{s,(1)}\right)(y, \eta, t)\left(\chi^{(2)} * \varrho_\varepsilon^{s,(2)}\right)(y, \eta, t) dy d\eta$$
$$+ \int \operatorname{sgn}(\eta)(\chi^{(1)} * \varrho_\varepsilon^{s,(1)})(y, \eta, t) dy d\eta + \int \operatorname{sgn}(\eta)(\chi^{(2)} * \varrho_\varepsilon^{s,(2)})(y, \eta, t) dy d\eta.$$

Lemma A.3 in Appendix A yields

$$|G_\varepsilon(t) - G(t)| \leq \left(\varepsilon + \|f(u^{(1)}(t))\|_{L^\infty(\mathbb{R}^N)}|z^{(1)}(t) - z^{(2)}(t)|\right)$$
$$\times \left(BV(u^{(1)}(t)) + BV(u^{(2)}(t))\right).$$

Hence,

$$G(t) - G(s) \leq G_\varepsilon(t) - G_\varepsilon(s) + 2\left(\varepsilon + \|f(u^{(1)})\|_{L^\infty(\mathbb{R}^N \times [s,t])}\|z^{(1)} - z^{(2)}\|_{C([s,t];\mathbb{R}^N)}\right)$$
$$\times (\|u^{(1)}\|_{L^\infty([s,t];BV(\mathbb{R}^N))} + \|u^{(2)}\|_{L^\infty([s,t];BV(\mathbb{R}^N))}), \tag{3.5}$$

and to conclude it is enough to derive an appropriate bound on $G_\varepsilon(t) - G_\varepsilon(s)$.

We show, and this will be the main part of the proof, that

$$G_\varepsilon(t) - G_\varepsilon(s) \lesssim \varepsilon^{-1}\|z^{(1)} - z^{(2)}\|_{C([s,t];\mathbb{R}^N)}, \tag{3.6}$$

and combining (3.5) and (3.6) we find

$$G(t) - G(s) \lesssim \varepsilon^{-1}\|z^{(1)} - z^{(2)}\|_{C([s,t];\mathbb{R}^N)}$$
$$+ 2\left(\varepsilon + \|f(u^{(1)})\|_{L^\infty(\mathbb{R}^N \times [s,t])}\|z^{(1)} - z^{(2)}\|_{C([s,t];\mathbb{R}^N)}\right)$$
$$\times (\|u^{(1)}\|_{L^\infty([s,t];BV(\mathbb{R}^N))} + \|u^{(2)}\|_{L^\infty([s,t];BV(\mathbb{R}^N))}).$$

The conclusion follows by choosing $\varepsilon^2 = \|z^{(1)} - z^{(2)}\|_{C([s,t];\mathbb{R}^N)}$. It thus remains to prove (3.6).

Step 2: *The product term.* We note that, while $\chi^{(i)} * \varrho_\varepsilon^{s,(i)}$ is smooth in y, in order to apply Definition 2.1 we still need to mollify with respect to η, that is to consider $\chi^{(i)} * \varrho_{\varepsilon,\delta}^{(i)}$ and, in the end, take $\delta \to 0$. For this mollification to converge, we first need to localize in y and η, a fast which requires yet another layer of approximation.

Therefore, for every $\psi \in C_c^\infty(\mathbb{R}^{N+1})$ with $\|\psi\|_{C(\mathbb{R}^N)} \le 1$, we introduce

$$
\begin{aligned}
G_{\varepsilon,\psi,\delta}(t) := &-2 \int \psi(y,\eta) \left(\chi^{(1)} * \varrho_{\varepsilon,\delta}^{(1)}\right)(y,\eta,t) \left(\chi^{(2)} * \varrho_{\varepsilon,\delta}^{(2)}\right)(y,\eta,t) dy d\eta \\
&+ \int \psi(y,\eta)(\text{sgn} * \varrho_\delta^v)(\eta)(\chi^{(1)} * \varrho_{\varepsilon,\delta}^{(1)})(y,\eta,t) dy d\eta \\
&+ \int \psi(y,\eta)(\text{sgn} * \varrho_\delta^v)(\eta)(\chi^{(2)} * \varrho_{\varepsilon,\delta}^{(2)})(y,\eta,t) dy d\eta.
\end{aligned}
$$

To simplify the notation we suppress the ε, ψ, δ-dependence in the next few steps. We have:

$$
\begin{aligned}
&-2\frac{d}{dt} \int \psi(y,\eta) \left(\chi^{(1)} * \varrho^{(1)}\right) \left(\chi^{(2)} * \varrho^{(2)}\right) dy d\eta \\
&= -2 \int \left[\psi(y,\eta) \left(\chi^{(1)} * \varrho^{(1)}\right) \frac{d}{dt}\left(\chi^{(2)} * \varrho^{(2)}\right) \right. \\
&\qquad \left. + \psi(y,\eta) \left(\chi^{(2)} * \varrho^{(2)}\right) \frac{d}{dt}\left(\chi^{(1)} * \varrho^{(1)}\right) \right] dy d\eta \\
&= -2 \sum_{i,j=1}^{N} \int [\psi(y,\eta)\chi^{(1)}(x,\xi,t)\varrho^{(1)}(x,y,\xi,\eta,t) \\
&\qquad \times \chi^{(2)}(x',\xi',t)a_{ij}(\xi')\partial_{x_i x_j}^2 \varrho^{(2)}(x',y,\xi',\eta,t) \\
&\qquad + \psi(y,\eta)\chi^{(2)}(x',\xi',t)\varrho^{(2)}(x',y,\xi',\eta,t)\chi^{(1)}(x,\xi,t)a_{ij}(\xi) \\
&\qquad \times \partial_{x_i x_j}^2 \varrho^{(1)}(x,y,\xi,\eta,t)]dx d\xi dx' d\xi' dy d\eta \\
&\quad - 2 \int [\psi(y,\eta)\chi^{(1)}(x,\xi,t)\varrho^{(1)}(x,y,\xi,\eta,t) \\
&\qquad \times \varrho^{(2)}(x',y,\xi',\eta,t)\partial_{\xi'}q^{(2)}(x',\xi',t) \\
&\qquad + \psi(y,\eta)\chi^{(2)}(x',\xi',t)\varrho^{(2)}(x',y,\xi',\eta,t) \\
&\qquad \times \varrho^{(1)}(x,y,\xi,\eta,t)\partial_{\xi}q^{(1)}(x,\xi,t)]dx d\xi dx' d\xi' dy d\eta.
\end{aligned}
$$

We first consider the last two terms on the right hand side. As in the hyperbolic case we note

$$
\begin{aligned}
&-2\int \left(\psi(y,\eta)\chi^{(1)}(x,\xi,t)\varrho^{(1)}(x,y,\xi,\eta,t)\varrho^{(2)}(x',y,\xi',\eta,t)\partial_{\xi'}q^{(2)}(x',\xi',t) \right. \\
&\qquad + \psi(y,\eta)\chi^{(2)}(x',\xi',t)\varrho^{(2)}(x',y,\xi',\eta,t) \\
&\qquad \left. \times \varrho^{(1)}(x,y,\xi,\eta,t)\partial_{\xi}q^{(1)}(x,\xi,t)\right)dx d\xi dx' d\xi' dy d\eta \\
&= -2\int \left(\psi(y,\eta)\chi^{(1)}(x,\xi,t)\partial_{\xi}\varrho^{(1)}(x,y,\xi,\eta,t) \right. \\
&\qquad \times \varrho^{(2)}(x',y,\xi',\eta,t)q^{(2)}(x',\xi',t) \\
&\qquad + \psi(y,\eta)\chi^{(2)}(x',\xi',t)\partial_{\xi'}\varrho^{(2)}(x',y,\xi',\eta,t) \\
&\qquad \left. \times \varrho^{(1)}(x,y,\xi,\eta,t)q^{(1)}(x,\xi,t)\right)dx d\xi dx' d\xi' dy d\eta \\
&\quad + Err^{(1,2)}(t),
\end{aligned}
$$

where

$$\begin{cases} Err^{(1,2)}(t) := 2\int \psi(y,\eta)\left(\chi^{(1)}(x,\xi,t)q^{(2)}(x',\xi',t) + \chi^{(2)}(x',\xi',t)q^{(1)}(x,\xi,t)\right) \\ \quad \times (\varrho^{(1)}(x,y,\xi,\eta,t)\partial_{\xi'}\varrho^{(2)}(x',y,\xi',\eta,t) \\ \quad + \partial_\xi\varrho^{(1)}(x,y,\xi,\eta,t)\varrho^{(2)}(x',y,\xi',\eta,t))dyd\eta dx'd\xi'dxd\xi. \end{cases} \quad (3.7)$$

We next use the nonnegativity of the Dirac masses δ to observe that

$$\begin{aligned} -2\int &[\psi(y,\eta)\chi^{(1)}(x,\xi,t)\partial_\xi\varrho^{(1)}(x,y,\xi,\eta,t)\varrho^{(2)}(x',y,\xi',\eta,t)q^{(2)}(x',\xi',t) \\ &+ \psi(y,\eta)\chi^{(2)}(x',\xi',t)\partial_{\xi'}\varrho^{(2)}(x',y,\xi',\eta,t) \\ &\times \varrho^{(1)}(x,y,\xi,\eta,t)q^{(1)}(x,\xi,t)]dxd\xi dx'd\xi'dyd\eta \\ =2\int &\Big[\psi(y,\eta)\left(\delta(\xi) - \delta(\xi - u^{(1)}(x,t))\right)\varrho^{(1)}(x,y,\xi,\eta,t) \\ &\times \varrho^{(2)}(x',y,\xi',\eta,t)q^{(2)}(x',\xi',t) \\ &+ \psi(y,\eta)\left(\delta(\xi') - \delta(\xi' - u^{(2)}(x',t))\right) \\ &\times \varrho^{(2)}(x',y,\xi',\eta,t)\varrho^{(1)}(x,y,\xi,\eta,t)q^{(1)}(x,\xi,t)\Big]dxd\xi dx'd\xi'dyd\eta \\ \le -2\int &[\psi(y,\eta)\delta(\xi - u^{(1)}(x,t))\varrho^{(1)}(x,y,\xi,\eta,t)\varrho^{(2)}(x',y,\xi',\eta,t)n^{(2)}(x',\xi',t) \\ &+ \psi(y,\eta)\delta(\xi' - u^{(2)}(x',t))\varrho^{(2)}(x',y,\xi',\eta,t) \\ &\times \varrho^{(1)}(x,y,\xi,\eta,t)n^{(1)}(x,\xi,t)]dxd\xi dx'd\xi'dyd\eta \\ &+ 2\int \psi(y,\eta)\varrho^{(1)}(x,y,0,\eta,t)\varrho^{(2)}(x',y,\xi',\eta,t)q^{(2)}(x',\xi',t)dxdx'd\xi'dyd\eta \\ &+ 2\int \psi(y,\eta)\varrho^{(1)}(x,y,\xi,\eta,t)\varrho^{(2)}(x',y,0,\eta,t)q^{(1)}(x,\xi,t)dxd\xi dx'dyd\eta, \quad (3.8) \end{aligned}$$

and, hence,

$$-2\frac{d}{dt}\int \psi(y,\eta)\left(\chi^{(1)}*\varrho^{(1)}\right)\left(\chi^{(2)}*\varrho^{(2)}\right)dyd\eta \le I^{par}(t) + I^{hyp}(t) + Err^{(1,2)}(t),$$

with I^{par} and I^{hyp} denoting respectively the parabolic and hyperbolic terms, that is

$$\begin{aligned} I^{par}(t) := -2\sum_{i,j=1}^N \int &\left(\psi(y,\eta)\chi^{(1)}(x,\xi,t)\varrho^{(1)}(x,y,\xi,\eta,t)\chi^{(2)}(x',\xi',r)a_{ij}(\xi')\right. \\ &\times \partial_{x_i'x_j'}\varrho^{(2)}(x',y,\xi',\eta,s) \\ &+ \psi(y,\eta)\chi^{(2)}(x',\xi',t)\varrho^{(2)}(x',y,\xi',\eta,t)\chi^{(1)}(x,\xi,r)a_{ij}(\xi) \\ &\times \partial_{x_ix_j}^2\varrho^{(1)}(x,y,\xi,\eta,s)\big)dxd\xi dx'd\xi'dyd\eta \\ &- 2\int \big(\psi(y,\eta)\delta(\xi - u^{(1)}(x,t))\varrho^{(1)}(x,y,\xi,\eta,t) \\ &\times \varrho^{(2)}(x',y,\xi',\eta,t)n^{(2)}(x',\xi',t) \\ &+ \psi(y,\eta)\delta(\xi' - u^{(2)}(x',t))\varrho^{(2)}(x',y,\xi',\eta,t) \\ &\times \varrho^{(1)}(x,y,\xi,\eta,t)n^{(1)}(x,\xi,t)\big)dxd\xi dx'd\xi'dyd\eta, \end{aligned}$$

and

$$I^{hyp}(t) := 2\int \psi(y,\eta)\varrho^{(1)}(x,y,0,\eta,t)\varrho^{(2)}(x',y,\xi',\eta,t)q^{(2)}(x',\xi',t)dxdx'd\xi'dyd\eta$$
$$+ 2\int \psi(y,\eta)\varrho^{(1)}(x,y,\xi,\eta,t)\varrho^{(2)}(x',y,0,\eta,t)q^{(1)}(x,\xi,t)dxd\xi dx'dyd\eta;$$

note that we call the above expression a hyperbolic term in spite of the occurrence of the parabolic dissipation measures $n^{(1)}$ and $n^{(2)}$, because they have the same structure as and are treated similarly to the hyperbolic case.

Step 3: *The hyperbolic terms.* Since both terms in I^{hyp} are treated similarly, we provide details only for the first one. We have

$$2\int_s^t \int \psi(y,\eta)\varrho^{(1)}(x,y,0,\eta,r)\varrho^{(2)}(x',y,\xi',\eta,r)q^{(2)}(x',\xi',r)dxdx'd\xi'dyd\eta dr$$
$$= \int_s^t \int \psi(y,\eta)\partial_\xi \operatorname{sgn}(\xi)\varrho^{(1)}(x,y,\xi,\eta,r)$$
$$\times \varrho^{(2)}(x',y,\xi',\eta,r)q^{(2)}(x',\xi',r)dxd\xi dx'd\xi'dyd\eta dr$$
$$= -\int_s^t \int \psi(y,\eta)\operatorname{sgn}(\xi)\partial_\xi\varrho^{(1)}(x,y,\xi,\eta,r)\varrho^{(2)}(x',y,\xi',\eta,r)$$
$$\times q^{(2)}(x',\xi',r)dxd\xi dx'd\xi'dyd\eta dr$$
$$= \int_s^t \int \psi(y,\eta)\left(\int \operatorname{sgn}(\xi)\varrho^{(1)}(x,y,\xi,\eta,r)dxd\xi\right)$$
$$\times \left(\int \partial_{\xi'}\varrho^{(2)}(x',y,\xi',\eta,r)q^{(2)}(x',\xi',r)dx'd\xi'\right)dyd\eta dr + \int_s^t Err^{(1)}(r)dr,$$

with

$$Err^{(1)}(r) := -\int \psi(y,\eta)\operatorname{sgn}(\xi)\Big(\partial_\xi\varrho^{(1)}(x,y,\xi,\eta,r)\varrho^{(2)}(x',y,\xi',\eta,r)$$
$$+ \varrho^{(1)}(x,y,\xi,\eta,r)\partial_{\xi'}\varrho^{(2)}(x',y,\xi',\eta,r)\Big)q^{(2)}(x',\xi',r)dxd\xi dx'd\xi'dyd\eta.$$

Using the solution property (2.9) for $\chi^{(2)}$ and

$$\int \operatorname{sgn}(\xi)\varrho^{(1)}(x,y,\xi,\eta,r)dxd\xi = (\operatorname{sgn}*\varrho^v)(\eta)$$

we find

$$2\int_s^t \int \psi(y,\eta)\varrho^{(1)}(x,y,0,\eta,r)\varrho^{(2)}(x',y,\xi',\eta,r)q^{(2)}(x',\xi',r)dxd\xi'dx'dyd\eta dr$$
$$= \int_s^t \int \psi(y,\eta)(\operatorname{sgn}*\varrho^v)(\eta)\left(\int \partial_{\xi'}\varrho^{(2)}(x',y,\xi',\eta,r)q^{(2)}(x',\xi',r)dx'd\xi'\right)dyd\eta dr$$
$$+ \int_s^t Err^{(1)}(r)dr$$
$$= -\int \psi(y,\eta)(\operatorname{sgn}*\varrho^v)(\eta)(\chi^{(2)}*\varrho^{(2)})(y,\eta,\cdot)|_s^t dyd\eta + \int_s^t \int (\operatorname{sgn}*\varrho^v)(\eta)$$
$$\times \left(\psi(y,\eta)\sum_{i,j=1}^N \int \chi^{(2)}(x',\xi',t)a_{ij}(\xi')\partial_{x'_i x'_j}\varrho^{(2)}(x',y,\xi',\eta,t)dx'd\xi'\right)dyd\eta dr$$
$$+ \int_s^t Err^{(1)}(r)dr.$$

Since

$$\partial_{x_i}\varrho^{(i)}(x,y,\xi,\eta,t) = -\partial_{y_i}\varrho^{(i)}(x,y,\xi,\eta,t)$$
$$\partial^2_{x_i x_j}\varrho^{(i)}(x,y,\xi,\eta,t) = \partial^2_{y_i y_j}\varrho^{(i)}(x,y,\xi,\eta,t) \tag{3.9}$$

we have

$$\int_s^t \int (\mathrm{sgn} * \varrho^v)(\eta)\psi(y,\eta)$$
$$\times \sum_{i,j=1}^N \int \chi^{(2)}(x',\xi',t)a_{ij}(\xi')\partial_{x'_i x'_j}\varrho^{(2)}(x',y,\xi',\eta,t)dx'd\xi'dyd\eta dr$$
$$= \int_s^t \int (\mathrm{sgn} * \varrho^v)(\eta)\partial_{y_i}\psi(y,\eta)$$
$$\times \sum_{i,j=1}^N \int \chi^{(2)}(x',\xi',t)a_{ij}(\xi')\partial_{x'_j}\varrho^{(2)}(x',y,\xi',\eta,t)dx'd\xi'dyd\eta dr$$
$$=: \int_s^t Err^{loc,(1)}(r)dr.$$

In conclusion,

$$\int_s^t I^{hyp}(r)dr = -\int \psi(y,\eta)(\mathrm{sgn} * \varrho^v)(\eta)(\chi^{(1)} * \varrho^{(1)})(y,\eta,\cdot)|_s^t dyd\eta$$
$$- \int \psi(y,\eta)(\mathrm{sgn} * \varrho^v)(\eta)(\chi^{(2)} * \varrho^{(2)})(y,\eta,\cdot)|_s^t dyd\eta$$
$$+ \int_s^t Err^{(1)}(r) + Err^{(2)}(r)dr + \int_s^t Err^{loc,(1)}(r) + Err^{loc,(2)}(r)dr,$$

with $Err^{(2)}$, $Err^{loc,(2)}$ defined analogously to $Err^{(1)}$, $Err^{loc,(1)}$ respectively.

***Step* 4**: *The parabolic terms.* Using (3.9) we get

$$\sum_{i,j=1}^N \int \psi(y,\eta)\chi^{(1)}(x,\xi,t)\varrho^{(1)}(x,y,\xi,\eta,t)\chi^{(2)}(x',\xi',t)a_{ij}(\xi')$$
$$\times \partial_{x'_i x'_j}\varrho^{(2)}(x',y,\xi',\eta,t)dxd\xi dx'd\xi'dyd\eta$$
$$= -\sum_{i,j=1}^N \int \psi(y,\eta)\chi^{(1)}(x,\xi,t)\chi^{(2)}(x',\xi',t)a_{ij}(\xi')$$
$$\times \partial_{x_i}\varrho^{(1)}(x,y,\xi,\eta,t)\partial_{x'_j}\varrho^{(2)}(x',y,\xi',\eta,t)dxd\xi dx'd\xi'dyd\eta + Err^{loc,(3)},$$

where

$$Err^{loc,(3)} := -\sum_{i,j=1}^N \int \partial_{y_i}\psi(y,\eta)\chi^{(1)}(x,\xi,t)\chi^{(2)}(x',\xi',t)a_{ij}(\xi')$$
$$\times \varrho^{(1)}(x,y,\xi,\eta,t)\partial_{x'_j}\varrho^{(2)}(x',y,\xi',\eta,t)dxd\xi dx'd\xi'dyd\eta.$$

Hence,

$$-2\sum_{i,j=1}^N \int [\psi(y,\eta)\chi^{(1)}(x,\xi,t)\varrho^{(1)}(x,y,\xi,\eta,t)$$
$$\times \chi^{(2)}(x',\xi',t)a_{ij}(\xi')\partial_{x'_i x'_j}\varrho^{(2)}(x',y,\xi',\eta,t)$$

$$
+ \psi(y,\eta)\chi^{(2)}(x',\xi',t)\varrho^{(2)}(x',y,\xi',\eta,t)\chi^{(1)}(x,\xi,t)a_{ij}(\xi)
$$
$$
\times \partial^2_{x_i x_j}\varrho^{(1)}(x,y,\xi,\eta,t)]dxd\xi dx'd\xi'dyd\eta
$$
$$
= 4\sum_{i,j=1}^{N}\sum_{k=1}^{N}\int [\psi(y,\eta)\chi^{(1)}(x,\xi,t)\chi^{(2)}(x',\xi',t)\sigma_{ik}(\xi)\sigma_{kj}(\xi')
$$
$$
\times \partial_{x_i}\varrho^{(1)}(x,y,\xi,\eta,t)\partial_{x'_j}\varrho^{(2)}(x',y,\xi',\eta,t)]dxd\xi dx'd\xi'dyd\eta
$$
$$
+ Err^{par} + Err^{loc,(3)} + Err^{loc,(4)}, \tag{3.10}
$$

with

$$
Err^{par} := 2\sum_{i,j=1}^{N}\int \psi(y,\eta)\chi^{(1)}(x,\xi,t)\chi^{(2)}(x',\xi',t)
$$
$$
\times \left(a_{ij}(\xi) - 2\sum_{k=1}^{N}\sigma_{ik}(\xi)\sigma_{kj}(\xi') + a_{ij}(\xi')\right)
$$
$$
\times \partial_{x_i}\varrho^{(1)}(x,y,\xi,\eta,t)\partial_{x'_j}\varrho^{(2)}(x',y,\xi',\eta,t)dxd\xi dx'd\xi'dyd\eta
$$

and $Err^{loc,(4)}$ defined analogously to $Err^{loc,(3)}$.

We now use the chain rule (2.8) and get

$$
\sum_{i,j=1}^{N}\int \psi(y,\eta)\chi^{(1)}(x,\xi,t)\chi^{(2)}(x',\xi',t)\sigma_{ik}(\xi)\sigma_{kj}(\xi')
$$
$$
\times \partial_{x_i}\varrho^{(1)}(x,y,\xi,\eta,t)\partial_{x'_j}\varrho^{(2)}(x',y,\xi',\eta,t)dxd\xi dx'd\xi'dyd\eta
$$
$$
= \int \psi(y,\eta)\left(\sum_{i=1}^{N}\int \chi^{(1)}(x,\xi,t)\sigma_{ik}(\xi)\partial_{x_i}\varrho^{(1)}(x,y,\xi,\eta,t)dxd\xi\right)
$$
$$
\times \left(\sum_{j=1}^{N}\int \chi^{(2)}(x',\xi',t)\sigma_{kj}(\xi')\partial_{x'_j}\varrho^{(2)}(x',y,\xi',\eta,t)dx'd\xi'\right)dyd\eta
$$
$$
= \int \psi(y,\eta)\left(\sum_{i=1}^{N}\int \partial_{x_i}\beta_{ik}(u^{(1)}(x,t))\varrho^{(1)}(x,y,u^{(1)}(x,t),\eta,t)dx\right)
$$
$$
\times \left(\sum_{j=1}^{N}\int \partial_{x'_j}\beta_{kj}(u^{(2)}(x',t))\varrho^{(2)}(x',y,u^{(2)}(x',t),\eta,t)dx'\right)dyd\eta,
$$

and, therefore,

$$
- 2\sum_{i,j=1}^{N}\int \left(\psi(y,\eta)\chi^{(1)}(x,\xi,t)\varrho^{(1)}(x,y,\xi,\eta,t)\chi^{(2)}(x',\xi',t)\right.
$$
$$
\times a_{ij}(\xi')\partial^2_{x_i x_j}\varrho^{(2)}(x',y,\xi',\eta,t)
$$
$$
+ \psi(y,\eta)\chi^{(2)}(x',\xi',t)\varrho^{(2)}(x',y,\xi',\eta,t)
$$
$$
\left. \times \chi^{(1)}(x,\xi,t)a_{ij}(\xi)\partial^2_{x_i x_j}\varrho^{(1)}(x,y,\xi,\eta,t)\right)dxd\xi dx'd\xi'dyd\eta
$$
$$
= 4\sum_{i,j=1}^{N}\sum_{k=1}^{N}\int \psi(y,\eta)\left(\int \partial_{x_i}\beta_{ik}(u^{(1)}(x,t))\varrho^{(1)}(x,y,u^{(1)}(x,t),\eta,t)\right.
$$

$$\times\, \partial_{x_j}\beta_{kj}(u^{(2)}(x',t))\varrho^{(2)}(x',y,u^{(2)}(x',t),\eta,t)dxdx'\Big)dyd\eta$$

$$+\, Err^{par} + Err^{loc,(3)} + Err^{loc,(4)}.$$

Using (2.10) we analyze next the last two terms in I^{par}. We have:

$$\int \psi(y,\eta)\delta(\xi - u^{(1)}(x,t))\varrho^{(1)}(x,y,\xi,\eta,t)$$

$$\times\, \varrho^{(2)}(x',y,\xi',\eta,t)n^{(2)}(x',\xi',t)dxd\xi dx'd\xi' dyd\eta$$

$$+\, \int \psi(y,\eta)\delta(\xi' - u^{(2)}(x',t))\varrho^{(2)}(x',y,\xi',\eta,t)$$

$$\times\, \varrho^{(1)}(x,y,\xi,\eta,t)n^{(1)}(x,\xi,t)dxd\xi dx'd\xi' dyd\eta$$

$$=\, \int \psi(y,\eta)\varrho^{(1)}(x,y,u^{(1)}(x,t),\eta,t)\varrho^{(2)}(x',y,\xi',\eta,t)n^{(2)}(x',\xi',t)dxdx'd\xi' dyd\eta$$

$$+\, \int \psi(y,\eta)\varrho^{(2)}(x',y,u^{(2)}(x',t),\eta,t)\varrho^{(1)}(x,y,\xi,\eta,t)n^{(1)}(x,\xi,t)dxd\xi dx'dyd\eta$$

$$=\, \int \psi(y,\eta)\varrho^{(1)}(x,y,u^{(1)}(x,t),\eta,t)\varrho^{(2)}(x',y,u^{(2)}(t,x'),\eta,t)$$

$$\times\, \sum_{k=1}^{N}\left(\sum_{i=1}^{N}\partial_{x_i'}\beta_{ik}(u^{(2)}(t,x'))\right)^2 dxdx'dyd\eta$$

$$+\, \int \psi(y,\eta)\varrho^{(2)}(x',y,u^{(2)}(x',t),\eta,t)\varrho^{(1)}(x,y,u^{(1)}(t,x),\eta,t)$$

$$\times\, \sum_{k=1}^{N}\left(\sum_{i=1}^{N}\partial_{x_i}\beta_{ik}(u^{(1)}(t,x))\right)^2 dxdx'dyd\eta.$$

Since

$$\left(\sum_{i=1}^{N}\partial_{x_i}\beta_{ik}(u^{(1)}(t,x))\right)^2 + \left(\sum_{i=1}^{N}\partial_{x_i}\beta_{ik}(u^{(2)}(t,x'))\right)^2$$

$$\geq 2\sum_{i,j=1}^{N}\partial_{x_i}\beta_{ik}(u^{(1)}(t,x))\partial_{x_j'}\beta_{kj}(u^{(2)}(t,x')),$$

we obtain

$$-2\int \psi(y,\eta)\delta(\xi - u^{(1)}(x,t))\varrho^{(1)}(x,y,\xi,\eta,t)$$

$$\times\, \varrho^{(2)}(x',y,\xi',\eta,t)n^{(2)}(x',\xi',t)dxd\xi dx'd\xi' dyd\eta$$

$$+\, \int \psi(y,\eta)\delta(\xi' - u^{(2)}(x',t))\varrho^{(2)}(x',y,\xi',\eta,t)$$

$$\times\, \varrho^{(1)}(x,y,\xi,\eta,t)n^{(1)}(x,\xi,t)dxd\xi dx'd\xi' dyd\eta$$

$$\leq\, -4\int \psi(y,\eta)\varrho^{(1)}(x,y,u^{(1)}(x,t),\eta,t)\varrho^{(2)}(x',y,u^{(2)}(t,x'),\eta,t)$$

$$\times\, \sum_{k=1}^{N}\left(\sum_{i,j=1}^{N}\partial_{x_i}\beta_{ik}(u^{(1)}(t,x))\partial_{x_j'}\beta_{jk}(u^{(2)}(t,x'))\right)dxdx'dyd\eta,$$

and, then,

$$\int_s^t I^{par}(r)dr \le \int_s^t Err^{par}(r) + Err^{loc,(3)}(r) + Err^{loc,(4)}(r)dr.$$

Step 5: *The end of the proof.* Combining the estimates in the last two steps we find, now explicitly writing the ε, ψ, δ-dependence,

$$
\begin{aligned}
-2\int \psi(y,\eta)&(\chi^{(1)} * \varrho_{\varepsilon,\delta}^{(1)})(y,\eta,\cdot)(\chi^{(2)} * \varrho_{\varepsilon,\delta}^{(2)})(y,\eta,\cdot)dyd\eta\Big|_s^t \\
\le &-\int \psi(y,\eta)(\operatorname{sgn} * \varrho_\delta^v)(\eta)(\chi^{(1)} * \varrho_{\varepsilon,\delta}^{(1)})(y,\eta,\cdot)|_s^t dyd\eta \\
&-\int \psi(y,\eta)(\operatorname{sgn} * \varrho_\delta^v)(\eta)(\chi^{(2)} * \varrho_{\varepsilon,\delta}^{(2)})(y,\eta,\cdot)|_s^t dyd\eta \\
&+\int_s^t \Big(Err_{\varepsilon,\psi,\delta}^{(1)}(r) + Err_{\varepsilon,\psi,\delta}^{(2)}(r) + Err_{\varepsilon,\psi,\delta}^{(1,2)}(r) + Err_{\varepsilon,\psi,\delta}^{par}(r)\Big)dr \\
&+\int_s^t \Big(Err_{\varepsilon,\psi,\delta}^{loc,(1)}(r) + Err_{\varepsilon,\psi,\delta}^{loc,(2)}(r) + Err_{\varepsilon,\psi,\delta}^{loc,(3)}(r) + Err_{\varepsilon,\psi,\delta}^{loc,(4)}(r)\Big)dr,
\end{aligned}
\tag{3.11}
$$

that is

$$
\begin{aligned}
G_{\varepsilon,\psi,\delta}(t) - G_{\varepsilon,\psi,\delta}(s) \le &\int_s^t \Big(Err_{\varepsilon,\psi,\delta}^{(1)}(r) + Err_{\varepsilon,\psi,\delta}^{(2)}(r) + Err_{\varepsilon,\psi,\delta}^{(1,2)}(r) + Err_{\varepsilon,\psi,\delta}^{par}(r)\Big)dr \\
&+\int_s^t \Big(Err_{\varepsilon,\psi,\delta}^{loc,(1)}(r) + Err_{\varepsilon,\psi,\delta}^{loc,(2)}(r) + Err_{\varepsilon,\psi,\delta}^{loc,(3)}(r) + Err_{\varepsilon,\psi,\delta}^{loc,(4)}(r)\Big)dr.
\end{aligned}
$$

It follows from (A.2), (A.4), Lemma A.4 and Lemma A.5 in Appendix A that, for $\psi \in C_c^\infty(\mathbb{R}^{N+1})$ and $\varepsilon > 0$ fixed,

$$
\begin{aligned}
\lim_{\delta \to 0} G_{\varepsilon,\psi,\delta}(t) &- \lim_{\delta \to 0} G_{\varepsilon,\psi,\delta}(s) \\
&\lesssim \int_s^t \int \|\partial_\eta \psi(\cdot,\xi)\|_{C(\mathbb{R}^N)}(q^{(1)} + q^{(2)})(x,\xi,r)dxd\xi dr + \varepsilon^{-1}\|z^{(1)} - z^{(2)}\|_{C([s,t];\mathbb{R}^N)} \\
&\quad + \int_s^t \Big(Err_{\varepsilon,\psi}^{loc,(1)}(r) + Err_{\varepsilon,\psi}^{loc,(2)}(r) + Err_{\varepsilon,\psi}^{loc,(3)}(r) + Err_{\varepsilon,\psi}^{loc,(4)}(r)\Big)dr.
\end{aligned}
\tag{3.12}
$$

Choosing $\psi_R \in C_c^\infty(\mathbb{R}^{N+1})$ such that

$$\psi_R(y,\eta) = \begin{cases} 1, & \text{if } |(y,\eta)| \le R \\ 0, & \text{if } |(y,\eta)| > R+1, \end{cases} \quad \text{and} \quad \|D\psi_R\| \le 1,
\tag{3.13}$$

yields

$$\lim_{R \to \infty} \int \|\partial_\eta \psi_R(\cdot,\xi)\|_{C(\mathbb{R}^N)}(q^{(1)} + q^{(2)})(x,\xi,t)dxd\xi = 0.$$

In view of Lemma A.6 we also have

$$\lim_{R \to \infty} \int_s^t \Big(Err_{\varepsilon,\psi_R}^{loc,(1)}(r) + Err_{\varepsilon,\psi_R}^{loc,(2)}(r) + Err_{\varepsilon,\psi_R}^{loc,(3)}(r) + Err_{\varepsilon,\psi_R}^{loc,(4)}(r)\Big)dr = 0$$

and it is easy to see that

$$\lim_{R \to \infty} \lim_{\delta \to 0} G_{\varepsilon,\psi_R,\delta}(t) = G_\varepsilon(t).$$

Hence, letting $R \to \infty$ in (3.12) yield

$$G_\varepsilon(t) - G_\varepsilon(s) \le \varepsilon^{-1} \|z^{(1)} - z^{(2)}\|_{C([s,t];\mathbb{R}^N)}, \tag{3.14}$$

which finishes the proof by step one.

4. The existence of pathwise entropy solutions

We prove the existence of pathwise entropy solutions to (1.1) and establish the a priori estimates stated in Theorem 2.4. The solution is found as a limit of solutions of a three-step approximation procedure. In the first step, the initial condition $u_0 \in (BV \cap L^\infty)(\mathbb{R}^N)$ is approximated by smooth functions $u_0^\delta \in C_c^\infty(\mathbb{R}^N)$ such that $\|u_0^\delta\|_2 \le \|u_0\|_2$. In the second step, the driving signal z is approximated by smooth driving signals $z^{(l)}$. In the third step, A is approximated by $A^\varepsilon + \varepsilon I$, where I is the $N \times N$ identity matrix and $A^\varepsilon(\xi) := A * \varphi^\varepsilon(\xi)$ with φ^ε being a standard Dirac family.

In conclusion, we consider the smooth solution $u^{(\varepsilon,\delta,l)}$ to

$$\begin{cases} \partial_t u^{(\varepsilon,\delta,l)} + \sum_{i=1}^N \partial_{x_i} F^i(u^{(\varepsilon,\delta,l)})(\dot{z}^{(l)})^i = \mathrm{div}(A^\varepsilon(u^{(\varepsilon,\delta,l)})Du^{(\varepsilon,\delta,l)}) + \varepsilon \Delta u^{(\varepsilon,\delta,l)} \text{ in } \mathbb{R}^N \times (0,T), \\ u^{(\varepsilon,\delta,l)} = u_0^\delta \text{ on } \mathbb{R}^N \times \{0\}. \end{cases} \tag{4.1}$$

The existence and uniqueness of $u^{(\varepsilon,\delta,l)}$ for each fixed ε, δ and l is classical; see, for example, Volpert and Hudjaev [45]. The proof of the bounds in Theorem 2.4 is based on establishing similar bounds for $u^{(\varepsilon,\delta,l)}$ and then passing to the limit.

Proof of Theorem 2.4. Since the proof is long, we divide it into several steps.

***Step* 1**: *The approximating* equation (4.1). The kinetic function $\chi^{(\varepsilon,\delta,l)}(x,\xi,t) := \chi(u^{(\varepsilon,\delta,l)}(x,t),\xi)$ solves

$$\begin{cases} \partial_t \chi^{(\varepsilon,\delta,l)} + f^{(l)}(\xi,t) \cdot D_x \chi^{(\varepsilon,\delta,l)} - \sum_{i,j=1}^d a_{ij}^\varepsilon(\xi)\partial_{x_i x_j}^2 \chi^{(\varepsilon,\delta,l)} - \varepsilon \Delta \chi^{(\varepsilon,\delta,l)} \\ \qquad = \partial_\xi q^{(\varepsilon,\delta,l)} \text{ in } \mathbb{R}^N \times \mathbb{R} \times (0,T), \\ \chi^{(\varepsilon,\delta,l)} = \chi(u_0^\delta(\cdot),\cdot) \text{ on } \mathbb{R}^N \times \mathbb{R} \times \{0\}, \end{cases} \tag{4.2}$$

where

$$f^{(l)}(\xi,t) := F'(\xi)\dot{z}^{(l)}(t)$$

and

$$q^{(\varepsilon,\delta,l)} = m^{(\varepsilon,\delta,l)} + n^{(\varepsilon,\delta,l)}$$

with the entropy dissipation and parabolic measures $m^{(\varepsilon,\delta,l)}$ and $n^{(\varepsilon,\delta,l)}$ given respectively by

$$m^{(\varepsilon,\delta,l)}(x,\xi,t) = \delta(\xi - u^{(\varepsilon,\delta,l)})\varepsilon|Du^{(\varepsilon,\delta,l)}|^2$$

and

$$n^{(\varepsilon,\delta,l)}(x,\xi,t) = \delta(\xi - u^{(\varepsilon,\delta,l)})\sum_{k=1}^N \left(\sum_{i=1}^N \partial_{x_i}\beta_{ik}^\varepsilon(u^{(\varepsilon,\delta,l)})\right)^2.$$

Moreover, it is shown in [45] that, for all $t \geq 0$ and $p \in [1, \infty]$, there exist constants $C^{(l)} > 0$ such that

$$\|u^{(\varepsilon,\delta,l)}(t)\|_{L^p} \leq \|u_0\|_p, \quad \|Du^{(\varepsilon,\delta,l)}(t)\|_{L^1} \leq BV(u_0) \quad \text{and} \quad \|\partial_t u^{(\varepsilon,\delta,l)}(t)\|_{L^1} \leq C^{(l)}.$$
$$(4.3)$$

Multiplying (4.2) by ξ and integrating yields

$$\int_{\mathbb{R}^N \times \mathbb{R} \times \mathbb{R}_+} m^{(\varepsilon,\delta,l)}(x,\xi,t)dxd\xi dt + \int_{\mathbb{R}^N \times \mathbb{R}_+} \sum_{k=1}^{N} \left(\sum_{i=1}^{N} \partial_{x_i} \beta_{ik}^{\varepsilon}(u^{(\varepsilon,\delta,l)}) \right)^2 dxdt$$
$$\leq \frac{1}{2} \|u_0^{\delta}\|_2^2 \leq \frac{1}{2} \|u_0\|_2^2. \quad (4.4)$$

In view of Lemma 2.2, $u^{(\varepsilon,\delta,l)}$ is a pathwise entropy solution, that is, in the sense of distributions in t,

$$\frac{d}{dt}(\varrho^{(l)} * \chi^{(\varepsilon,\delta,l)})(y,\eta,t) = \sum_{i,j=1}^{N} \int \chi^{(\varepsilon,\delta,l)}(x,\xi,t)a_{ij}^{\varepsilon}(\xi)\partial_{x_i x_j}^2 \varrho^{(l)}(x,y,\xi,\eta,t)dxd\xi$$
$$+ \varepsilon \int \chi^{(\varepsilon,\delta,l)}(x,\xi,t)\Delta\varrho^{(l)}(x,y,\xi,\eta,t)dxd\xi$$
$$- \int \partial_\xi \varrho^{(l)}(x,y,\xi,\eta,t)q^{(\varepsilon,\delta,l)}(x,\xi,t)dxd\xi. \quad (4.5)$$

Elementary calculations also give

$$\sum_{i=1}^{N} \int \chi^{(\varepsilon,\delta,l)}(x,\xi,t)\sigma_{ik}^{\varepsilon}(\xi)\partial_{x_i}\varrho^{(l)}(x,y,\xi,\eta,t)dxd\xi$$
$$= -\sum_{i=1}^{N} \int \partial_{x_i}\beta_{ik}^{\varepsilon}(u^{(\varepsilon,\delta,l)}(x,t))\varrho^{(l)}(x,y,u^{(\varepsilon,\delta,l)}(x,t),\eta,t)dx. \quad (4.6)$$

Step 2: *The singular degenerate limit $\varepsilon \to 0$.* In view of the estimates (4.3) and (4.4), there exist subsequences, which we denote again by $u^{(\varepsilon,\delta,l)}$, $m^{(\varepsilon,\delta,l)}$ and $n^{(\varepsilon,\delta,l)}$, such that, as $\varepsilon \to 0$,

$$u^{(\varepsilon,\delta,l)} \to u^{(\delta,l)} \quad \text{in } C([0,T]; L^1(\mathbb{R}^N)),$$
$$m^{(\varepsilon,\delta,l)} \rightharpoonup m^{(\delta,l)},$$
$$n^{(\varepsilon,\delta,l)} \rightharpoonup n^{(\delta,l)} \text{ weak } \star,$$
$$\sum_{i=1}^{N} \partial_{x_i}\beta_{ik}^{\varepsilon}(u^{(\varepsilon,\delta,l)}) \rightharpoonup \sum_{i=1}^{N} \partial_{x_i}\beta_{ik}(u^{(\delta,l)}) \quad \text{in } L^2([0,T] \times \mathbb{R}^N),$$

and thus

$$n^{(\varepsilon,\delta,l)} \rightharpoonup n^{(\delta,l)} = \delta(\xi - u^{(\delta,l)}) \sum_{k=1}^{N} \left(\sum_{i=1}^{N} \partial_{x_i}\beta_{ik}(u^{(\delta,l)}) \right)^2.$$

Passing to the limit in (4.5) and (4.6) yields respectively (2.9) for $z = z^{(l)}$ and

$$\sum_{i=1}^{N} \int \chi^{(\delta,l)}(x, \xi, t) \sigma_{ik}(\xi) \partial_{x_i} \varrho^{(l)}(x, y, \xi, \eta, t) dx d\xi$$

$$= \sum_{i=1}^{N} \int \partial_{x_i} \beta_{ik}(u^{(\delta,l)}(x, t)) \varrho^{(l)}(x, y, u^{(\delta,l)}(x, t), \eta, t) dx.$$

Using the lower semicontinuity with respect to weak convergence, we also get

$$\int_0^T \int_{\mathbb{R}^N \times \mathbb{R}} m^{(\delta,l)}(x, \xi, t) dx d\xi dt + \int_0^T \int_{\mathbb{R}^N} \sum_{k=1}^N \left(\sum_{i=1}^N \partial_{x_i} \beta_{ik}(u^{(\delta,l)}) \right)^2 dx dt \le \frac{1}{2} \|u_0\|_2^2.$$

Step 3: *The rough signal limit $l \to \infty$ and initial condition $\delta \to 0$*. Theorem 2.3 yields that, as $l, m \to \infty$,

$$\|u^{(\delta,l)} - u^{(\delta,m)}\|_{C([0,T];L^1(\mathbb{R}^N))} \le C\|z^{(l)} - z^{(m)}\|_{C([0,T];\mathbb{R}^N)}^{1/2} \to 0,$$

that is, $u^{(\delta,l)}$ is a Cauchy sequence in $C([0, T]; L^1(\mathbb{R}^N))$ and thus has, as $l \to \infty$, a limit $u^{(\delta)}$ in $C([0, T]; L^1(\mathbb{R}^N))$.

We then argue as in Step 2 to obtain a pathwise entropy solution $u^{(\delta)}$ to (1.1) with initial condition u_0^δ. Again, Theorem 2.3 yields that, as $\delta_1, \delta_2 \to 0$,

$$\|u^{(\delta_1)} - u^{(\delta_1)}\|_{C([0,T];L^1(\mathbb{R}^N))} \le C\|u_0^{\delta_1} - u_0^{\delta_2}\|_{L^1(\mathbb{R}^N)}$$

and arguing as before we obtain a pathwise entropy solution u.

The bound (2.18) follows easily by testing with $\xi^{[m+1]} = |\xi|^m \xi$ and a cut-off argument. Since the argument is routine, we leave the details to the reader.

5. Long-time behavior—the proof of Theorem 2.6

The general approach is based on averaging techniques related to the classical averaging lemmata for scalar conservation laws. Typically, the proofs of averaging lemmata use Fourier transforms in space and time. In the stochastic context this is not possible due to the time-dependence of the flux. Therefore, we only use Fourier transforms in the spatial variable x. This technique was developed by Bouchut and Desvillettes [6] and was used for semilinear stochastic scalar conservation laws by Debussche and Vovelle in [15]. Although our proof follows the arguments of [15] and more closely [25], new difficulties arise because of the second order term in (1.6). In particular, we have to adapt the important technical result (Lemma 5.1), which relies on the genuine nonlinearity condition (2.20). In what follows we are brief about parts similar to [25] and we concentrate on the differences. Since the proof is rather long, we divide it in several subsections. We also remark that we use properties of the Brownian paths and our approach does not extend to general continuous time dependence.

Split-up of the solution. Without loss of generality, we restrict to initial conditions with zero average, that is we assume that

$$\int u_0(x) dx = 0; \tag{5.1}$$

the case of non-zero spatial average can be easily reduced to this case. Moreover, in view of (2.14) and the density of $L^\infty(\mathbb{T}^N)$ in $L^1(\mathbb{T}^N)$, it is enough to consider $u_0 \in L^\infty(\mathbb{T}^N)$.

As a regularizing term we use the fractional Laplace operator (see [6,15] for similar types of arguments)

$$B := (-\Delta)^\alpha + Id \quad \text{with } \alpha \in (0, 1].$$ (5.2)

Let $S_{A_\gamma(\xi)}(s, t)$ denote the solution operator of

$$\partial_t v + \sum_i f^i(\xi)\dot{\beta}^i(t)\partial_{x_i} v - \sum_{i,j=1}^d a_{ij}(\xi)\partial^2_{x_i x_j} v + \gamma B v = 0 \quad \text{in } \mathbb{T}^N \times \mathbb{R} \times (s, \infty).$$ (5.3)

Since $A(\xi) : D^2 = \sum_{i,j=1}^d a_{ij}(\xi)\partial^2_{x_i x_j}$ and B commute, it is immediate that, for all f in the appropriate function space,

$$\begin{aligned}
S_{A_\gamma(\xi)}(s, t)f(x) &= (e^{-(t-s)(A(\xi):D^2+\gamma B)} f)(x - f(\xi)(\beta(t) - \beta(s))), \\
&= (e^{-(t-s)A(\xi):D^2} e^{-(t-s)\gamma B} f)(x - f(\xi)(\beta(t) - \beta(s))),
\end{aligned}$$ (5.4)

where e^{tA} denotes the solution semigroup to the operator A and

$$f(\xi)(\beta(t) - \beta(s)) = (f^1(\xi)(\beta^1(t) - \beta^1(s)), \ldots, f^N(\xi)(\beta^N(t) - \beta^N(s))).$$

For $n \in \mathbb{Z}^N$, the Fourier transform of $S_{A_\gamma(\xi)}$ corresponds to multiplication by

$$\exp\left(-if(\xi)(\beta(t) - \beta(s)) \cdot n - \left(nA(\xi)n - \gamma(|n|^{2\alpha} + 1)\right)(t - s)\right).$$

It follows from the variation of constants formula that, for $t \in [0, T]$ and $\varphi \in C^\infty(\mathbb{T}^N)$,

$$\begin{aligned}
\int_{\mathbb{T}^N} \varphi(x)u(x, t)dx &= \int_{\mathbb{T}^N \times \mathbb{R}} \varphi(x)\chi(x, \xi, t)dxd\xi \\
&= \int_{\mathbb{T}^N \times \mathbb{R}} \varphi(x)S_{A_\gamma(\xi)}(0, t)\chi_0(x, \xi)dxd\xi \\
&\quad + \int_0^t \int_{\mathbb{T}^N \times \mathbb{R}} \gamma B(S^*_{A_\gamma(\xi)}(s, t)\varphi)(x)\chi(x, \xi, s)dxd\xi ds \\
&\quad - \int_0^t \int_{\mathbb{T}^N \times \mathbb{R}} \partial_\xi (S^*_{A_\gamma(\xi)}(s, t)\varphi)(x)dq(x, \xi, s),
\end{aligned}$$ (5.5)

where $S^*_{A_\gamma(\xi)}$ denotes the adjoint semigroup to $S_{A_\gamma(\xi)}$ and for simplicity we set again $q = m+n$. Noting that (2.18) implies $q(\mathbb{T}^N \times \{0\} \times \mathbb{R}_+) = 0$ and that $\xi \mapsto (S^*_{A_\gamma(\xi)}(s, t)\varphi)(x)$ is continuously differentiable on $\mathbb{R} \setminus \{0\}$, (5.5) can be justified following the same arguments as in [25, Appendix C] in combination with dominated convergence based on (2.18) and (1.7). We note that, in comparison to [25], the additional parabolic term of (1.6) is represented in (5.5) via the parabolic term in the semigroup $S_{A_\gamma(\xi)}$.

Recall that, a.s. in ω, $u \in C([0, \infty); L^1(\mathbb{T}^N))$. Accordingly, in the sense of distributions in x, we have

$$u(t) = u^0(t) + u^1(t) + Q(t),$$ (5.6)

where, for $\varphi \in C^\infty(\mathbb{T}^N)$,

$$\langle u^0(t), \varphi \rangle := \int_{\mathbb{T}^N \times \mathbb{R}} \varphi(x)S_{A_\gamma(\xi)}(0, t)\chi_0(x, \xi)dxd\xi,$$ (5.7)

$$\langle u^1(t), \varphi \rangle := \int_0^t \int_{\mathbb{T}^N \times \mathbb{R}} \gamma B(S^*_{A_\gamma(\xi)}(s,t)\varphi)(x)\chi(x,\xi,s)dxd\xi ds \tag{5.8}$$

$$\langle Q(t), \varphi \rangle := -\int_0^t \int_{\mathbb{T}^N \times \mathbb{R}} \partial_\xi(S^*_{A_\gamma(\xi)}(s,t)\varphi)(x)q(x,\xi,s)dxd\xi ds. \tag{5.9}$$

Next we estimate (5.7)–(5.9) separately using averaging techniques. In the analysis we need a basic integral estimate which is proved in Appendix C. For its statement it is convenient to introduce, for each measurable $b : \mathbb{R} \to \mathbb{R}^N$, $a : \mathbb{R} \to \mathbb{R}_+$ and $f \in L^2$, the function $\phi(\cdot; a, b, f) : \mathbb{R}^N \to \mathbb{R}$ given by

$$\phi(w; a, b, f) := e^{-\frac{|w|^2}{\delta}} \int_{\mathbb{R}} e^{ib(\xi)\cdot w - \delta a(\xi)} f(\xi)d\xi. \tag{5.10}$$

Lemma 5.1. *Let $b : \mathbb{R} \to \mathbb{R}^N$, $a : \mathbb{R} \to \mathbb{R}_+$ be measurable functions, such that, for all $\varepsilon > 0, z \in \mathbb{R}^N$ and some nondecreasing $\iota : [0, \infty) \to [0, \infty)$ with $\lim_{\varepsilon \to 0} i(\varepsilon) = 0$,*

$$|\{\xi \in \mathbb{R} : |b(\xi) - z|^2 + a(\xi) \le \varepsilon\}| \le \iota(\varepsilon).$$

Then, for all $\delta > 0$ and $f \in L^2(\mathbb{R})$,

$$\|\phi(\cdot; a, b, f)\|_{L^2}^2 \le \frac{\sqrt{\delta\pi}}{4} \int_0^\infty e^{-\frac{\tau}{4}} \iota\left(\frac{\tau}{\delta}\right)d\tau \|f\|_2^2.$$

The estimate of u^0. Taking Fourier transforms in (5.7) yields, for each $n \in \mathbb{Z}^N$,

$$\hat{u}^0(n, t) = \int e^{-if(\xi)\beta(t)\cdot n - (nA(\xi)n + \gamma(|n|^{2\alpha}+1))t} \hat{\chi}_0(n, \xi)d\xi.$$

As in [25] we have $\hat{u}^0(0, t) = 0$. For $n \in \mathbb{Z}^N \setminus \{0\}$, integrating in time, taking expectations and using the scaling properties of the Brownian paths, we find

$$\mathbb{E} \int_0^T |\hat{u}^0(n, t)|^2 dt$$

$$= \mathbb{E} \int_0^T \left| \int e^{-if(\xi)\beta(t)\cdot n - (nA(\xi)n + \gamma(|n|^{2\alpha}+1))t} \hat{\chi}_0(n, \xi)d\xi \right|^2 dt$$

$$= \int_0^T e^{-2\gamma(|n|^{2\alpha}+1)t} \mathbb{E} \left| \int e^{-if(\xi)\beta(t|n|^2)\cdot \frac{n}{|n|} - nA(\xi)nt} \hat{\chi}_0(n, \xi)d\xi \right|^2 dt$$

$$= \int_0^T \frac{e^{-2\gamma(|n|^{2\alpha}+1)t}}{\sqrt{2\pi|n|^2 t}} \int \left| e^{-\frac{|w|^2}{4|n|^2 t}} \int e^{-if(\xi)\frac{n}{|n|}\cdot w - \frac{n}{|n|}A(\xi)\frac{n}{|n|}|n|^2 t} \hat{\chi}_0(n, \xi)d\xi \right|^2 dw dt.$$

Lemma 5.1 with $\delta = 4|n|^2 t$, $b(\xi) = f(\xi) \cdot \frac{n}{|n|}$, $a(\xi) = \frac{1}{4}\frac{n}{|n|}A(\xi)\frac{n}{|n|}$ and $i(\varepsilon) \sim \varepsilon^{\frac{\theta}{2}}$ gives

$$\mathbb{E} \int_0^T |\hat{u}^0(n, t)|^2 dt \le \frac{\sqrt{\delta\pi}}{4} \int_0^T \frac{e^{-2\gamma(|n|^{2\alpha}+1)t}}{\sqrt{2\pi|n|^2 t}} \int_0^\infty e^{-\frac{\tau}{4}} \iota\left(\frac{\tau}{\delta}\right)d\tau dt \|\hat{\chi}_0(n, \cdot)\|_2^2$$

$$\lesssim \delta^{\frac{1-\theta}{2}} \int_0^T \frac{e^{-2\gamma(|n|^{2\alpha}+1)t}}{\sqrt{2\pi|n|^2 t}} \int_0^\infty \tau^{\frac{\theta}{2}} e^{-\tau} d\tau dt \|\hat{\chi}_0(n, \cdot)\|_2^2$$

$$\lesssim \int_0^T e^{-2\gamma(|n|^{2\alpha}+1)t} (|n|^2 t)^{-\frac{\theta}{2}} dt \|\hat{\chi}_0(n, \cdot)\|_2^2$$

$$\lesssim |n|^{-\theta} \int_0^T e^{-2\gamma(|n|^{2\alpha}+1)t} t^{-\frac{\theta}{2}} dt \|\hat{\chi}_0(n,\cdot)\|_2^2$$

$$\lesssim |n|^{-\theta} \gamma^{-\frac{2-\theta}{2}} (|n|^{2\alpha}+1)^{-\frac{2-\theta}{2}} \int_0^\infty e^{-t} t^{-\frac{\theta}{2}} dt \|\chi_0(\cdot,\cdot)\|_2^2,$$

and, hence,

$$\mathbb{E} \int_0^T |\hat{u}^0(n,t)|^2 dt \leq C |n|^{-\theta} \gamma^{-\frac{2-\theta}{2}} (|n|^{2\alpha}+1)^{-\frac{2-\theta}{2}} \|\hat{\chi}_0(n,\cdot)\|_2^2$$

$$\lesssim \gamma^{-\frac{2-\theta}{2}} \|\hat{\chi}_0(n,\cdot)\|_2^2. \tag{5.11}$$

Combining the previous estimates, after summing over n, we obtain

$$\mathbb{E} \int_0^T \|u^0(t)\|_2^2 dt = \mathbb{E} \int_0^T \|\hat{u}^0(t)\|^2 dt \lesssim \gamma^{-\frac{2-\theta}{2}} \|\chi_0(\cdot,\cdot)\|_2^2$$

$$\lesssim \gamma^{-\frac{2-\theta}{2}} \|\chi_0(\cdot,\cdot)\|_1 = \gamma^{-\frac{2-\theta}{2}} \|u_0\|_1. \tag{5.12}$$

The estimate of u^1. Let $\bar{\omega}_n := \gamma(|n|^{2\alpha}+1)$. For each $n \in \mathbb{Z}^N$, the Fourier transform $\hat{u}^1(n,t)$ of $u^1(t)$ in x is given by

$$\hat{u}^1(n,t) = \bar{\omega}_n \int_0^t \int e^{-if(\xi)(\beta(t)-\beta(s))\cdot n - (nA(\xi)n + \gamma(|n|^{2\alpha}+1))(t-s)} \hat{\chi}(n,\xi,s) d\xi ds$$

$$= \bar{\omega}_n \int_0^t e^{-\bar{\omega}_n(t-s)} \int e^{-if(\xi)(\beta(t)-\beta(s))\cdot n - nA(\xi)n(t-s)} \hat{\chi}(n,\xi,s) d\xi ds.$$

Integrating in t, taking expectation and using that $\int_0^t \bar{\omega}_n e^{-\bar{\omega}_n r} dr \leq 1$, we find

$$\mathbb{E} \int_0^T |\hat{u}^1|^2(n,t) dt$$

$$= \mathbb{E} \int_0^T \left| \int_0^t \bar{\omega}_n e^{-\bar{\omega}_n(t-s)} \int e^{-if(\xi)(\beta(t)-\beta(s))\cdot n - nA(\xi)n(t-s)} \hat{\chi}(n,\xi,s) d\xi ds \right|^2 dt$$

$$= \mathbb{E} \int_0^T \left| \int_0^t \bar{\omega}_n e^{-\bar{\omega}_n r} \int e^{-if(\xi)(\beta(t)-\beta(t-r))\cdot n - nA(\xi)nr} \hat{\chi}(n,\xi,t-r) d\xi dr \right|^2 dt$$

$$\leq \mathbb{E} \int_0^T \int_0^t \bar{\omega}_n e^{-\bar{\omega}_n r} \left| \int e^{-if(\xi)(\beta(t)-\beta(t-r))\cdot n - nA(\xi)nr} \hat{\chi}(n,\xi,t-r) d\xi \right|^2 dr dt.$$

Using that $\hat{\chi}$ is $\mathcal{F}_t$-adapted, that the increments $\beta(t) - \beta(t-r)$ are independent of $\mathcal{F}_{t-r}$ and the scaling properties of the Brownian motion, as in [25], we find

$$\mathbb{E} \left| \int e^{-if(\xi)(\beta(t)-\beta(t-r))\cdot n - nA(\xi)nr} \hat{\chi}(n,\xi,t-r) d\xi \right|^2$$

$$= \mathbb{E}\tilde{\mathbb{E}} \left| \int e^{-if(\xi)\beta(|n|^2 r)(\tilde{\omega})\cdot \frac{n}{|n|} - \left(\frac{n}{|n|} A(\xi) \frac{n}{|n|}\right)|n|^2 r} \hat{\chi}(n,\xi,t-r)(\omega) d\xi \right|^2$$

$$= \frac{1}{\sqrt{2\pi |n|^2 r}} \mathbb{E} \int \left| e^{-\frac{|w|^2}{4|n|^2 r}} \int e^{-if(\xi)w\cdot \frac{n}{|n|} - \left(\frac{n}{|n|} A(\xi) \frac{n}{|n|}\right)|n|^2 r} \hat{\chi}(n,\xi,t-r) d\xi \right|^2 dw,$$

where $\tilde{E}$ is the expectation with respect to $\tilde{\omega}$. Using again Lemma 5.1 with $\delta = 4|n|^2 r$, $b(\xi) = f(\xi) \cdot \frac{n}{|n|}$ and $a(\xi) = \frac{1}{4}\left(\frac{n}{|n|} A(\xi) \frac{n}{|n|}\right)$ we get

$$\int \left| e^{-\frac{|w|^2}{4|n|^2 r}} \int e^{-if(\xi)w\cdot\frac{n}{|n|}-\left(\frac{n}{|n|}A(\xi)\frac{n}{|n|}\right)|n|^2 r} \hat{\chi}(n,\xi,t-r)d\xi \right|^2 dw$$
$$\lesssim \sqrt{\delta} \int_0^\infty e^{-\frac{\tau}{4}} \iota\left(\frac{\tau}{\delta}\right) d\tau \|\hat{\chi}(n,\cdot,t-r)\|_2^2$$
$$\lesssim \delta^{\frac{1-\theta}{2}} \int_0^\infty \tau^{\frac{\theta}{2}} e^{-\tau} d\tau \|\hat{\chi}(n,\cdot,t-r)\|_2^2 \lesssim (|n|^2 r)^{\frac{1-\theta}{2}} \|\hat{\chi}(n,\cdot,t-r)\|_2^2,$$

and, hence,

$$\mathbb{E}\left| \int e^{-if(\xi)(\beta(t)-\beta(t-r))\cdot n - nA(\xi)nr} \hat{\chi}(n,t-r)d\xi \right|^2 \lesssim (|n|^2 r)^{-\frac{\theta}{2}} \mathbb{E}\|\hat{\chi}(n,\cdot,t-r)\|_2^2.$$

Combining all the above estimates we find

$$\mathbb{E}\int_0^T |\hat{u}^1|^2(n,t)dt \lesssim \int_0^T \int_0^t \bar{\omega}_n e^{-\bar{\omega}_n r}(|n|^2 r)^{-\frac{\theta}{2}} \mathbb{E}\|\hat{\chi}(n,\cdot,t-r)\|_2^2 dr dt.$$

Young's inequality then yields

$$\mathbb{E}\int_0^T |\hat{u}^1|^2(n,t)dt \lesssim \int_0^T \bar{\omega}_n e^{-\bar{\omega}_n r}(|n|^2 r)^{-\frac{\theta}{2}} dr \int_0^T \mathbb{E}\|\hat{\chi}(n,\cdot,r)\|_2^2 dr,$$

and, in view of the fact that, for $\theta \in [0,1]$,

$$\int_0^T \bar{\omega}_n e^{-\bar{\omega}_n r} r^{-\frac{\theta}{2}} dr \le \bar{\omega}_n^{1+\frac{\theta}{2}} \int_{\mathbb{R}} e^{-\bar{\omega}_n r}(\bar{\omega}_n r)^{-\frac{\theta}{2}} dr = \bar{\omega}_n^{\frac{\theta}{2}} \int_{\mathbb{R}} e^{-r} r^{-\frac{\theta}{2}} dr < \infty,$$

we conclude that, for $n \in \mathbb{Z}^N \setminus \{0\}$,

$$\mathbb{E}\int_0^T |\hat{u}^1|^2(n,t)dt \lesssim \gamma^{\frac{\theta}{2}}(|n|^{2\alpha-2} + |n|^{-2})^{\frac{\theta}{2}} \mathbb{E}\|\hat{\chi}(n,\cdot,\cdot)\|_{L^2(\mathbb{R}\times[0,T])}^2. \tag{5.13}$$

Hence,

$$\mathbb{E}\int_0^T \|u^1(t)\|_2^2 dt \lesssim \gamma^{\frac{\theta}{2}} \mathbb{E}\int_0^T \|u(t)\|_1 dt.$$

The estimate of Q. For $\lambda \ge 0$ and $\varphi \in L^\infty(\Omega \times [0,T]; C^\infty(\mathbb{T}^N))$ (in what follows, unless necessary, we do not display the dependence of φ in ω and t), let

$$\left\langle (-\Delta)^{\frac{\lambda}{2}} Q(t), \varphi(t) \right\rangle$$
$$:= \int_0^t \int \partial_\xi \left(S_{A_\gamma(\xi)}(s,t)\left((-\Delta)^{\frac{\lambda}{2}}\varphi(t)\right)\right)(x)dq(x,\xi,s)$$
$$= \int_0^t \int_{\mathbb{T}^N \times (\mathbb{R}\setminus\{0\})} \partial_\xi \left(S_{A_\gamma(\xi)}(s,t)\left((-\Delta)^{\frac{\lambda}{2}}\varphi(t)\right)\right)(x)dq(x,\xi,s)$$
$$= \int_0^t \int_{\mathbb{T}^N \times (\mathbb{R}\setminus\{0\})} \left(f'(\xi)(\beta_t - \beta_s)\cdot D + A'(\xi) : D^2(t-s)\right)$$
$$\times S_{A_\gamma(\xi)}^*(s,t)(-\Delta)^{\frac{\lambda}{2}}\varphi(t)(x)dq(x,\xi,s),$$

where the first equality follows from (2.18) and the second equality is immediate from the definition of $S_{A_\gamma(\xi)}$ and $S^*_{A_\gamma(\xi)}$.

In view of Lemma B.1, for any $\psi \in C^\infty(\mathbb{T}^N)$, we have

$$
\begin{aligned}
&\left\| D(S^*_{A_\gamma(\xi)}(s,t)(-\Delta)^{\frac{\lambda}{2}}\psi) \right\|_\infty \\
&= \left\| D\left(e^{-\gamma(t-s)(\gamma B+A(\xi):D^2)}(-\Delta)^{\frac{\lambda}{2}}\psi(\cdot)\right)(x-f(\xi)(\beta(t)-\beta(s))) \right\|_\infty \\
&= e^{-\gamma(t-s)}\left\| \left(e^{-\gamma(t-s)A(\xi):D^2}De^{-\gamma(t-s)(-\Delta)^\alpha}(-\Delta)^{\frac{\lambda}{2}}\psi(\cdot)\right)(x-f(\xi)(\beta(t)-\beta(s))) \right\|_\infty \\
&\leq e^{-\gamma(t-s)}\left\| De^{-\gamma(t-s)(-\Delta)^\alpha}(-\Delta)^{\frac{\lambda}{2}}\psi \right\|_\infty \lesssim e^{-\gamma(t-s)}(\gamma(t-s))^{-\frac{\lambda+1}{2\alpha}}\|\psi\|_\infty.
\end{aligned}
$$

Again using Lemma B.1 we have

$$
\begin{aligned}
&\left\| \partial^2_{x_i x_j}\left(S^*_{A_\gamma(\xi)}(s,t)(-\Delta)^{\frac{\lambda}{2}}\psi\right) \right\|_\infty \\
&= \left\| \partial^2_{x_i x_j}\left(e^{-(t-s)(\gamma B+A(\xi):D^2)}(-\Delta)^{\frac{\lambda}{2}}\psi(\cdot)\right)(x-f(\xi)(\beta(t)-\beta(s))) \right\|_\infty \\
&= e^{-\gamma(t-s)}\left\| \left(e^{-(t-s)A(\xi):D^2}\partial^2_{x_i x_j}e^{-(t-s)\gamma(-\Delta)^\alpha}(-\Delta)^{\frac{\lambda}{2}}\psi(\cdot)\right)(x-f(\xi)(\beta(t)-\beta(s))) \right\|_\infty \\
&\leq e^{-\gamma(t-s)}\left\| \partial^2_{x_i x_j}e^{-(t-s)\gamma(-\Delta)^\alpha}(-\Delta)^{\frac{\lambda}{2}}\psi(\cdot) \right\|_\infty \lesssim e^{-\gamma(t-s)}(\gamma(t-s))^{-\frac{\lambda+2}{2\alpha}}\|\psi\|_\infty.
\end{aligned}
$$

In view of (1.7), it follows that, for all $\varphi \in L^\infty(\Omega \times [0,T]; C^\infty(\mathbb{T}^N))$,

$$
\begin{aligned}
&\mathbb{E}\int_0^T \left\langle (-\Delta)^{\frac{\lambda}{2}}Q(t), \varphi(t) \right\rangle dt \\
&\leq C\|\varphi\|_\infty \mathbb{E}\int_0^T \int_0^t e^{-\gamma(t-s)}\left(\gamma^{-\frac{\lambda+1}{2\alpha}}|\beta(t)-\beta(s)|(t-s)^{-\frac{\lambda+1}{2\alpha}}+\gamma^{-\frac{\lambda+2}{2\alpha}}(t-s)^{1-\frac{\lambda+2}{2\alpha}}\right) \\
&\quad \times \int_{\mathbb{T}^N \times (\mathbb{R}\setminus\{0\})}(1+|\xi|^{p_1}+|\xi|^{p_2})dq(x,\xi,s)dt \\
&\leq C\|\varphi\|_\infty \mathbb{E}\int_0^T \int_0^t e^{-\gamma(t-s)}\left(\gamma^{-\frac{\lambda+1}{2\alpha}}|\beta(t)-\beta(s)|(t-s)^{-\frac{\lambda+1}{2\alpha}}+\gamma^{-\frac{\lambda+2}{2\alpha}}(t-s)^{1-\frac{\lambda+2}{2\alpha}}\right) \\
&\quad \times \int_{\mathbb{T}^N \times \mathbb{R}}(1+|\xi|^{p_1}+|\xi|^{p_2})dq(x,\xi,s)dt.
\end{aligned}
$$

At this point we need to argue as in the previous step by taking conditional expectation and using that $\beta(t)-\beta(t-r)$ is independent of $\mathcal{F}_{t-r}$. Since q is only a measure, to make the following argument rigorous it is necessary to perform another approximation. This is done in [25] and we omit the details here.

Using the independence of $\beta(t)-\beta(s)$ from $\mathcal{F}_s$ and $\mathcal{F}_s$-measurability of q we find

$$
\mathbb{E}|\beta(t)-\beta(s)|q(s) = \mathbb{E}[\mathbb{E}|\beta(t)-\beta(s)|q(s)|\mathcal{F}_s] = \mathbb{E}|\beta(t)-\beta(s)|\mathbb{E}q(s) = \sqrt{t-s}\,\mathbb{E}q(s),
$$

and employing once more Young's inequality we obtain

$$
\begin{aligned}
&\mathbb{E}\int_0^T \int_0^t e^{-\gamma(t-s)}\left(|\beta(t)-\beta(s)|\gamma^{-\frac{\lambda+1}{2\alpha}}(t-s)^{-\frac{\lambda+1}{2\alpha}}+\gamma^{-\frac{\lambda+2}{2\alpha}}(t-s)^{1-\frac{\lambda+2}{2\alpha}}\right) \\
&\quad \times \int_{\mathbb{T}^N \times \mathbb{R}}(1+|\xi|^{p_1}+|\xi|^{p_2})dq(x,\xi,s)dt \\
&= \int_0^T \int_0^t e^{-\gamma(t-s)}\left(\gamma^{-\frac{\lambda+1}{2\alpha}}(t-s)^{\frac{1}{2}-\frac{\lambda+1}{2\alpha}}+\gamma^{-\frac{\lambda+2}{2\alpha}}(t-s)^{1-\frac{\lambda+2}{2\alpha}}\right)\mathbb{E}
\end{aligned}
$$

$$\times \int_{\mathbb{T}^N \times \mathbb{R}} (1 + |\xi|^{p_1} + |\xi|^{p_2}) dq(x, \xi, s) dt$$

$$\leq \int_0^T e^{-\gamma t} \left(\gamma^{-\frac{\lambda+1}{2\alpha}} t^{\frac{1}{2} - \frac{\lambda+1}{2\alpha}} + \gamma^{-\frac{\lambda+2}{2\alpha}} t^{1 - \frac{\lambda+2}{2\alpha}} \right) dt$$

$$\times \int_0^T \mathbb{E} \int_{\mathbb{T}^N \times \mathbb{R}} (1 + |\xi|^{p_1} + |\xi|^{p_2}) dq(x, \xi, s) dt.$$

In conclusion,

$$\mathbb{E} \int_0^T \left\langle (-\Delta)^{\frac{\lambda}{2}} Q(t), \varphi(t) \right\rangle dt \lesssim \|\varphi\|_\infty \int_0^T e^{-\gamma t} \left(\gamma^{-\frac{\lambda+1}{2\alpha}} t^{\frac{1}{2} - \frac{\lambda+1}{2\alpha}} + \gamma^{-\frac{\lambda+2}{2\alpha}} t^{1 - \frac{\lambda+2}{2\alpha}} \right) dt$$

$$\times \mathbb{E} \int_0^T \int_{\mathbb{T}^N \times \mathbb{R}} (1 + |\xi|^{p_1} + |\xi|^{p_2}) dq(x, \xi, s).$$

Moreover note that, if $\delta > -1$, then

$$\int_0^T t^\delta e^{-\gamma t} dt = \gamma^{-\delta} \int_0^T (\gamma t)^\delta e^{-\gamma t} dt = \gamma^{-\delta-1} \int_{\mathbb{R}_+} t^\delta e^{-t} dt \leq C \gamma^{-\delta-1},$$

and for $\delta = \frac{1}{2} - \mu_{\alpha,\lambda}^{(1)} = \frac{1}{2} - \frac{\lambda+1}{2\alpha}$, $\delta = 1 - \mu_{\alpha,\lambda}^{(2)} = 1 - \frac{\lambda+2}{2\alpha}$ and assuming $\mu_{\alpha,\lambda}^{(2)} < 2$ (note that this also implies $\mu_{\alpha,\lambda}^{(1)} < \frac{3}{2}$), we get

$$\int_0^T e^{-\gamma t} \left(\gamma^{-\mu_{\alpha,\lambda}^{(1)}} t^{\frac{1}{2} - \mu_{\alpha,\lambda}^{(1)}} + \gamma^{-\mu_{\alpha,\lambda}^{(2)}} t^{1 - \mu_{\alpha,\lambda}^{(2)}} \right) dt \leq C \left(\gamma^{-\frac{3}{2}} + \gamma^{-2} \right).$$

We use next (2.18) and get, in view of all the above,

$$\mathbb{E} \int_0^T \left\langle (-\Delta)^{\frac{\lambda}{2}} Q(t), \varphi(t) \right\rangle dt \lesssim \|\varphi\|_\infty \left(\gamma^{-\frac{3}{2}} + \gamma^{-2} \right) (\|u_0\|_2^2 + \|u_0\|_{p_1+2}^{p_1+2} + \|u_0\|_{p_2+2}^{p_2+2}). \tag{5.14}$$

For now it is enough to take $\lambda = 0$. Then $\mu_{\alpha,0}^{(1)} < \frac{3}{2}$, $\mu_{\alpha,0}^{(2)} < 2$ is satisfied if $\alpha > \frac{1}{2}$, and, for $\gamma \leq 1$, we obtain

$$\mathbb{E} \int_0^T \langle Q(t), \varphi(t) \rangle dt \lesssim \gamma^{-2} \|\varphi\|_\infty (\|u_0\|_2^2 + \|u_0\|_{p_1+2}^{p_1+2} + \|u_0\|_{p_2+2}^{p_2+2}).$$

Proof of Theorem 2.6. We conclude the proof as in [25]. Note that as compared to [25] the constants change due to the changed constants in the estimate of Q. We obtain

$$\mathbb{E}\|u(T)\|_1 \lesssim T^{-\frac{1}{2} + a\left(\frac{2-\theta}{4}\right)} \|u_0\|_1^{\frac{1}{2}} + 2T^{-\frac{a\theta}{2}} + T^{2a-1} (\|u_0\|_2^2 + \|u_0\|_{p_1+2}^{p_1+2} + \|u_0\|_{p_2+2}^{p_2+2})$$

and letting $a = \frac{2}{4+\theta}$, for $T \geq 1$, we get

$$\mathbb{E}\|u(T)\|_1 \lesssim T^{-\frac{\theta}{4+\theta}} \left(\|u_0\|_1^{\frac{1}{2}} + 1 + \|u_0\|_2^2 + \|u_0\|_{p_1 \vee p_2 + 2}^{p_1 \vee p_2 + 2} \right);$$

note that the rate is independent of the choice of α. The proof is concluded as in [25]. $\square$

6. Regularity (proof of Theorem 2.7)

The remaining argument is precisely the same as in [25]. First, assuming $\chi = \chi(u) \in L^2(\mathbb{R} \times [0, T] \times \Omega; H^\tau(\mathbb{R}^N))$ for some $\tau \in [0, 1]$ this implies

$$\mathbb{E} \int_0^T \|u(t)\|_{W^{\lambda,1}} dt \le C(1 + \|u_0\|_{p_1+2}^{p_1+2} + \|u_0\|_{p_2+2}^{p_2+2}), \tag{6.1}$$

for λ satisfying the constraint

$$\lambda < (4\alpha - 2) \wedge \left(\frac{\theta(1-\alpha)}{2} + \tau \right).$$

Note that the constants change as compared to [25] due to the changed constants in the estimation of Q. Maximizing the right hand side yields $\alpha = \frac{\theta+4+2\tau}{\theta+8} \in (0, 1)$ and we obtain

$$\lambda < 4\alpha - 2 = \frac{2\theta}{\theta + 8} + \frac{8\tau}{\theta + 8}.$$

As in [25] this bound is now bootstrapped, which yields that (6.1) holds for all $\lambda \in \left(0, \frac{2\theta}{\theta+4} \right)$. The proof is then concluded as in [25].

Acknowledgments

Appendix A. Convolution error estimates

We study here the behavior of the several error terms introduced in the proof of the comparison principle for pathwise entropy solutions.

The first type of errors, which are studied in the first subsection below, are due to the convolution along characteristics, that is, when we replace $\chi(x, \xi, t)$ by $\tilde{\chi}(x, \xi, t) := \chi(x + f(\xi)z_t, \xi, t)$ and $\tilde{\chi}$ by its approximation $\chi * \varrho_\varepsilon^s$. We then consider the parabolic error terms arising from doubling of variables/mollification of the velocity variable ξ in the second subsection. In the last subsection we estimate the hyperbolic error which is a consequence of the failure of the right hand side in (3.1) to be the ξ-derivative of a nonnegative measure.

Convolution along characteristics. Let

$$\varrho_\varepsilon^s(x, y, \xi, t) := \varrho_\varepsilon^s(x - y + f(\xi)z(t)),$$

where ϱ_ε^s is a standard Dirac family.

Lemma A.1. *Let $u \in BV(\mathbb{R}^N)$. Then, for all $t \in [0, T]$, $p \in [1, \infty)$,*

$$\| \left(\chi * \varrho_\varepsilon^s \right)(y, \xi, t) - \chi(y - f(\xi)z(t), \xi) \|_{L^p(\mathbb{R}^N \times \mathbb{R})} \le \varepsilon BV(u).$$

Proof. Using Hölder's inequality and $u \in BV$ we find

$$\| (\chi * \varrho_\varepsilon^s) (y, \xi, t) - \chi(y - f(\xi)z(t), \xi) \|_{L^p(\mathbb{R}^N \times \mathbb{R})}^p$$

$$= \int \left| \int \chi(x, \xi) \varrho_\varepsilon^s(x, y, \xi, t) dx - \chi(y - f(\xi)z(t), \xi) \right|^p dy d\xi$$

$$= \int \left| \int [\chi(x, \xi) - \chi(y - f(\xi)z(t), \xi)] \varrho_\varepsilon^s(x, y, \xi, t) dx \right|^p dy d\xi$$

$$\leq \int \left| \left(\int \varrho_\varepsilon^s(x, y, \xi, t) dx \right)^{\frac{1}{q}} \right.$$

$$\left. \left(\int |\chi(x, \xi) - \chi(y - f(\xi)z(t), \xi)|^p \varrho_\varepsilon^s(x, y, \xi, t) dx \right)^{\frac{1}{p}} \right|^p dy d\xi$$

$$\leq \iint |\chi(x, \xi) - \chi(y - f(\xi)z(t), \xi)|^p \varrho_\varepsilon^{s,0}(x - y + f(\xi)z(t)) dx dy d\xi$$

$$= \iint |\chi(x, \xi) - \chi(y, \xi)| \varrho_\varepsilon^s(x - y) dx dy d\xi$$

$$= \int |u(x) - u(y)| \varrho_\varepsilon^s(x - y) dx dy$$

$$\leq \varepsilon BV(u). \quad \square$$

Lemma A.2. *Let $u \in (BV \cap L^\infty)(\mathbb{R}^N)$. Then, for all $t \in [0, T]$,*

$$\| \chi(y - f(\xi)z(t), \xi) - \chi(y, \xi) \|_{L^1(\mathbb{R}^N \times \mathbb{R})} \leq \| f(u) \|_\infty |z(t)| BV(u).$$

Proof. Due to [23, Lemma C.1] we have that

$$\| u \|_{BV} = \int_\mathbb{R} BV(\chi(\cdot, \xi)) d\xi,$$

with $\chi(x, \xi) := \chi(u(x), \xi)$. It follows that

$$\int |\chi(y - f(\xi)z(t), \xi) - \chi(y, \xi)| dy d\xi \leq \int BV(\chi(\cdot, \xi)) |f(\xi)z(t)| d\xi$$

$$\leq |z(t)| \| f(u) \|_\infty \int BV(\chi(\cdot, \xi)) d\xi. \quad \square$$

Let

$$\tilde{G}(t) := -2 \int \chi^{(1)}(y - f(\xi)z^{(1)}(t), \xi, t) \chi^{(2)}(y - f(\xi)z^{(1)}(t), \xi, t) dy d\xi$$

$$+ \int \text{sgn}(\eta) \chi^{(1)}(y - f(\xi)z^{(1)}(t), \xi, t) dy d\xi$$

$$+ \int \text{sgn}(\eta) \chi^{(2)}(y - f(\xi)z^{(2)}(t), \xi, t) dy d\xi$$

$$= -2 \int \chi^{(1)}(y - f(\xi)(z^{(1)}(t) - z^{(2)}(t)), \xi, t) \chi^{(2)}(y, \xi, t) dy d\xi$$

$$+ \int \text{sgn}(\eta) \chi^{(1)}(y, \xi, t) dy d\xi + \int \text{sgn}(\eta) \chi^{(2)}(y, \xi, t) dy d\xi.$$

Lemma A.1, with G_ε as in Section 3, yields

$$|G_\varepsilon(t) - \tilde{G}(t)| \leq \varepsilon(BV(u^{(1)}(t)) + BV(u^{(2)}(t)))$$

and, with G as in Section 3, Lemma A.2 implies

$$|\tilde{G}(t) - G(t)| \leq \|f(u^{(1)})\|_\infty |z^{(1)}(t) - z^{(2)}(t)| BV(u^{(1)}(t)).$$

In conclusion, we have:

Lemma A.3. *If $u \in (BV \cap L^\infty)(\mathbb{R}^N)$, then for all $t \in [0, T]$,*

$$|G_\varepsilon(t) - G(t)| \leq \left(\varepsilon + \|f(u^{(1)}(t))\|_\infty |z^{(1)}(t) - z^{(2)}(t)|\right)(BV(u^{(1)}(t)) + BV(u^{(2)}(t))).$$

Parabolic error. We study the parabolic error $Err^{par}_{\varepsilon,\psi,\delta}$ occurring in the proof of Theorem 2.3.

Lemma A.4. *Let $s, t \in [0, T]$, $s < t$, $\psi \in C^\infty_c(\mathbb{R}^{N+1})$, $\varepsilon > 0$. Then, as $\delta \to 0$, $\int_s^t Err^{par}_{\varepsilon,\psi,\delta}(r)dr \to 0$.*

Proof. Straightforward calculations lead to

$$
\begin{aligned}
Err^{par}_{\varepsilon,\psi,\delta}(r) = {}& 2\sum_{i,j=1}^N \iint \psi(y,\eta)\chi^{(1)}(x,\xi,r)\chi^{(2)}(x',\xi',r) \\
& \times \left(a_{ij}(\xi) - 2\sum_{k=1}^N \sigma_{ik}(\xi)\sigma_{kj}(\xi') + a_{ij}(\xi')\right) \\
& \times \partial_{x_i}\varrho^{(1)}_{\varepsilon,\delta}(x,y,\xi,\eta,r)\partial_{x'_j}\varrho^{(2)}_{\varepsilon,\delta}(x',y,\xi',\eta,r)dxd\xi dx'd\xi'dyd\eta \\
= {}& 2\sum_{i,j=1}^N \int \psi(y,\eta)\left(\int \chi^{(1)}(x,\xi,r)a_{ij}(\xi)\partial_{x_i}\varrho^{(1)}_{\varepsilon,\delta}(x,y,\xi,\eta,r)dxd\xi\right) \\
& \times \left(\int \chi^{(2)}(x',\xi',r)\partial_{x'_j}\varrho^{(2)}_{\varepsilon,\delta}(x',y,\xi',\eta,r)dx'd\xi'\right)dyd\eta \\
& - 4\sum_{i,j=1}^N \sum_{k=1}^N \int \psi(y,\eta) \\
& \times \left(\int \chi^{(1)}(x,\xi,r)\sigma_{ik}(\xi)\partial_{x_i}\varrho^{(1)}_{\varepsilon,\delta}(x,y,\xi,\eta,r)dxd\xi\right) \\
& \times \left(\int \chi^{(2)}(x',\xi',r)\sigma_{kj}(\xi')\partial_{x'_j}\varrho^{(2)}_{\varepsilon,\delta}(x',y,\xi',\eta,r)dx'd\xi'\right)dyd\eta \\
& + 2\sum_{i,j=1}^N \int \psi(y,\eta)\left(\int \chi^{(1)}(x,\xi,r)\partial_{x_i}\varrho^{(1)}_{\varepsilon,\delta}(x,y,\xi,\eta,r)dxd\xi\right) \\
& \times \left(\int \chi^{(2)}(x',\xi',r)a_{ij}(\xi')\partial_{x'_j}\varrho^{(2)}_{\varepsilon,\delta}(x',y,\xi',\eta,r)dx'd\xi'\right)dyd\eta.
\end{aligned}
$$

We now aim to take the limit $\delta \to 0$ in each of these three terms. Since the first and third terms are similar, here we only give the details for the first two.

Note that, since ψ has compact support, the integration in (x, y) and (x', y') is taking place over bounded sets. This together with $a_{ij} \in L^{\infty}_{loc}$ yields that, in the limit $\delta \to 0$, and in $L^2_{loc}(\mathbb{R})$

$$\int \chi^{(1)}(x, \xi, r) a_{ij}(\xi) \partial_{x_i} \varrho^{(1)}_{\varepsilon, \delta}(x, y, \xi, \eta, r) dx d\xi$$

$$= \left(\int \chi^{(1)}(x, \cdot, r) a_{ij}(\cdot) \partial_{x_i} \varrho^{s,(1)}_{\varepsilon}(x, y, \cdot, r) dx \right) * \varrho^v_{\delta}(\eta)$$

$$\to \int \chi^{(1)}(x, \eta, r) a_{ij}(\eta) \partial_{x_i} \varrho^{s,(1)}_{\varepsilon}(x, y, \eta, r) dx.$$

Since all functions are locally bounded, dominated convergence implies that, as $\delta \to 0$,

$$\sum_{i,j=1}^{N} \int_s^t \int \psi(y, \eta) \left(\int \chi^{(1)}(x, \xi, r) a_{ij}(\xi) \partial_{x_i} \varrho^{(1)}_{\varepsilon, \delta}(x, y, \xi, \eta, r) dx d\xi \right)$$

$$\times \left(\int \chi^{(2)}(x', \xi', r) \partial_{x'_j} \varrho^{(2)}_{\varepsilon, \delta}(x', y, \xi', \eta, r) dx' d\xi' \right) dy d\eta dr$$

$$\to \sum_{i,j=1}^{N} \int_s^t \int \psi(y, \eta) \chi^{(1)}(x, \eta, r) \chi^{(2)}(x', \eta, r) a_{ij}(\eta)$$

$$\times \partial_{x_i} \varrho^{s,(1)}_{\varepsilon}(x, y, \eta, r) \partial_{x'_j} \varrho^{s,(2)}_{\varepsilon}(x', y, \eta, r) dx' dx dy d\eta dr.$$

Similarly, as $\delta \to 0$,

$$\sum_{i,j,k=1}^{N} \int_s^t \int \psi(y, \eta) \chi^{(1)}(x, \xi, r) \sigma_{ik}(\xi) \partial_{x_i} \varrho^{(1)}_{\varepsilon, \delta}(x, y, \xi, \eta, r) dx d\xi$$

$$\times \int \chi^{(2)}(x', \xi', r) \sigma_{kj}(\xi') \partial_{x'_j} \varrho^{(2)}_{\varepsilon, \delta}(x', y, \xi', \eta, r) dx' d\xi' dy d\eta dr$$

$$\to \sum_{i,j,k=1}^{N} \int_s^t \int \psi(y, \eta) \chi^{(1)}(x, \eta, r) \sigma_{ik}(\eta) \partial_{x_i} \varrho^{s,(1)}_{\varepsilon, \delta}(x, y, \eta, r) dx$$

$$\times \int \chi^{(2)}(x', \eta, r) \sigma_{kj}(\eta) \partial_{x'_j} \varrho^{s,(2)}_{\varepsilon, \delta}(x', y, \eta, r) dx' dy d\eta dr$$

$$= \sum_{i,j=1}^{N} \int_s^t \int \psi(y, \eta) \chi^{(1)}(x, \eta, r) \chi^{(2)}(x', \eta, r) a_{ij}(\eta)$$

$$\times \partial_{x_i} \varrho^{s,(1)}_{\varepsilon, \delta}(x, y, \eta, r) \partial_{x'_j} \varrho^{s,(2)}_{\varepsilon}(x', y, \eta, r) dx dx' dy d\eta dr. \quad \square$$

The localization error. We study here the error terms appearing in Section 3 due to the localization in the (y, η) variables, that is $Err^{loc,(i)}_{\varepsilon, \psi, \delta}$, $i = 1, 2, 3, 4$. Since their analysis is similar, we concentrate on $Err^{loc,(1)}_{\varepsilon, \psi, \delta}$ and $Err^{loc,(3)}_{\varepsilon, \psi, \delta}$.

Recall

$$Err^{loc,(1)}_{\varepsilon, \psi, \delta}(r) = \int (sgn * \varrho^v_{\delta})(\eta) \left(\partial_{y_i y_j} \psi(y, \eta) \sum_{i,j=1}^{N} \int \chi^{(2)}(x', \xi', r) a_{ij}(\xi') \right.$$

$$\left. \times \varrho^{(2)}_{\varepsilon, \delta}(x', y, \xi', \eta, r) dx' d\xi' \right) dy d\eta$$

$$Err_{\varepsilon,\psi,\delta}^{loc,(3)}(r) = -\sum_{i,j=1}^{N}\int \partial_{y_i}\psi(y,\eta)\chi^{(1)}(x,\xi,r)\chi^{(2)}(x',\xi',r)a_{ij}(\xi')$$
$$\times \varrho^{(1)}(x,y,\xi,\eta,r)\partial_{x'_j}\varrho^{(2)}(x',y,\xi',\eta,r)dxd\xi dx'd\xi'dyd\eta$$

and set

$$Err_{\varepsilon,\psi}^{loc,(1)}(r) = \sum_{i,j=1}^{N}\int \mathrm{sgn}(\eta)\partial_{y_iy_j}\psi(y,\eta)\chi^{(2)}(x',\eta,r)a_{ij}(\eta)\varrho_\varepsilon^{s,(2)}(x',y,\eta,r)dx'dyd\eta$$

$$Err_{\varepsilon,\psi}^{loc,(3)}(r) = -\sum_{i,j=1}^{N}\int \partial_{y_i}\psi(y,\eta)\chi^{(1)}(x,\eta,r)\chi^{(2)}(x',\eta,r)a_{ij}(\eta)$$
$$\times \varrho^{s,(1)}(x,y,\eta,r)\partial_{x'_j}\varrho^{s,(2)}(x',y,\eta,r)dxdx'dyd\eta.$$

Lemma A.5. *For all* $s < t$, $s, t \in [0,T]$, *as* $\delta \to 0$, *we have*

$$\int_s^t Err_{\varepsilon,\psi,\delta}^{loc,(1)}(r)dr \to \int_s^t Err_{\varepsilon,\psi}^{loc,(1)}(r)dr \ \text{and} \ \int_s^t Err_{\varepsilon,\psi,\delta}^{loc,(3)}(r)dr \to \int_s^t Err_{\varepsilon,\psi}^{loc,(3)}(r)dr.$$

Proof. Noting that ψ has compact support and $a_{ij} \in L_{loc}^\infty$, the proof is a simple consequence of convergence of mollifications along Dirac families. □

Lemma A.6. *Let* ψ_R *be as in* (3.13). *Then, as* $R \to \infty$,

$$\int_s^t Err_{\varepsilon,\psi_R}^{loc,(1)}(r)dr \to 0 \quad \text{and} \quad \int_s^t Err_{\varepsilon,\psi_R}^{loc,(3)}(r)dr \to 0.$$

Proof. The convergence of $Err_{\varepsilon,\psi_R}^{loc,(1)}$ is a consequence of the dominated convergence theorem. We first observe that

$$|\mathrm{sgn}(\eta)\partial_{y_iy_j}\psi_R(y,\eta)\chi^{(2)}(x',\eta,r)a_{ij}(\eta)\varrho_\varepsilon^{s,(2)}(x',y,\eta,r)|$$
$$\le C|\chi^{(2)}|(x',\eta,r)a_{ij}(\eta)\varrho_\varepsilon^{s,(2)}(x',y,\eta,r) \le C\|a_{ij}\|_{L_{loc}^\infty}|\chi^{(2)}|(x',\eta,r)\varrho_\varepsilon^{s,(2)}(x',y,\eta,r)$$

and

$$\int_s^t\int |\chi^{(2)}|(x',\eta,r)\varrho_\varepsilon^{s,(2)}(x',y,\eta,r)dydx'd\eta dr = \int_s^t\int |\chi^{(2)}|(x',\eta,r)dx'd\eta dr$$
$$= \int_s^t\int |u^{(2)}|(x',r)dx'dt < \infty.$$

Moreover, as $R \to \infty$ and for all $(y,\eta) \in \mathbb{R}^{N+1}$ and $r \in [0,T]$,

$$\mathrm{sgn}(\eta)\partial_{y_iy_j}\psi_R(y,\eta)\chi^{(2)}(x',\eta,r)a_{ij}(\eta)\varrho_\varepsilon^{s,(2)}(x',y,\eta,r) \to 0,$$

and the claim now follows again from dominated convergence.

For the convergence of $Err_{\varepsilon,\psi_R}^{loc,(3)}$ we observe that

$$\partial_{y_i}\psi(y,\eta)\chi^{(1)}(x,\eta,r)\chi^{(2)}(x',\eta,r)a_{ij}(\eta)\varrho^{s,(1)}(x,y,\eta,r)\partial_{x'_j}\varrho^{s,(2)}(x',y,\eta,r)$$
$$\le C|\chi^{(1)}|(x,\eta,r)|\chi^{(2)}|(x',\eta,r)|a_{ij}|(\eta)|\varrho^{s,(1)}|(x,y,\eta,r)|\partial_{x'_j}\varrho^{s,(2)}|(x',y,\eta,r)$$
$$\le C\|a_{ij}\|_{L_{loc}^\infty}|\chi^{(1)}|(x,\eta,r)|\chi^{(2)}|(x',\eta,r)|\varrho^{s,(1)}|(x,y,\eta,r)|\partial_{x'_j}\varrho^{s,(2)}|(x',y,\eta,r)$$

and

$$
\int_s^t \int |\chi^{(1)}|(x, \eta, r)|\chi^{(2)}|(x', \eta, r)\varrho^{s,(1)}|(x, y, \eta, r)|\partial_{x'_j}\varrho^{s,(2)}|(x', y, \eta, r)dxdx'dyd\eta dr
$$

$$
\leq \int_s^t \int |\chi^{(2)}|(x', \eta, r)|\partial_{x'_j}\varrho^{s,(2)}|(x', y, \eta, r)dx'dyd\eta dr
$$

$$
\leq C \int_s^t \int |\chi^{(2)}|(x', \eta, r)dx'd\eta dr \leq C \int_s^t \int |u^{(2)}|(x', r)dx'dr < \infty.
$$

Since, as $R \to \infty$ and for all $(y, \eta) \in \mathbb{R}^{N+1}$ and $r \in [0, T]$,

$$
\partial_{y_i}\psi(y, \eta)\chi^{(1)}(x, \eta, r)\chi^{(2)}(x', \eta, r)a_{ij}(\eta)\varrho^{s,(1)}(x, y, \eta, r)\partial_{x'_j}\varrho^{s,(2)}(x', y, \eta, r) \to 0
$$

dominated convergence finishes the proof.   □

A.1. The hyperbolic errors

We first provide the following cancellation result.

Lemma A.7. *Let* $\varrho^{(1)}_{\varepsilon,\delta}$, $\varrho^{(2)}_{\varepsilon,\delta}$ *be as in* (2.6) *with* $z^{(1)}, z^{(2)} \in C_0([0, T]; \mathbb{R}^N)$ *and fix* $\varepsilon > 0$, $\psi \in C_c^\infty(\mathbb{R}^{N+1})$. *There exists a bounded function* $c_{\psi,\delta} : \mathbb{R} \to \mathbb{R}$ *so that* $\lim_{\delta \to 0} c_{\psi,\delta}(\xi') = 0$ *for all* ξ' *and*

$$
\int \left| \int \left(\psi(y, \eta)(\varrho^{(1)}_{\varepsilon,\delta}(x, y, \xi, \eta, t)\partial_{\xi'}\varrho^{(2)}_{\varepsilon,\delta}(x', y, \xi', \eta, t) \right. \right.
$$
$$
\left. \left. + \partial_\xi \varrho^{(1)}_{\varepsilon,\delta}(x, y, \xi, \eta, t)\varrho^{(2)}_{\varepsilon,\delta}(x', y, \xi', \eta, t)) \right) dyd\eta \right| dxd\xi
$$
$$
\leq \frac{C}{\varepsilon}|z^{(2)}_t - z^{(1)}_t| + \frac{c_{\psi,\delta}(\xi')}{\varepsilon} + (\|\partial_\eta \psi(\cdot, \cdot)\|_{C(\mathbb{R}^N)} * \varrho^v_\delta)(\xi').
$$

Proof. We have

$$
\varrho^{(1)}_{\varepsilon,\delta}(x, y, \xi, \eta, t)\partial_{\xi'}\varrho^{(2)}_{\varepsilon,\delta}(x', y, \xi', \eta, t) + \partial_\xi \varrho^{(1)}_{\varepsilon,\delta}(x, y, \xi, \eta, t)\varrho^{(2)}_{\varepsilon,\delta}(x', y, \xi', \eta, t)
$$
$$
= \varrho^{s,(1)}_\varepsilon(x, y, \xi, t)\varrho^v_\delta(\xi - \eta)\partial_{\xi'}[\varrho^{s,(2)}_\varepsilon(x', y, \xi', t)\varrho^v_\delta(\xi' - \eta)]
$$
$$
+ \partial_\xi[\varrho^{s,(1)}_\varepsilon(x, y, \xi, t)\varrho^v_\delta(\xi - \eta)]\varrho^{s,(2)}_\varepsilon(x', y, \xi', t)\varrho^v_\delta(\xi' - \eta)
$$
$$
= \varrho^v_\delta(\xi - \eta)\varrho^v_\delta(\xi' - \eta)\Big(\varrho^{s,(1)}_\varepsilon(x, y, \xi, t)\partial_{\xi'}\varrho^{s,(2)}_\varepsilon(x', y, \xi', t)
$$
$$
+ \partial_\xi \varrho^{s,(1)}_\varepsilon(x, y, \xi, t)\varrho^{s,(2)}_\varepsilon(x', y, \xi', t)\Big)
$$
$$
+ \varrho^{s,(1)}_\varepsilon(x, y, \xi, t)\varrho^{s,(2)}_\varepsilon(x', y, \xi', t)\big(\partial_\xi \varrho^v_\delta(\xi - \eta)\varrho^v_\delta(\xi' - \eta)
$$
$$
+ \varrho^v_\delta(\xi - \eta)\partial_{\xi'}\varrho^v_\delta(\xi' - \eta)\big).
$$

We first note that

$$
\int \psi(y, \eta)\varrho^{s,(1)}_\varepsilon(x, y, \xi, t)\varrho^{s,(2)}_\varepsilon(x', y, \xi', t)\big(\partial_\xi \varrho^v_\delta(\xi - \eta)\varrho^v_\delta(\xi' - \eta)
$$
$$
+ \varrho^v_\delta(\xi - \eta)\partial_{\xi'}\varrho^v_\delta(\xi' - \eta)\big)d\eta
$$
$$
= \int \partial_\eta \psi(y, \eta)\varrho^{s,(1)}_\varepsilon(x, y, \xi, t)\varrho^{s,(2)}_\varepsilon(x', y, \xi', t)\varrho^v_\delta(\xi - \eta)\varrho^v_\delta(\xi' - \eta)d\eta
$$

and, after integrating, we get

$$
\int \Big| \int \psi(y,\eta) \varrho_\varepsilon^{s,(1)}(x,y,\xi,t) \varrho_\varepsilon^{s,(2)}(x',y,\xi',t) \big(\partial_\xi \varrho_\delta^v(\xi-\eta) \varrho_\delta^v(\xi'-\eta)
$$
$$
+ \varrho_\delta^v(\xi-\eta) \partial_{\xi'} \varrho_\delta^v(\xi'-\eta) \big) dy d\eta \Big| dx d\xi
$$
$$
\leq \int |\partial_\eta \psi(y,\eta)| \varrho_\varepsilon^{s,(1)}(x,y,\xi,t) \varrho_\varepsilon^{s,(2)}(x',y,\xi',t) \varrho_\delta^v(\xi-\eta) \varrho_\delta^v(\xi'-\eta) dy d\eta dx d\xi
$$
$$
= \int |\partial_\eta \psi(y,\eta)| \varrho_\varepsilon^{s,(2)}(x',y,\xi',t) \varrho_\delta^v(\xi-\eta) \varrho_\delta^v(\xi'-\eta) dy d\eta d\xi
$$
$$
\leq \int \|\partial_\eta \psi(\cdot,\eta)\|_{C(\mathbb{R}^N)} \varrho_\delta^v(\xi-\eta) \varrho_\delta^v(\xi'-\eta) d\eta d\xi
$$
$$
= \int \|\partial_\eta \psi(\cdot,\eta)\|_{C(\mathbb{R}^N)} \varrho_\delta^v(\xi'-\eta) d\eta = (\|\partial_\eta \psi(\cdot,\cdot)\|_{C(\mathbb{R}^N)} * \varrho_\delta^v)(\xi').
$$

Next we observe that

$$
\partial_\xi \varrho_\varepsilon^{s,(i)}(x,y,\xi,t) = f'(\xi) z_t^{(i)} \partial_y \varrho_\varepsilon^{s,(i)}(x,y,\xi,t),
$$

and, hence,

$$
\varrho_\varepsilon^{s,(1)}(x,y,\xi,t) \partial_{\xi'} \varrho_\varepsilon^{s,(2)}(x',y,\xi',t) + \partial_\xi \varrho_\varepsilon^{s,(1)}(x,y,\xi,t) \varrho_\varepsilon^{s,(2)}(x',y,\xi',t)
$$
$$
= \varrho_\varepsilon^{s,(1)}(x,y,\xi,t) f'(\xi') z_t^{(2)} \partial_y \varrho_\varepsilon^{s,(2)}(x',y,\xi',t)
$$
$$
+ f'(\xi) z_t^{(1)} \partial_y \varrho_\varepsilon^{s,(1)}(x,y,\xi,t) \varrho_\varepsilon^{s,(2)}(x',y,\xi',t)
$$
$$
= \varrho_\varepsilon^{s,(1)}(x,y,\xi,t) f'(\xi) z_t^{(2)} \partial_y \varrho_\varepsilon^{s,(2)}(x',y,\xi',t)
$$
$$
+ f'(\xi) z_t^{(1)} \partial_y \varrho_\varepsilon^{s,(1)}(x,y,\xi,t) \varrho_\varepsilon^{s,(2)}(x',y,\xi',t)
$$
$$
- \varrho_\varepsilon^{s,(1)}(x,y,\xi,t) (f'(\xi) - f'(\xi')) z_t^{(2)} \partial_y \varrho_\varepsilon^{s,(2)}(x',y,\xi',t). \tag{A.1}
$$

For the last term we have the estimate

$$
\int \Big| \int \psi(y,\eta) \varrho_\delta^v(\xi-\eta) \varrho_\delta^v(\xi'-\eta) \varrho_\varepsilon^{s,(1)}(x,y,\xi,t) (f'(\xi)
$$
$$
- f'(\xi')) z_t^{(2)} \partial_y \varrho_\varepsilon^{s,(2)}(x',y,\xi',t) dy d\eta \Big| dx d\xi
$$
$$
\leq \frac{|z_t^{(2)}|}{\varepsilon} \iint \psi(y,\eta) \varrho_\delta^v(\xi'-\eta) \varrho_\delta^v(\xi-\eta) |f'(\xi) - f'(\xi')| d\eta d\xi =: \frac{c_{\psi,\delta}(\xi')}{\varepsilon}.
$$

Since f' is continuous, $c_{\psi,\delta}$ is bounded and $\lim_{\delta \to 0} c_{\psi,\delta} = 0$. The conclusion now follows from the following estimate

$$
\int \Big| \int \psi(y,\eta) \varrho_\delta^v(\xi-\eta) \varrho_\delta^v(\xi'-\eta) \big[\varrho_\varepsilon^{s,(1)}(x,y,\xi,t) f'(\xi) z_t^{(2)} \partial_y \varrho_\varepsilon^{s,(2)}(x',y,\xi',t)
$$
$$
+ f'(\xi) z_t^{(1)} \partial_y \varrho_\varepsilon^{s,(1)}(x,y,\xi,t) \varrho_\varepsilon^{s,(2)}(x',y,\xi',t) \big] dy d\eta \Big| dx d\xi
$$
$$
\leq |z_t^{(1)} - z_t^{(2)}| \int \psi(y,\eta) \varrho_\delta^v(\xi-\eta) \varrho_\delta^v(\xi'-\eta)
$$
$$
\times \varrho_\varepsilon^{s,(1)}(x,y,\xi,t) |\partial_y \varrho_\varepsilon^{s,(2)}|(x',y,\xi',t) |f'|(\xi) dy d\eta dx d\xi
$$
$$
= |z_t^{(1)} - z_t^{(2)}| \int \psi(y,\eta) \varrho_\delta^v(\xi-\eta) \varrho_\delta^v(\xi'-\eta) |\partial_y \varrho_\varepsilon^{s,(2)}|(x',y,\xi',t) |f'|(\xi) dy d\eta d\xi
$$
$$
\leq \frac{|z_t^{(1)} - z_t^{(2)}|}{\varepsilon} \int \|\psi(\cdot,\eta)\|_{C(\mathbb{R}^N)} \varrho_\delta^v(\xi-\eta) \varrho_\delta^v(\xi'-\eta) |f'|(\xi) d\eta d\xi
$$

$$= \frac{|z_t^{(1)} - z_t^{(2)}|}{\varepsilon} \int \|\psi(\cdot, \eta)\|_{C(\mathbb{R}^N)} \varrho_\delta^v(\xi - \eta)|f'|(\xi)d\xi \varrho_\delta^v(\xi' - \eta)d\eta$$

$$\leq \sup_{\text{supp}(\eta \to \|\psi(\cdot, \eta)\|_{C(\mathbb{R}^N)}) + [-\delta, \delta]} |f'|\varepsilon^{-1}|z_t^{(1)} - z_t^{(2)}|. \quad \square$$

For the estimate for $Err^{(i)}$, $i = 1, 2$ we recall that

$$Err_{\varepsilon, \psi, \delta}^{(1)}(r) = -\int \Big(\psi(y, \eta) \text{sgn}(\xi)(\partial_\xi \varrho_{\varepsilon, \delta}^{(1)}(x, y, \xi, \eta, r) \varrho_{\varepsilon, \delta}^{(2)}(x', y, \xi', \eta, r)$$
$$+ \varrho_{\varepsilon, \delta}^{(1)}(x, y, \xi, \eta, r) \partial_{\xi'} \varrho_{\varepsilon, \delta}^{(2)}(x', y, \xi', \eta, r)) q^{(2)}(x', \xi', r) \Big) dx d\xi dx' d\xi' dy d\eta.$$

Hence, using Lemma A.7, we find

$$|Err_{\varepsilon, \psi, \delta}^{(1)}(r)|$$
$$\leq \int q^{(2)}(x', \xi', r) \Big(\int \Big| \int \psi(y, \eta)(\partial_\xi \varrho_{\varepsilon, \delta}^{(1)}(x, y, \xi, \eta, r) \varrho_{\varepsilon, \delta}^{(2)}(x', y, \xi', \eta, r)$$
$$+ \varrho_{\varepsilon, \delta}^{(1)}(x, y, \xi, \eta, r) \partial_{\xi'} \varrho_{\varepsilon, \delta}^{(2)}(x', y, \xi', \eta, r)) dy d\eta \Big| dx d\xi \Big) dx' d\xi'$$
$$\leq \int q^{(2)}(x', \xi', r) \Big(\frac{C}{\varepsilon}|z_t^{(2)} - z_t^{(1)}| + \frac{c_{\psi, \delta}(\xi')}{\varepsilon} + (\|\partial_\eta \psi(\cdot, \cdot)\|_{C(\mathbb{R}^N)} * \varrho_\delta^v)(\xi')\Big) dx' d\xi',$$

and, letting $\delta \to 0$, we get

$$\lim_{\delta \to 0} \int_s^t |Err_{\varepsilon, \psi, \delta}^{(1)}(r)| dr$$
$$\leq \frac{C}{\varepsilon} \|z^{(2)} - z^{(1)}\|_{C([s,t]; \mathbb{R}^N)} + \int_s^t \int q^{(2)}(x', \xi', r) \|\partial_\eta \psi(\cdot, \xi')\|_{C(\mathbb{R}^N)} dx' d\xi' dr. \quad (A.2)$$

Now we present the proof of the error estimates for

$$\begin{cases} Err_{\varepsilon, \psi, \delta}^{(1,2)}(t) := 2 \int \psi(y, \eta) \Big(\chi^{(1)}(x, \xi, t) q^{(2)}(x', \xi', t) + \chi^{(2)}(x', \xi', t) q^{(1)}(x, \xi, t) \Big) \\ \quad \times \Big(\varrho_{\varepsilon, \delta}^{(1)}(x, y, \xi, \eta, t) \partial_{\xi'} \varrho_{\varepsilon, \delta}^{(2)}(x', y, \xi', \eta, t) \\ \quad + \partial_\xi \varrho_{\varepsilon, \delta}^{(1)}(x, y, \xi, \eta, t) \varrho_{\varepsilon, \delta}^{(2)}(x', y, \xi', \eta, t) \Big) dy d\eta dx' d\xi' dx d\xi, \end{cases}$$

and we note that

$$|Err_{\varepsilon, \psi, \delta}^{(1,2)}(t)|$$
$$\leq 2 \int q^{(2)}(x', \xi', t) \Big(\int \Big| \int (\psi(y, \eta)(\varrho_{\varepsilon, \delta}^{(1)}(x, y, \xi, \eta, t) \partial_{\xi'} \varrho_{\varepsilon, \delta}^{(2)}(x', y, \xi', \eta, t)$$
$$+ \partial_\xi \varrho_{\varepsilon, \delta}^{(1)}(x, y, \xi, \eta, t) \varrho_{\varepsilon, \delta}^{(2)}(x', y, \xi', \eta, t))) dy d\eta \Big| dx d\xi \Big) dx' d\xi'$$
$$+ 2 \int q^{(1)}(x, \xi, t) \Big(\int \Big| \int (\psi(y, \eta)(\varrho_{\varepsilon, \delta}^{(1)}(x, y, \xi, \eta, t) \partial_{\xi'} \varrho_{\varepsilon, \delta}^{(2)}(x', y, \xi', \eta, t)$$
$$+ \partial_\xi \varrho_{\varepsilon, \delta}^{(1)}(x, y, \xi, \eta, t) \varrho_{\varepsilon, \delta}^{(2)}(x', y, \xi', \eta, t))) dy d\eta \Big| dx' d\xi' \Big) dx d\xi. \quad (A.3)$$

Using Lemma A.7 in (A.3) yields

$$|Err^{(1,2)}(t)| \leq 2 \int q^{(2)}(x', \xi', t) \Big(\frac{C}{\varepsilon}|z_t^{(2)} - z_t^{(1)}| + \frac{c_{\psi, \delta}(\xi')}{\varepsilon}$$
$$+ (\|\partial_\eta \psi(\cdot, \cdot)\|_{C(\mathbb{R}^N)} * \varrho_\delta^v)(\xi') \Big) dx' d\xi'$$

$$+ 2 \int q^{(1)}(x, \xi, t) \left(\frac{C}{\varepsilon} |z_t^{(2)} - z_t^{(1)}| + \frac{c_{\psi,\delta}(\xi)}{\varepsilon} \right.$$
$$\left. + (\|\partial_\eta \psi(\cdot, \cdot)\|_{C(\mathbb{R}^N)} * \varrho_\delta^v)(\xi) \right) dx d\xi.$$

Letting $\delta \to 0$ and again using dominated convergence we get

$$\lim_{\delta \to 0} \int_s^t |Err^{(1,2)}(r)| dr \leq 2 \int_s^t \int q^{(2)}(x', \xi', r)$$
$$\times \left(\frac{C}{\varepsilon} |z_r^{(2)} - z_r^{(1)}| + \|\partial_\eta \psi(\cdot, \xi')\|_{C(\mathbb{R}^N)} \right) dx' d\xi' dr$$
$$+ 2 \int_s^t \int q^{(1)}(x, \xi, r)$$
$$\times \left(\frac{C}{\varepsilon} |z_r^{(2)} - z_r^{(1)}| + \|\partial_\eta \psi(\cdot, \xi)\|_{C(\mathbb{R}^N)} \right) dx d\xi dr.$$
$$\leq \frac{C}{\varepsilon} \|z^{(2)} - z^{(1)}\|_{C([s,t];\mathbb{R}^N)}$$
$$+ 2 \int_s^t \int q^{(2)}(x', \xi', r) \|\partial_\eta \psi(\cdot, \xi')\|_{C(\mathbb{R}^N)} dx' d\xi' dr$$
$$+ 2 \int_s^t \int q^{(1)}(x, \xi, r) \|\partial_\eta \psi(\cdot, \xi)\|_{C(\mathbb{R}^N)} dx d\xi dr. \tag{A.4}$$

Appendix B. The Regularity for the Fractional Heat semigroup

We recall without a proof [15, Lemma 9].

Lemma B.1. *For $\gamma > 0$, $\alpha \in (0, 1]$, let $B_\gamma = \gamma(-\Delta)^\alpha$ on $\mathbb{T}^N$ with corresponding semigroup $e^{-t B_\gamma}$. Then there exists a $C = C(N, n, m, \alpha, \beta)$ such that, for all $1 \leq m \leq n \leq \infty$ and $\beta \geq 0$,*

$$\left\| (-\Delta)^{\frac{\beta}{2}} e^{-t B_\gamma} \right\|_{L^m \to L^n} \leq \frac{C}{|\gamma t|^{\frac{N}{2\alpha} \left(\frac{1}{m} - \frac{1}{n} \right) + \frac{\beta}{2\alpha}}}.$$

Appendix C. The proof of Lemma 5.1

We present here the proof of Lemma 5.1.

Proof. We compute the Fourier transform $\hat{\phi}$ of ϕ. Using the elementary fact that $\int e^{-2\pi i z \cdot w} e^{-\frac{|w|^2}{\delta}} dw = \sqrt{\delta \pi} e^{-\delta \pi^2 |z|^2}$, we find

$$\hat{\phi}(z) = \int e^{-2\pi i z \cdot w} \phi(w) dw = \int e^{-2\pi i z \cdot w} e^{-\frac{|w|^2}{\delta}} \int e^{ib(\xi) \cdot w - \delta a(\xi)} f(\xi) d\xi dw$$
$$= \iint e^{-2\pi i \left(z - \frac{1}{2\pi} b(\xi) \right) \cdot w} e^{-\frac{|w|^2}{\delta}} dw e^{-\delta a(\xi)} f(\xi) d\xi$$
$$= \int \sqrt{\delta \pi} e^{-\left| \sqrt{\delta \pi} \left(\frac{1}{2\pi} b(\xi) - z \right) \right|^2} e^{-\delta a(\xi)} f(\xi) d\xi,$$

and, hence,

$$|\hat{\phi}(z)|^2 = \left| \int \sqrt{\delta\pi} e^{-\left|\sqrt{\delta\pi}\left(\frac{1}{2\pi}b(\xi)-z\right)\right|^2 - \delta a(\xi)} f(\xi)d\xi \right|^2$$

$$\le \delta\pi \int e^{-\left(\left|\sqrt{\delta\pi}\left(\frac{1}{2\pi}b(\xi)-z\right)\right|^2 + \delta a(\xi)\right)} d\xi \int e^{-\left|\sqrt{\delta\pi}\left(\frac{1}{2\pi}b(\xi)-z\right)\right|^2 - \delta a(\xi)} f^2(\xi)d\xi.$$

Next we use the assumption on a, b which enters in the following straightforward estimate:

$$\int e^{-\delta\left(\frac{1}{4}|b(\xi)-2\pi z|^2 + a(\xi)\right)} d\xi = \int_0^\infty \delta e^{-\delta\tau} \left|\left\{\xi : \frac{1}{4}|b(\xi)-2\pi z|^2 + a(\xi) < \tau\right\}\right| d\tau$$

$$\le \int_0^\infty \delta e^{-\delta\tau} |\{\xi : |b(\xi)-2\pi z|^2 + a(\xi) < 4\tau\}| d\tau$$

$$\le \int_0^\infty \delta e^{-\delta\tau} \iota(4\tau) d\tau \le \frac{1}{4} \int_0^\infty e^{-\frac{\tau}{4}} \iota\left(\frac{\tau}{\delta}\right) d\tau.$$

Hence,

$$|\hat{\phi}(z)|^2 \le \frac{\delta\pi}{4} \int_0^\infty e^{-\frac{\tau}{4}} \iota\left(\frac{\tau}{\delta}\right) d\tau \int e^{-\left|\sqrt{\delta\pi}\left(\frac{1}{2\pi}b(\xi)-z\right)\right|^2 - \delta a(\xi)} f^2(\xi)d\xi.$$

Integrating the above inequality in z and using that $\int e^{-|\sqrt{\delta\pi}\left(\frac{1}{2\pi}b(\xi)-z\right)|^2 - \delta a(\xi)} dz = \int e^{-\delta\pi^2|z|^2 - \delta a(\xi)} dz \le \frac{1}{\sqrt{\delta}}$ yields

$$\int |\hat{\phi}(z)|^2 dz \le \frac{\sqrt{\delta\pi}}{4} \int_0^\infty e^{-\frac{\tau}{4}} \iota\left(\frac{\tau}{\delta}\right) d\tau \|f\|_2^2. \quad \square$$

Appendix D. The definition of the kinetic solution for smooth driving signals

For the convenience of the reader we recall here the definition of the kinetic solution to (1.1) for smooth signals given in [10]. Note that this reference considers the special case of $z(t) = (t, \ldots, t)$ but the argument extends trivially to the case of $z \in C^1([0, T]; \mathbb{R}^N)$. For the notation β_{ik}^ψ recall (2.13).

Definition D.1 (*Definition 2.2 in [10]*). Assume that $z \in C^1([0, T]; \mathbb{R}^N)$. Then $u \in C([0, T]; L^1(\mathbb{R}^N)) \cap L^\infty(\mathbb{R}^N \times [0, T])$ for each $T > 0$ is a kinetic solution to (1.1) if

(i) for all $k \in \{1, \ldots, N\}$ and all nonnegative $\psi \in \mathcal{D}(\mathbb{R})$,

$$\sum_{i=1}^N \partial_{x_i} \beta_{ik}^\psi(u) \in L^2(\mathbb{R}^N \times [0, T]), \tag{D.1}$$

(ii) for any nonnegative $\psi_1, \psi_2 \in \mathcal{D}(\mathbb{R})$ and for all $k \in \{1, \ldots, N\}$

$$\sqrt{\psi_1(u(x,t))} \sum_{i=1}^N \partial_{x_i} \beta_{ik}^{\psi_2}(u(x,t)) = \sum_{i=1}^N \partial_{x_i} \beta_{ik}^{\psi_1 \psi_2}(u(x,t)), \tag{D.2}$$

(iii) there exists a non-negative bounded measure m on $\mathbb{R}^N \times \mathbb{R} \times [0, T]$ such that (2.3) holds in the sense of distributions, where n is a non-negative measure on $\mathbb{R}^N \times \mathbb{R} \times [0, T]$ such that, for any nonnegative $\psi \in \mathcal{D}(\mathbb{R})$,

$$\int \psi(\xi) dn(x, \xi, t) = \sum_{k=1}^{N} \left(\sum_{i=1}^{N} \partial_{x_i} \beta_{ik}^{\psi}(u(x, t)) \right)^2 . \tag{D.3}$$

References

[1] V. Barbu, G. Da Prato, M. Röckner, Existence and uniqueness of nonnegative solutions to the stochastic porous media equation, Indiana Univ. Math. J. 57 (1) (2008) 187–211.

[2] V. Barbu, G. Da Prato, M. Röckner, Existence of strong solutions for stochastic porous media equation under general monotonicity conditions, Ann. Probab. 37 (2) (2009) 428–452.

[3] V. Barbu, G. Da Prato, M. Röckner, Stochastic porous media equations and self-organized criticality, Comm. Math. Phys. 285 (3) (2009) 901–923.

[4] C. Bauzet, G. Vallet, P. Wittbold, The Cauchy problem for conservation laws with a multiplicative stochastic perturbation, J. Hyperbolic Differ. Equ. 9 (4) (2013) 661–709.

[5] C. Bauzet, G. Vallet, P. Wittbold, A degenerate parabolic-hyperbolic Cauchy problem with a stochastic force, J. Hyperbolic Differ. Equ. 12 (3) (2015) 501–533.

[6] F. Bouchut, L. Desvillettes, Averaging lemmas without time Fourier transform and application to discretized kinetic equations, Proc. Roy. Soc. Edinburgh Sect. A 129 (1) (1999) 19–36.

[7] G.-Q. Chen, Q. Ding, K.H. Karlsen, On nonlinear stochastic balance laws, Arch. Ration. Mech. Anal. 204 (3) (2012) 707–743.

[8] G.-Q. Chen, K.H. Karlsen, Quasilinear anisotropic degenerate parabolic equations with time-space dependent diffusion coefficients, Commun. Pure Appl. Anal. 4 (2) (2005) 241–266.

[9] G.-Q. Chen, K.H. Karlsen, L^1-framework for continuous dependence and error estimates for quasilinear anisotropic degenerate parabolic equations, Trans. Amer. Math. Soc. 358 (3) (2006) 937–963.

[10] 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.

[11] G.-Q. Chen, B. Perthame, Large-time behavior of periodic entropy solutions to anisotropic degenerate parabolic-hyperbolic equations, Proc. Amer. Math. Soc. 137 (9) (2009) 3003–3011.

[12] C.M. Dafermos, Long time behavior of periodic solutions to scalar conservation laws in several space dimensions, SIAM J. Math. Anal. 45 (4) (2013) 2064–2070.

[13] A. Debussche, M. Hofmanová, J. Vovelle, Degenerate parabolic stochastic partial differential equations: Quasilinear case. preprint, 2013, pp. 1–34.

[14] A. Debussche, J. Vovelle, Scalar conservation laws with stochastic forcing, J. Funct. Anal. 259 (4) (2010) 1014–1042.

[15] A. Debussche, J. Vovelle, Invariant measure of scalar first-order conservation laws with stochastic forcing, Probab. Theory Related Fields 163 (3–4) (2015) 575–611.

[16] W. E, K.M. Khanin, A.E. Mazel, Y.G. Sinaĭ, Invariant measures for Burgers equation with stochastic forcing, Ann. of Math. (2) 151 (3) (2000) 877–960.

[17] B. Engquist, W.E. , Large time behavior and homogenization of solutions of two-dimensional conservation laws, Comm. Pure Appl. Math. 46 (1) (1993) 1–26.

[18] J. Feng, D. Nualart, Stochastic scalar conservation laws, J. Funct. Anal. 255 (2) (2008) 313–373.

[19] F. Flandoli, Regularity Theory and Stochastic Flows for Parabolic SPDEs, in: Stochastics Monographs, vol. 9, Gordon and Breach Science Publishers, Yverdon, 1995.

[20] P.K. Friz, B. Gess, Stochastic scalar conservation laws driven by rough paths, Ann. Inst. H. Poincaré Anal. Non Linéaire 33 (4) (2016) 933–963.

[21] B. Gess, Random attractors for stochastic porous media equations perturbed by space–time linear multiplicative noise, Ann. Probab. 42 (2) (2014) 818–864.

[22] B. Gess, M. Hofmanova, Well-posedness and regularity for quasilinear degenerate parabolic-hyperbolic spde. arXiv:1611.03600, 2016.

[23] B. Gess, B. Perthame, P.E. Souganidis, Semi-discretization for stochastic scalar conservation laws with multiple rough fluxes, SIAM J. Numer. Anal. 54 (4) (2016) 2187–2209.

[24] B. Gess, P.E. Souganidis, Scalar conservation laws with multiple rough fluxes, Commun. Math. Sci. 13 (6) (2015) 1569–1597.
[25] B. Gess, P.E. Souganidis, Long-time behavior, invariant measures and regularizing effects for stochastic scalar conservation laws, Comm. Pure Appl. Math. (2016) arXiv:1411.3939 (in press).
[26] M. Hofmanová, Degenerate parabolic stochastic partial differential equations, Stochastic Process. Appl. 123 (12) (2013) 4294–4336.
[27] M. Hofmanová, Scalar conservation laws with rough flux and stochastic forcing, Stoch. Partial Differ. Equ. Anal. Comput. 4 (3) (2016) 635–690.
[28] H. Holden, N.H. Risebro, Conservation laws with a random source, Appl. Math. Optim. 36 (2) (1997) 229–241.
[29] N.V. Krylov, B.L. Rozovskiĭ, Stochastic evolution equations, in: Current Problems in Mathematics, Vol. 14, Akad. Nauk SSSR, Vsesoyuz. Inst. Nauchn. i Tekhn. Informatsii, Moscow, 1979, pp. 71–147. (in Russian) 256.
[30] P.D. Lax, Hyperbolic systems of conservation laws. II, Comm. Pure Appl. Math. 10 (1957) 537–566.
[31] P.-L. Lions, B. Perthame, P.E. Souganidis, Scalar conservation laws with rough (stochastic) fluxes, Stoch. Partial Differ. Equ. Anal. Comput. 1 (4) (2013) 664–686.
[32] P.-L. Lions, B. Perthame, P.E. Souganidis, Stochastic averaging lemmas for kinetic equations. Séminaire Laurent Schwartz - EDP et applications, 2013, pp. 1–17.
[33] P.-L. Lions, B. Perthame, P.E. Souganidis, Scalar conservation laws with rough (stochastic) fluxes; the spatially dependent case. arXiv:1403.4424, 2014, pp. 1–17.
[34] P.-L. Lions, P.E. Souganidis, Fully nonlinear stochastic partial differential equations, C. R. Acad. Sci. Paris Sér. I 326 (9) (1998) 1085–1092.
[35] P.-L. Lions, P.E. Souganidis, Fully nonlinear stochastic partial differential equations: non-smooth equations and applications, C. R. Acad. Sci. Paris Sér. I 327 (8) (1998) 735–741.
[36] P.-L. Lions, P.E. Souganidis, Fully nonlinear stochastic PDE with semilinear stochastic dependence, C. R. Acad. Sci. Paris Sér. I 331 (8) (2000) 617–624.
[37] P.-L. Lions, P.E. Souganidis, Uniqueness of weak solutions of fully nonlinear stochastic partial differential equations, C. R. Acad. Sci. Paris Sér. I 331 (10) (2000) 783–790.
[38] P.-L. Lions, P.E. Souganidis, Viscosity solutions of fully nonlinear stochastic partial differential equations, Sūrikaisekikenkyūsho Kōkyūroku (1287) (2002) 58–65. Viscosity solutions of differential equations and related topics (Japanese) (Kyoto, 2001).
[39] S.-E.A. Mohammed, T. Zhang, H. Zhao, The stable manifold theorem for semilinear stochastic evolution equations and stochastic partial differential equations, Mem. Amer. Math. Soc. 196 (917) (2008) vi+105.
[40] É. Pardoux, Equations aux dérivées partielles stochastiques non linéaires monotones (Ph.D. thesis), 1975.
[41] B. Perthame, Kinetic Formulation of Conservation Laws, in: Oxford Lecture Series in Mathematics and its Applications, vol. 21, Oxford University Press, Oxford, 2002.
[42] B. Perthame, P. Souganidis, Dissipative and entropy solutions to non-isotropic degenerate parabolic balance laws, Arch. Ration. Mech. Anal. 170 (4) (2003) 359–370.
[43] C. Prévôt, M. Röckner, A Concise Course on Stochastic Partial Differential Equations, in: Lecture Notes in Mathematics, vol. 1905, Springer, Berlin, 2007.
[44] E. Tadmor, T. Tao, Velocity averaging, kinetic formulations, and regularizing effects in quasi-linear PDEs, Comm. Pure Appl. Math. 60 (10) (2007) 1488–1521.
[45] A.I. Vol'rpert, S.I. Hudjaev, The Cauchy problem for second order quasilinear degenerate parabolic equations, Mat. Sb. (NS) 78 (120) (1969) 374–396.