Mathematical Models and Methods in Applied Sciences

Multi-dimensional scalar conservation laws with fluxes discontinuous in the unknown and the spatial variable*

Miroslav Bulíček

Mathematical Institute, Faculty of Mathematics and Physics, Charles University
Sokolovská 83, 186 75 Praha 8, Czech Republic
mbul8060@karlin.mff.cuni.cz

Piotr Gwiazda

Institute of Applied Mathematics, University of Warsaw
ul. Banacha 2, 02-097 Warsaw, Poland
pgwiazda@mimuw.edu.pl

Agnieszka Świerczewska-Gwiazda

Institute of Applied Mathematics, University of Warsaw
ul. Banacha 2, 02-097 Warsaw, Poland
aswiercz@mimuw.edu.pl

Received (Day Month Year)
Revised (Day Month Year)
Communicated by (xxxxxxxxxx)

The paper deals with a scalar conservation law in an arbitrary dimension d with a discontinuous flux. The flux is supposed to be a discontinuous function in the spatial variable x and in an unknown function u. Under some additional hypothesis on the structure of possible discontinuities, we formulate an appropriate notion of entropy solution and establish its existence and uniqueness. The framework for proving the existence and uniqueness of entropy weak solutions is provided by the studies on entropy measure-valued solutions and may be viewed as a corollary of the uniqueness theorem for entropy measure-valued solutions.

Keywords: scalar hyperbolic conservation laws, measure-valued solutions, entropy weak solutions, averaged contraction property, Young measures, discontinuous flux, heterogenous flux

AMS Subject Classification: 35L65, 35R05

*Miroslav Bulíček thanks to the project GAČR 201/09/0917. Piotr Gwiazda thanks to National Science Centre (project no. 6085/B/H03/2011/40). Agnieszka Świerczewska-Gwiazda thanks to Jindřich Nečas Center for Mathematical Modeling, the project LC06052 financed by MŠMT, and National Science Centre (project no. 6085/B/H03/2011/40).

1. Introduction

We focus on the Cauchy problem for a scalar hyperbolic conservation law

$$\begin{aligned} u_{,t} + \operatorname{div} \mathbf{F}(x,u) &= 0 \quad \text{in } \mathbb{R}^{d+1}_+, \\ u(0,\cdot) &= u_0 \quad \text{in } \mathbb{R}^d, \end{aligned} \tag{1.1}$$

where $\mathbb{R}^{d+1}_+ := (0,\infty)\times\mathbb{R}^d$, d denotes an arbitrary spatial dimension, $u : \mathbb{R}^{d+1}_+ \to \mathbb{R}$ is an unknown and $\mathbf{F} : \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}^d$ is a given flux of the quantity u. In addition, we assume that u vanishes as $|x| \to \infty$. Our main goal in the paper is to identify a class of fluxes $\mathbf{F}$ for which one can develop "well-posedness" of the problem (1.1). By well-posedness we mean that we want to find a proper notion of (entropy) solution to (1.1) for which we are able to prove the existence, uniqueness and stability of the solution with respect to data in an a priori chosen class. In particular, our primary motivation is that such class is equivalent to Kružkov entropy solution in case that $\mathbf{F}$ is a sufficiently smooth function. For brevity, we recall that for smooth $\mathbf{F}$ a weak solution to (1.1) is called the Kružkov entropy solution if it satisfies for all $k \in \mathbb{R}$ the following entropy inequality in the distributional sense in $\mathbb{R}^{d+1}_+$

$$\begin{aligned} |u-k|_{,t} + \operatorname{div}\left(\operatorname{sgn}(u-k)(\mathbf{F}(x,u) - \mathbf{F}(x,k))\right) \\ + \operatorname{sgn}(u-k) \operatorname{div} \mathbf{F}(x,k) \le 0. \end{aligned} \tag{1.2}$$

To generalize this notion to a class of fluxes $\mathbf{F}(x,u)$ discontinuous with respect both to x and to u we shall follow several recent results and combine them in a proper way to develop a unified theory. An important approach to problems with $x-$discontinuous fluxes appears for $d = 1$ in the Ref. 9 and later in Ref. 7. The authors generalized the entropy inequality (1.2) in a way that instead of a constant k they considered a stationary solution, namely a function $k(x)$ solving the equation

$$\partial_x \mathbf{F}(x,k(x)) = 0. \tag{1.3}$$

This idea of extending the definition of Kružkov solutions to a discontinuous case consisted in introducing adapted entropies $E(x,u) = |u - k(x)|$. For such choice of entropies we observe that the last term in (1.2), which is not well-defined in case of non-smooth $\mathbf{F}$ vanishes. For a sufficiently large class of k's satisfying (1.3) the uniqueness of solutions to (1.1)–(1.2) can be proved. We discuss here the assumptions under which one is able to find this rich class of k's fulfilling (1.3). The Ref. 9 concerns the case of injective fluxes. In Ref. 7 the authors assumed that there exist continuous functions f, g; $f(u) \to \infty$ as $|u| \to \infty$ such that

$$\mathbf{F}(x,u) \qquad \text{is Carathéodory,} \tag{1.4}$$

$$f(u) \le |\mathbf{F}(x,u)| \le g(u), \tag{1.5}$$

$$\mathbf{F}(x,u) \qquad \text{is for a.a. } x \text{ one to one locally Lipschitz .} \tag{1.6}$$

Alternatively, instead of (1.6) the authors assumed the following

$$\begin{aligned}&\text{There is a function } u_M(x):\mathbb{R}\to\mathbb{R} \text{ such that for a.a. } x\in\mathbb{R}\\ &F(x,\cdot) \text{ is a one to one locally Lipschitz function from } [-\infty,u_M(x)]\\ &\text{and } [u_M(x),\infty] \text{ to } [0,\infty] \text{ that satisfies } F(x,u_M(x))=0.\end{aligned} \tag{1.7}$$

Under such conditions the authors were able to prove uniqueness (in a given class) of the solution to (1.1), that is equivalent to the Kružkov entropy solution in case that F is smooth. This equivalence, as well as an existence of such a solution was proved in Ref. 18. In case of (1.7) there are multiple solutions to (1.3) and then the choice of k's needs to be restricted only to some of them.

Later Panov[33] generalized the method developed in Ref. 7 in the following way. He assumed that F is of the form

$$\mathsf{F}(x,u)=\mathsf{G}(\theta(x,u)), \tag{1.8}$$

where $\mathsf{G}\in\mathcal{C}(\mathbb{R};\mathbb{R}^d)$ and $\theta(x,u):\mathbb{R}^d\times\mathbb{R}\to\mathbb{R}$ is a Carathéodory function that is for almost all x strictly increasing with respect to u and for which there exist continuous functions f and g fulfilling $f(u)\to\infty$ as $|u|\to\infty$ such that

$$f(u)\le|\theta(x,u)|\le g(u). \tag{1.9}$$

In this setting it is shown in Ref. 33 that it is natural to rewrite the notion of Kružkov solution in the following way: Let $\eta(x,v)$ be the inverse to θ, i.e., $\theta(x,\eta(x,v))=v$. Then assuming that G, θ, η and u are smooth one easily shows that u solves (1.1)–(1.2) if and only if $u(t,x):=\eta(x,v(t,x))$ (that is again smooth) solves for all $k\in\mathbb{R}$ the following system in the sense of distribution

$$\eta(x,v)_{,t}+\operatorname{div}\mathsf{G}(v)=0 \qquad \text{in } \mathbb{R}^{d+1}_+, \tag{1.10}$$

$$\eta(x,v(x,0))=u_0(x) \qquad \text{in } \mathbb{R}^d, \tag{1.11}$$

$$|\eta(x,v)-\eta(x,k)|_{,t}+\operatorname{div}(\operatorname{sgn}(v-k)(\mathsf{G}(v)-\mathsf{G}(k)))\le 0, \qquad \text{in } \mathbb{R}^{d+1}_+. \tag{1.12}$$

From Ref. 33 and 18 one observes that solving (1.10)–(1.12) is equivalent with solving (1.1)–(1.2) provided that $\mathsf{F}=\mathsf{G}(\theta(x,u))$ with sufficiently smooth G and θ. In addition, it is evident that for introducing a notion of a weak solution satisfying (1.12) we do not need to require any smoothness of θ with respect to x, which is the case if one assumes (1.2). Thus, the system (1.10)–(1.12) seems to be a natural extension of the concept of Kružkov entropy solution and the existence and uniqueness of such a solution is established in Ref. 33 provided that G is sufficiently smooth (Hölder continuous)[a].

Finally, in Ref. 10 the authors considered fluxes being independent of x but discontinuous with respect to u. They showed that for jump continuous fluxes[b], one can find nondecreasing U such that $\mathsf{F}\circ U$ is continuous and define a notion of entropy

[a]In Ref. 33 such theorem is stated in any d but the rigorous proof is provided for $d=1$

[b]A function F is called jump continuous if for all $s\in\mathbb{R}$ there exists a finite $\lim_{t\to s_\pm}\mathsf{F}(s)$ and there is at most countable set where F is not continuous.

weak solution that is equivalent to the Kružkov entropy weak solution for smooth fluxes in the following way: A function v is an entropy weak solution to (1.10)–(1.12) if it solves for all $k \in \mathbb{R}$ the following system in the sense of distribution

$$U(v)_{,t} + \operatorname{div} \mathbf{F}(U(v)) = 0 \qquad \text{in } \mathbb{R}^{d+1}_+, \quad (1.13)$$

$$U(v(x,0)) = u_0(x) \quad \text{in } \mathbb{R}^d, \quad (1.14)$$

$$|U(v) - U(k)|_{,t} + \operatorname{div}(\operatorname{sgn}(v-k)(\mathbf{F}(U(v)) - \mathbf{F}(U(k)))) \le 0, \qquad \text{in } \mathbb{R}^{d+1}_+. \quad (1.15)$$

The authors showed in Ref. 10 that for jump continuous $\mathbf{F}$ that is additionally Hölder continuous at zero there exists just one entropy weak solution to (1.13)–(1.15).

Thus, our main goal in the paper and also our strategy is to combine the methods developed in Ref. 7, 10 and 33 and to develop a theory that covers both possible discontinuities of the flux, i.e. discontinuities with respect to u and x.

Moreover, to simplify the presentation we frequently use the following notations: Small letters in italics always denote the scalar functions mapping $\mathbb{R}^{d+1}_+$, $\mathbb{R}^d$ respectively on $\mathbb{R}$, i.e., $v(t,x) : \mathbb{R}^{d+1}_+ \to \mathbb{R}$ and $v(x) : \mathbb{R}^d \to \mathbb{R}$. The real function of one real variable will be always denoted by capital letter, i.e., $P : \mathbb{R} \to \mathbb{R}$. On the other hand vector-valued function of one real variable is denoted by capital bold letter, i.e., $\mathbf{Q} : \mathbb{R} \to \mathbb{R}^d$. Similarly, vector-valued function of (t,x) or x respectively will be denoted by small bold letter, i.e., $\boldsymbol{v}(t,x) : \mathbb{R}^{d+1}_+ \to \mathbb{R}^d$ and $\boldsymbol{v}(x) : \mathbb{R}^d \to \mathbb{R}^d$.

1.1. *Assumptions on the flux and admissible parametrization*

In the whole paper we assume that there are $\mathbf{G} : \mathbb{R} \to \mathbb{R}^d$ and $\theta : \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}$ such that $\mathbf{F}(x,u) = \mathbf{G}(\theta(x,u))$. Moreover, we assume that

- **(A1)** $\mathbf{G}(v)$ is jump continuous
- **(A2)** θ is a Carathéodory strictly increasing function such that $\theta(x,0) = 0$ and there exists a Carathéodory function $\eta(x,v)$ such that $\theta(x, \eta(x,v)) = v$ for all $v \in \mathbb{R}$ and a.a. $x \in \mathbb{R}^d$.
- **(A3)** there exist continuous functions h_1 and h_2 such that for all $x \in \mathbb{R}^d$

$$h_1(u) \le |\theta(x,u)| \le h_2(u),$$

  and such that for all $R > 0$ there exists C_R so that $h_2(u) \le C_R h_1(u)$ for all $|u| \le R$; in addition we require that $\lim_{|u|\to\infty} h_1(u) = \infty$.

Note that it directly follows from **(A3)** that also $\eta(x,v)$ is bounded by some function dependent only on v and not on x, namely

- **(A3*)** there exists h_3 such that for all $x \in \mathbb{R}^d$

$$|\eta(x,v)| \le h_3(v)$$

Already Kružkov[29] pointed out that if $\mathbf{F}$ is not sufficiently regular at zero one can get non-uniqueness of entropy solution. On the other hand it was shown in Ref. 37 for continuous fluxes and in Ref. 10 that only appropriate Hölder continuity at zero is required for obtaining uniqueness of solution provided[c] that the initial data u_0 belongs to $L^1(\mathbb{R}^d)$. Thus, having this in mind it is natural to introduce a new assumption on a flux that will guarantee the uniqueness of a solution. Therefore, in addition to **(A1)**–**(A3)** we assume that for given G and θ

(A4) there exists $1 \le p \le \frac{d}{d-1}$ and constants $R_\infty > 0$ and $C_\infty > 0$ such that for all $x \in \mathbb{R}^d \setminus B_{R_\infty}(0)$

$$|\mathsf{G}(s)|^p \le C_\infty |\eta(x,s)|.$$

Finally, we introduce an admissible parametrization of (possibly) discontinuous G. We denote by $z_k \in \mathbb{R}$, $k \in \mathbb{N}$ the points such that $\lim_{s\to(z_k)_+} \mathsf{G}(s) \neq \lim_{s\to(z_k)_-} \mathsf{G}(s)$. Recall that the set of such points is countable[d]. Next we construct a multi-valued mapping $\mathcal{G}$ by filling the jumps of G with intervals connecting $\lim_{s\to(z_k)_+} \mathsf{G}(s)$ and $\lim_{s\to(z_k)_-} \mathsf{G}(s)$ and in what follows we identify G with $\mathcal{G}$ for simplicity. Then, we say that a couple (A, U) is an *admissible parametrization* of G if it satisfies the following conditions:

$$\begin{aligned}
&\bullet \text{ the function } U \in \mathcal{C}(\mathbb{R}) \text{ is nondecreasing and } \lim_{s\to\pm\infty} U(s) = \pm\infty,\\
&\bullet \text{ let } \alpha_k := \inf_{\alpha;\, U(\alpha)=z_k} \alpha, \qquad \beta_k := \sup_{\beta;\, U(\beta)=z_k} \beta,\\
&\quad \text{then for all } k \in \mathbb{N} \text{ there holds } \alpha_k < \beta_k < \alpha_{k+1},\\
&\bullet \text{ the function } U \text{ is constant on } [\alpha_k, \beta_k] \text{ and strictly increasing}\\
&\quad \text{on } (\beta_k, \alpha_{k+1}) \text{ for all } k \in \mathbb{N},\\
&\bullet \text{ the function } \mathsf{A} \in \mathcal{C}(\mathbb{R})^d \text{ satisfies } \mathsf{A}(s) \in \mathsf{G}(U(s)),\\
&\bullet \text{ the function } \mathsf{A} \text{ is linear on } [\alpha_k, \beta_k] \text{ for all } k \in \mathbb{N}.
\end{aligned} \tag{1.16}$$

For every jump continuous function G there exists such an admissible parametrization, what we discuss in more detail at the beginning of Section 2.

1.2. *Definition of entropy weak solutions and main result*

Definition 1.1. Let $\mathsf{F} = \mathsf{G} \circ \theta$ with G and θ satisfying **(A1)**–**(A3)** and let $u_0 \in L^1(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$. We say that $u \in L^\infty(0,\infty; L^1(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d))$ is an entropy weak solution to (1.1) related to (G, θ) and u_0 for an admissible parametrization (A, U) of G if there exists a function $g \in L^\infty(\mathbb{R}^{d+1}_+)$ such that

$$\eta(x, U(g(t,x))) = u(t,x), \quad \mathsf{A}(g(t,x)) \in \mathsf{G}(\theta(x, u(t,x))) \qquad \text{a.e. in } \mathbb{R}^{d+1}_+, \tag{1.17}$$

$$\liminf_{t\to 0} \int_K |u(t,x) - u_0(x)|\, dx = 0, \qquad \text{for any compact } K \subset \mathbb{R}^d, \tag{1.18}$$

[c]Note that if initial data are assumed to be only bounded, then the requirement on Hölder continuity of the flux at zero is not sufficient and one has to assume the Hölder continuity of the flux at each point.

[d]It follows from the fact that G is jump continuous function

and for all nonnegative $\psi \in \mathcal{D}(\mathbb{R}^{d+1}_+)$ and arbitrary $k \in \mathbb{R} \setminus \bigcup_{l\in\mathbb{N}} (\alpha_l, \beta_l)$ there holds

$$\begin{aligned}\int_{\mathbb{R}^{d+1}_+} &|\eta(x, U(g(t,x))) - \eta(x, U(k))| \psi_{,t}(t,x)\, dx\, dt \\ &+ \int_{\mathbb{R}^{d+1}_+} (\operatorname{sgn}(g(t,x)-k)(\mathbf{A}(g(t,x)) - \mathbf{A}(k))) \cdot \nabla\psi(t,x)\, dx\, dt \geq 0.\end{aligned} \tag{1.19}$$

The numbers α_l, β_l, $l \in \mathbb{N}$ are defined in (1.16).

Remark 1.1. Any entropy weak solution is the sense of Definition 1.1 is a weak solution to (1.1). Indeed, since $g \in L^\infty$ we may take $k := \pm\|g\|_\infty$ in (1.19) (or possibly we increase/decrease the value of k such that U is strictly increasing in k) and by using the strict monotonicity of η (guaranteed by **(A2)**) and the monotonicity of U we conclude that

$$u_{,t} + \operatorname{div} \mathbf{A}(g) = 0, \qquad \text{in the sense of distribution in } \mathbb{R}^{d+1}_+, \tag{1.20}$$

which is exactly $(1.1)_1$. Next, we can use the fact that by functions $|u - \cdot|$ one can generate any convex function and therefore it is a direct consequence of (1.19) that (see Ref. 10 for details) for all smooth convex $\tilde{E}$, such that $\tilde{E}$ is linear on (α_k, β_k) for all $\mathbb{N}$, where α_k and β_k are introduced in (1.16), there holds

$$Q_u(x,g)_{,t} + \operatorname{div} \mathbf{Q_A} \leq 0, \qquad \text{in sense of distribution in } \mathbb{R}^{d+1}_+ \tag{1.21}$$

with Q_u and $\mathbf{Q_A}$ given by

$$\partial_s Q'_u(x,s) = \partial_s \eta(x, U(s)) \tilde{E}'(s), \qquad \partial_s \mathbf{Q_A}(s) = \partial_s \mathbf{A}(s) \partial_s \tilde{E}(s). \tag{1.22}$$

Having a notion of an entropy weak solution to (1.1) we finally focus on the existence and uniqueness theorem. The main result of the paper is the following.

Theorem 1.1. *Let $\mathbf{F} = \mathbf{G} \circ \theta$ with $\mathbf{G}$ and θ satisfying **(A1)**–**(A4)** and let $u_0 \in L^1(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$. Then there exists a unique entropy weak solution to* (1.1) *related to* $(\mathbf{G}, \theta)$ *and u_0 in the sense of Definition 1.1.*

Remark 1.2. Note that an entropy weak solution is independent of the choice of parametrization U unless this parametrization depends on x. Another observation is that functions $\mathbf{G}$ and θ are not necessarily given uniquely and thus there arises a question whether different choice of $\mathbf{G}$ and θ may lead to different entropy solutions, namely in the case when $\mathbf{G}_2(\theta_2(x,u)) = \mathbf{G}_1(\theta_1(x,u))$, which was shown in 33. In the Appendix B we discuss sufficient conditions for independence of the choice of $\mathbf{G}$ and θ, see Theorem B.1.

Here, we briefly discuss the optimality of the assumptions of Theorem 1.1 and the relevance to other results. The assumption **(A1)** is inspired by Ref. 10 and covers also the case when $\mathbf{G}$ has infinitely many jumps. The assumptions **(A2)** and **(A3)** were introduced by Panov[33] and define the class of fluxes $\mathbf{F}$ for which one can relatively easily introduce a change of variables which eliminates a direct dependence of the flux on x. Note that it is in fact restrictive, e.g., it is clear that if **(A1)**–**(A2)** are satisfied then necessarily for fixed x the $\mathrm{Im}\,(F(x,u))$ does not depend on x. For some physical phenomena such an assumption may not be valid (e.g. flows in porous media, sedimentation, cf. Ref. 12, 14, 21, 27). Finally, the last assumption **(A4)** combines the requirements on the behavior of $\mathbf{F}(x,\cdot)$ near zero and behavior of $\mathbf{F}(\cdot,u)$ near infinity. Indeed, one could alternatively assume that $\mathbf{G}$ is $\alpha-$Hölder continuous in zero with an appropriate exponent α and that $|\theta(x,u)| \le C_\infty |u|$ for $|x| \to \infty$ and/or to introduce other assumptions on the behavior of $\mathbf{F}$ near ∞ for $u = 0$ that would lead to the uniqueness of the solution.

Besides the results described above we shall mention various other approaches to scalar conservation laws including discontinuities of fluxes. According to our knowledge the case of discontinuity in both variables x and u has not been considered.

The motivation for considering the discontinuity in the flux function with respect to u comes from the implicit constitutive theory, namely implicit relations between the flux and the unknown, see e.g. Ref. 35 and also Ref. 10. The strategy is the following: The discontinuous flux is identified with a continuous curve consisting of the graph of the function $\mathbf{F}$ and the intervals filling the jumps.

The framework of fluxes discontinuous in u was also studied in Ref. 4, 17 and 30 for the case of fluxes independent of x in a multi-dimensional case. We shall recall briefly the approach of Carrillo[16,17]. He considered a problem in a bounded domain

$$\begin{aligned} u_t + \mathrm{div}\,\Phi(u) \ni f \quad & \text{in } (0,T)\times\Omega \\ u(0) = u_0 \quad & \text{in } \Omega \end{aligned} \tag{1.23}$$

with a piecewise continuous flux function having a finite number of jumps. The author further provides an appropriate change of variables which leads to a "new" problem

$$\begin{aligned} g(v)_t + \mathrm{div}\,\Psi(u) \ni f \quad & \text{in } (0,T)\times\Omega, \\ g(v(0)) = u_0 \quad & \text{in } \Omega. \end{aligned} \tag{1.24}$$

For the details we refer to Ref. 17. The proof bases upon the comparison principle and the entropy inequality involving a version of semi Kružkov entropies, namely $E(v,k) = (g(v) - g(k))^+$. It is worth to notice that the change of variables introduced by Carrillo is a prototype of the change of variables used in Ref. 10.

Most of the studies on fluxes discontinuous in x confine to one-dimensional case. We recall here an overview paper of Risebro[36] concerned mostly with a front tracking method, see also Ref. 25, 28. The author describes various motivations for considering such problems. The multidimensional problem was considered in Ref. 5, 26 and 31.

In the works basing upon the classical Kružkov entropies the problem of giving sense to the last term in the entropy inequality (1.2) arises. Some approaches require introducing an interface entropy condition, see e.g. Ref. 1, which was mostly entailed with the assumptions on the total variation of the flux function. Nevertheless, one can show that the entropy solution with a total variation not necessarily bounded, admits strong boundary traces, cf. Ref 32.

Another interesting approaches to $x-$discontinuous fluxes are the strong precompactness result based on H-measures techniques by Panov[34] and the approach of kinetic formulation for conservation laws presented by Bachmann and Vovelle[8]. In the first case the method requires that the flux is non-degenerate (for almost all x, for all $\xi \in \mathbb{R}^d, \xi = 0$ the functions $\lambda \mapsto \xi F(x, \lambda)$ are not constant on non-degenerate intervals).

Among the variety of entropy conditions for the problems with $x-$discontinuous fluxes an interesting question arises which of them have an appropriate physical meaning. An important observation is that the same formal equation may admit different solutions (Ref. 2, 6, 11, 14, 24). However these differences refer to different dissipative mechanisms, cf. in particular Ref. 12, 14, 13 and 15.

Nevertheless, the classical framework for answering the question of appropriate physical meaning is derivation of target equations from primitive equations. With this strategy in mind, we shall be particularly interested in the entropy conditions of Audusse and Perthame (and hence also the ones of Panov), due to the rigorous derivation as a hydrodynamic limit from the particle system, see Ref. 18.

We organize the paper as follows: Section 2 concerns the existence of entropy measure-valued solutions. We define the entropy measure-valued solutions (Definition 2.1) and state the theorems on existence and uniqueness of solutions (Theorems 2.1 and 2.2). The crucial tool for these results is the so-called averaged contraction property formulated in Lemma 2.1. The meaning of uniqueness of entropy measure-valued solutions requires the discussion on its appropriate interpretation. Two entropy measure-valued solutions are meant to be unique in the sense specified by (2.22)–(2.23), more precisely they are unique up to the level sets of the function U. The results on measure-valued solutions are essential for passing to entropy weak solutions. The main difference in comparison to the result on $x-$independent fluxes (see Ref. 10) concentrates in showing the well-posedness of entropy measure-valued solutions and hence these steps are described in details. The passage from the level of measure-valued to weak solutions follows the same lines. We comment on this issue in Section 3, where the reader can find the sketch of the proof of Theorem 1.1.

Finally, Appendix A contains the discussion on relations between various notions of solutions and provides their equivalence for sufficiently smooth fluxes $\mathbf{F}$. In Appendix B we discuss the dependence of solution on different choices of functions $\mathbf{G}$ and θ.

2. Entropy measure-valued solutions

In this section we shall be concerned with entropy measure-valued solutions to (1.1) in case of non-smooth $\mathbf{F}$. More precisely, we focus here on fluxes $\mathbf{F}$ that satisfy (1.8) with $\mathbf{G}, \theta$ satisfying **(A1)**–**(A4)**. By $\mathcal{M}(\mathbb{R})$ we mean the space of bounded Radon measures and by Prob($\mathbb{R}$) the space of probability measures. As a Young measure ν we mean a weak* measurable map $\nu : \mathbb{R}^{d+1}_+ \to \mathcal{M}(\mathbb{R})$ and such that $\nu_{(t,x)} \geq 0, \|\nu_{(t,x)}\|_{\mathcal{M}(\mathbb{R})} \leq 1$ for a.a. $(t,x) \in \mathbb{R}^{d+1}_+$. Any bounded sequence $u^n : \mathbb{R}^{d+1}_+ \to \mathbb{R}$ generates a Young measure, which is a probability measure. By $L^\infty_w(\mathbb{R}^{d+1}_+; \mathcal{M}(\mathbb{R}))$ we understand the space of weak* measurable maps $\nu : \mathbb{R}^{d+1}_+ \to \mathcal{M}(\mathbb{R})$ that are essentially bounded.

We focus on properties of an admissible parametrization (A, U) of G, i.e., on functions satisfying (1.16). First note, that if U and A satisfy (1.16) then necessarily

$$\mathsf{A}(s) = \begin{cases} \mathsf{G}(U(s)) & \text{for } s \in (\beta_k, \alpha_{k+1}), \\ \dfrac{\mathsf{G}_+(z_k) - \mathsf{G}_-(z_k)}{\beta_k - \alpha_k}(s - \alpha_k) + \mathsf{G}_-(z_k) & \text{for } s \in [\alpha_k, \beta_k], \end{cases} \tag{2.1}$$

where $\mathsf{G}_\pm(z_k) := \lim_{s \to (z_k)_\pm} \mathsf{G}(s)$. It is also important to mention here, that there are many ways how to construct U and A such that (1.16) holds. Since according to Theorem 2.1 the solution will not depend on the choice of U, we introduce here one possible choice that shall also be used in the proof of the existence theorem. First, we define

$$\alpha_k = z_k + \sum_{n;\, z_n < z_k} \frac{1}{n^2} \qquad \text{and} \qquad \beta_k = z_k + \sum_{n; z_n \leq z_k} \frac{1}{n^2}$$

and define $\bar{U}$ as

$$\bar{U}(g) = \begin{cases} g - \displaystyle\sum_{k: z_k < g} \frac{1}{k^2} & \text{if } g \notin (\alpha_m, \beta_m) \text{ for all } m \in \mathbb{N}, \\ z_k - \displaystyle\sum_{n: z_n < z_k} \frac{1}{n^2} & \text{otherwise.} \end{cases} \tag{2.2}$$

To normalize U to be zero at zero we set

$$U(s) := \bar{U}(s) - \bar{U}(0). \tag{2.3}$$

It is an obvious observation that the function U is continuous, nondecreasing and $\operatorname{Im} U(\mathbb{R}) = \mathbb{R}$. Moreover, having U we can immediately find A such that (1.16) holds.

Next, we introduce a notion of a local entropy measure-valued solution, i.e., we do not prescribe any initial data and any behavior for $|x| \to \infty$.

Definition 2.1. Let $\mathsf{F} = \mathsf{G} \circ \theta$ with G and θ satisfying **(A1)**–**(A3)**. Assume that (A, U) is an admissible parametrization of G. We say that a Young measure $\nu : \mathbb{R}^{d+1}_+ \to \text{Prob}(\mathbb{R})$ is a local entropy measure valued solution to (1.1) corresponding to (G, θ) and (A, U) if there exists $R(t,x) \in L^\infty_{loc}(\mathbb{R}^{d+1}_+)$ such that

$$\operatorname{supp} \nu_{(t,x)} \subset [-R(t,x), R(t,x)] \qquad \text{for a.a. } (t,x) \in \mathbb{R}^{d+1}_+ \tag{2.4}$$

and if for all $\mu \in \mathbb{R}$ and all nonnegative $\varphi \in \mathcal{D}(\mathbb{R}^{d+1}_+)$ there holds

$$\begin{aligned} &\int_{\mathbb{R}^{d+1}_+} \langle |\eta(x,U(\lambda)) - \eta(x,U(\mu))|, \nu_{(t,x)}(\lambda)\rangle \varphi_{,t}(t,x)\, dx\, dt \\ &\quad + \int_{\mathbb{R}^{d+1}_+} \langle (\mathbf{A}(\lambda) - \mathbf{A}(\mu)) \operatorname{sgn}(\lambda-\mu), \nu_{(t,x)}(\lambda)\rangle \cdot \nabla\varphi(t,x)\, dx\, dt \ge 0. \end{aligned} \tag{2.5}$$

The key observation of the paper, on which all results presented here heavily rely, is the averaged contraction property, which is also the standard property for classical results with smooth fluxes $\mathbf{F}$, see Ref. 22, 37.

Lemma 2.1. *Let $\mathbf{F} = \mathbf{G} \circ \theta$ with $\mathbf{G}$ and θ satisfying **(A1)**–**(A3)** and let $(\mathbf{A}_1, U_1)$, $(\mathbf{A}_2, U_2)$ be two different admissible parametrizations of $\mathbf{G}$. Assume that ν and σ are two local entropy measure-valued solutions to* (1.1) *corresponding to $(\mathbf{A}_1, U_1)$, $(\mathbf{A}_2, U_2)$ respectively in the sense of Definition 2.1. Then*

$$\begin{aligned} &\int_{\mathbb{R}^{d+1}_+} \langle |\eta(x,U_1(\lambda)) - \eta(x,U_2(\mu))|, \nu_{(t,x)}(\lambda)\otimes\sigma_{(t,x)}(\mu)\rangle \psi_{,t}(t,x)\, dx\, dt \\ &\quad + \int_{\mathbb{R}^{d+1}_+} \langle \mathbf{Q}(\lambda,\mu), \nu_{(t,x)}(\lambda)\otimes\sigma_{(t,x)}(\mu)\rangle \cdot \nabla\psi(t,x)\, dx\, dt \ge 0 \end{aligned} \tag{2.6}$$

for all nonnegative $\psi \in \mathcal{D}(\mathbb{R}^{d+1}_+)$. Here, we defined $\mathbf{Q}(\lambda,\mu)$ through

$$\mathbf{Q}(\lambda,\mu) := \begin{cases} (\mathbf{A}_1(\lambda) - \mathbf{A}_2(\mu)) \operatorname{sgn}\left(\dfrac{\lambda-\alpha^1_k}{\beta^1_k-\alpha^1_k} - \dfrac{\mu-\alpha^2_k}{\beta^2_k-\alpha^2_k}\right) \\ \qquad \text{if there is } k \text{ such that } \lambda\in[\alpha^1_k,\beta^1_k],\ \mu\in[\alpha^2_k,\beta^2_k], \\ (\mathbf{A}_1(\lambda) - \mathbf{A}_2(\mu)) \operatorname{sgn}(U_1(\lambda) - U_2(\mu)) \qquad \text{otherwise.} \end{cases} \tag{2.7}$$

The numbers α^i_k, β^i_k, $i = 1, 2$ are defined in (1.16) *and correspond to U_1, U_2 respectively.*

Proof. Although the proof is similar to the one given in Ref. 10 and based on the idea of regularization of Young measures developed in Ref. 22, 37, but the possibly discontinuous dependence in x of the function η involves new difficulties. The main difference in the proof consists in using the smoothing kernel in the product form and then passing to the limit separately, first with the parameter of regularization with respect to x and then with the one with respect to t. It motivates us to conduct the whole proof rigorously in order to avoid any unclarity. Let $\omega \in \mathcal{D}(-1,1)$ be a regularizing kernel, i.e., $\omega(x) = \omega(-x)$ and $\int_{-1}^1 \omega(x)\, dx = 1$. Then, for any $\gamma > 0$, we define

$$\begin{aligned} &\omega^\gamma_1(t) := \gamma^{-1}\omega(t/\gamma) && \text{for all } t\in\mathbb{R}, \\ &\omega^\gamma_2(x) := \gamma^{-d}\omega(x_1/\gamma)\cdot\ldots\cdot\omega(x_d/\gamma) && \text{for all } x = (x_1,\ldots,x_d)\in\mathbb{R}^d. \end{aligned}$$

For arbitrary $\varepsilon, \delta > 0$ we set $\omega^{\delta,\varepsilon}(t,x) := \omega^\delta_1(t)\cdot\omega^\varepsilon_2(x)$. Notice that for any Young measure $\nu \in L^\infty_w([0,T]\times\mathbb{R}^d; \mathcal{M}(\mathbb{R}))$ there exists a Young measure $\nu^\delta \in$

$L^\infty_w(\mathbb{R}^d;\mathcal{C}^\infty([0,T];\mathcal{M}(\mathbb{R})))$ with $\|\nu^\delta\|_{L^\infty_w([0,T]\times\mathbb{R}^d;\mathcal{M}(\mathbb{R}))}\le 1$ such that for any $f\in\mathcal{C}(\mathbb{R})$ the following holds[e] $(\omega_1^\delta * \langle f,\nu\rangle) = \langle f,\nu^\delta\rangle$ for almost all $t\in\mathbb{R}$. Moreover, we can interchange the derivative as $\langle f,\partial_t\nu^\delta\rangle = \langle f,\nu^\delta\rangle_{,t}$ for all $t\in\mathbb{R}$. Similarly, there exists $\nu^\varepsilon\in L^\infty_w([0,T];\mathcal{C}^\infty(\mathbb{R}^d_{loc};\mathcal{M}(\mathbb{R})))$ with $\|\nu^\varepsilon\|_{L^\infty_w([0,T]\times\mathbb{R}^d;\mathcal{M}(\mathbb{R}))}\le 1$ such that $\omega_2^\varepsilon * \langle f,\nu\rangle = \langle f,\nu^\varepsilon\rangle$ and $\langle f,\partial_{x_i}\nu^\varepsilon\rangle = \partial_{x_i}\langle f,\nu^\varepsilon\rangle$ for all $x\in\mathbb{R}^d$, see Ref. 22.

In (2.5) we set

$$\varphi(t,x) := (\psi * \omega^{\delta,\varepsilon})(t,x) = \int_{\mathbb{R}^{d+1}_+} \psi(\tau,y)\omega^{\delta,\varepsilon}(t-\tau,x-y)\,dy\,d\tau$$

where nonnegative $\psi\in\mathcal{D}((\delta,\infty)\times\mathbb{R}^d)$ is arbitrary. Note that $\varphi\in\mathcal{D}(\mathbb{R}^{d+1}_+)$ is nonnegative and therefore such setting is possible. Observe that for all $\mu\in\mathbb{R}$ we have

$$\begin{aligned}&\int_{\mathbb{R}^{d+1}_+}\langle|\eta(x,U_1(\lambda))-\eta(x,U_1(\mu))|,\nu_{(t,x)}(\lambda)\rangle\ (\psi*(\omega_1^\delta\cdot\omega_2^\varepsilon))_{,t}\,dx\,dt\\ &=\int_{\mathbb{R}^{d+1}_+}\omega_2^\varepsilon*\langle|\eta(x,U_1(\lambda))-\eta(x,U_1(\mu))|,\nu^\delta_{(t,x)}(\lambda)\rangle\,\psi_{,t}\,dx\,dt.\end{aligned}\tag{2.8}$$

Similarly, we obtain for all $\mu\in\mathbb{R}$

$$\begin{aligned}&\int_{\mathbb{R}^{d+1}_+}\langle(\mathbf{A}_1(\lambda)-\mathbf{A}_1(\mu))\operatorname{sgn}(\lambda-\mu),\nu_{(t,x)}(\lambda)\rangle\cdot\nabla(\psi*(\omega_1^\delta\cdot\omega_2^\varepsilon))\,dx\,dt=\\ &=\int_{\mathbb{R}^{d+1}_+}\langle(\mathbf{A}_1(\lambda)-\mathbf{A}_1(\mu))\operatorname{sgn}(\lambda-\mu),\nu^{\delta,\varepsilon}_{(t,x)}(\lambda)\rangle\cdot\nabla\psi\,dx\,dt.\end{aligned}\tag{2.9}$$

Consequently, using (2.8) and (2.9) in (2.5), we deduce that for all $\mu\in\mathbb{R}$ and all nonnegative $\psi\in\mathcal{D}((\delta,\infty)\times\mathbb{R}^d)$ there holds

$$\begin{aligned}&\int_{\mathbb{R}^{d+1}_+}\omega_2^\varepsilon*\langle|\eta(x,U_1(\lambda))-\eta(x,U_1(\mu))|,\nu^\delta_{(t,x)}(\lambda)\rangle\psi_{,t}\,dx\,dt\\ &\quad+\int_{\mathbb{R}^{d+1}_+}\langle(\mathbf{A}_1(\lambda)-\mathbf{A}_1(\mu))\operatorname{sgn}(\lambda-\mu),\nu^{\delta,\varepsilon}_{(t,x)}(\lambda)\rangle\cdot\nabla\psi\,dx\,dt\ge 0,\end{aligned}\tag{2.10}$$

which in particular implies that for all $\tilde\mu\in\mathbb{R}$ and all $(t,x)\in(\delta,\infty)\times\mathbb{R}^d$ there holds

$$\begin{aligned}&\left(\omega_2^\varepsilon*\langle|\eta(x,U_1(\lambda))-\eta(x,U_1(\tilde\mu))|,\nu^\delta_{(t,x)}(\lambda)\rangle\right)_{,t}\\ &+\operatorname{div}\left(\langle(\mathbf{A}_1(\lambda)-\mathbf{A}_1(\tilde\mu))\operatorname{sgn}(\lambda-\tilde\mu),\nu^{\delta,\varepsilon}_{(t,x)}(\lambda)\rangle\right)\le 0.\end{aligned}\tag{2.11}$$

Similarly, for a Young measure σ and functions U_2 and $\mathbf{A}_2$, we can deduce that for any $\varepsilon>0$, $\tilde\lambda\in\mathbb{R}$ and all $(t,x)\in(\delta,\infty)\times\mathbb{R}^d$ we have

$$\begin{aligned}&\left(\omega_2^\varepsilon*\langle|\eta(x,U_2(\tilde\lambda))-\eta(x,U_2(\mu))|,\sigma^\delta_{(t,x)}(\mu)\rangle\right)_{,t}\\ &\quad+\operatorname{div}\left(\langle(\mathbf{A}_2(\tilde\lambda)-\mathbf{A}_2(\mu))\operatorname{sgn}(\tilde\lambda-\mu),\sigma^{\delta,\varepsilon}_{(t,x)}(\mu)\rangle\right)\le 0.\end{aligned}\tag{2.12}$$

[e]We extend the measure for $t<0$ and $t>T$ by zero.

Following Ref. 10 we show that (2.11) and (2.12) imply that for all $\mu \in \mathbb{R}$ the following inequality holds for any $\varepsilon > 0$ point-wisely

$$\left.\begin{aligned}&\left(\omega_2^\varepsilon * \langle|\eta(x,U_1(\lambda))-\eta(x,U_2(\mu))|,\nu^\delta_{(t,x)}(\lambda)\rangle\right)_{,t}\\&\quad+\operatorname{div}\left(\langle \mathbf{Q}(\lambda,\mu),\nu^{\delta,\varepsilon}_{(t,x)}(\lambda)\rangle\right)\le 0\end{aligned}\right\}\text{ in } (\delta,\infty)\times\mathbb{R}^d. \tag{2.13}$$

Here, $\mathbf{Q}$ was defined in (2.7). Similarly, one can observe that for all $\lambda \in \mathbb{R}$, there holds

$$\left.\begin{aligned}&\left(\omega_2^\varepsilon * \langle|\eta(x,U_1(\lambda))-\eta(x,U_2(\mu))|,\sigma^\delta_{(t,x)}(\mu)\rangle\right)_{,t}\\&\quad+\operatorname{div}\left(\langle \mathbf{Q}(\lambda,\mu),\sigma^{\delta,\varepsilon}_{(t,x)}(\mu)\rangle\right)\le 0\end{aligned}\right\}\text{ in } (\delta,\infty)\times\mathbb{R}^d. \tag{2.14}$$

In order to deduce (2.6) we combine (2.13) and (2.14). For brevity, set $\zeta(x,\lambda,\mu) := |\eta(x,U_1(\lambda))-\eta(x,U_2(\mu))|$. The main effort is directed to the function ζ, because of its dependence on the variable x. As a consequence of the Fubini theorem it holds (note that all expressions are well-defined)

$$\begin{aligned}&\operatorname{div}\left(\langle \mathbf{Q}(\lambda,\mu),\nu^{\delta,\varepsilon}_{(t,x)}(\lambda)\otimes\sigma^{\delta,\varepsilon}_{(t,x)}(\mu)\rangle\right)\\&\quad=\left\langle \operatorname{div}\left(\langle \mathbf{Q}(\lambda,\mu),\nu^{\delta,\varepsilon}_{(t,x)}(\lambda)\rangle\right),\sigma^{\delta,\varepsilon}_{(t,x)}(\mu)\right\rangle\\&\qquad+\left\langle \operatorname{div}\left(\langle \mathbf{Q}(\lambda,\mu),\sigma^{\delta,\varepsilon}_{(t,x)}(\mu)\rangle\right),\nu^{\delta,\varepsilon}_{(t,x)}(\lambda)\right\rangle.\end{aligned} \tag{2.15}$$

We apply $\sigma^{\delta,\varepsilon}_{(t,x)}$ onto (2.13) (note that it is continuous function of μ), similarly we apply $\nu^{\delta,\varepsilon}_{(t,x)}$ onto (2.14). Summing the resulting expressions and using (2.15) we find that for all $(t,x)\in(2\varepsilon,\infty)\times\mathbb{R}^d$ there holds

$$\begin{aligned}&\langle \omega_2^\varepsilon * \langle\zeta(x,\lambda,\mu),\nu^\delta_{(t,x)}(\lambda)\rangle_{,t},\sigma^{\delta,\varepsilon}_{(t,x)}(\mu)\rangle\\&+\langle \omega_2^\varepsilon * \langle\zeta(x,\lambda,\mu),\sigma^\delta_{(t,x)}(\mu)\rangle_{,t},\nu^{\delta,\varepsilon}_{(t,x)}(\lambda)\rangle\\&+\operatorname{div}\left(\langle \mathbf{Q}(\lambda,\mu),\nu^{\delta,\varepsilon}_{(t,x)}(\lambda)\otimes\sigma^{\delta,\varepsilon}_{(t,x)}(\mu)\rangle\right)\le 0.\end{aligned} \tag{2.16}$$

Thus, multiplying (2.16) by an arbitrary fixed nonnegative $\psi\in\mathcal{D}((2\delta,\infty)\times\mathbb{R}^d)$, integrating the result over $\mathbb{R}^{d+1}_+$ and using integration by parts, we find that

$$\begin{aligned}-\int_{\mathbb{R}^{d+1}_+}&\left(\left\langle \omega_2^\varepsilon * \langle\zeta(x,\lambda,\mu),\nu^\delta_{(t,x)}(\lambda)\rangle_{,t},\sigma^{\delta,\varepsilon}_{(t,x)}(\mu)\right\rangle\right.\\&\left.+\left\langle \omega_2^\varepsilon * \langle\zeta(x,\lambda,\mu),\sigma^\delta_{(t,x)}(\mu)\rangle_{,t},\nu^{\delta,\varepsilon}_{(t,x)}(\lambda)\right\rangle\right)\psi\,dx\,dt\\&+\int_{\mathbb{R}^{d+1}_+}\left\langle \mathbf{Q}(\lambda,\mu),\nu^{\delta,\varepsilon}_{(t,x)}(\lambda)\otimes\sigma^{\delta,\varepsilon}_{(t,x)}(\mu)\right\rangle\cdot\nabla\psi\,dx\,dt\ge 0.\end{aligned} \tag{2.17}$$

First, we let $\varepsilon\to 0_+$. Then let $\Omega_\psi := \operatorname{supp}\psi$. From (2.4) it follows that there exists a compact set K such that for $(t,x)\in\Omega_\psi$ we have $\operatorname{supp}\nu^\delta_{(t,x)}\subset K$ and then also $\operatorname{supp}\partial_t\nu^\delta_{(t,x)}\subset K$. The same holds for $\sigma^\delta_{(t,x)}$. Since **(A3*)** provides that $\eta\in$

$L^\infty(\mathbb{R}^{d+1}_+;\mathcal{C}(K))$, then also $\eta \in L^1(\Omega_\psi;\mathcal{C}(K))$ and consequently $\zeta \in L^1(\Omega_\psi;\mathcal{C}(K))$. Thus we can extract a subsequence, that we do not relabel, such that

$$\begin{aligned}
\omega_2^\varepsilon * \langle \zeta, \partial_t \nu^\delta\rangle &\to \langle \zeta, \partial_t \nu^\delta\rangle && \text{strongly in} && L^1(\Omega_\psi;\mathcal{C}(K)),\\
\omega_2^\varepsilon * \langle \zeta, \partial_t \sigma^\delta\rangle &\to \langle \zeta, \partial_t \sigma^\delta\rangle && \text{strongly in} && L^1(\Omega_\psi;\mathcal{C}(K)),\\
\sigma^{\delta,\varepsilon} &\rightharpoonup^* \sigma^\delta && \text{weakly}^* \text{ in} && L^\infty_w(\Omega_\psi;\mathcal{M}(K)),\\
\nu^{\delta,\varepsilon} &\rightharpoonup^* \nu^\delta && \text{weakly}^* \text{ in} && L^\infty_w(\Omega_\psi;\mathcal{M}(K)).
\end{aligned}$$

Using these convergence results, we observe from (2.17) that

$$\begin{aligned}
&-\int_{\mathbb{R}^{d+1}_+} \langle\langle \zeta(x,\lambda,\mu), \nu^\delta_{(t,x)}(\lambda)\rangle_{,t}, \sigma^\delta_{(t,x)}(\mu)\rangle \psi \, dx\, dt\\
&\quad -\int_{\mathbb{R}^{d+1}_+} \langle\langle \zeta(x,\lambda,\mu), \sigma^\delta_{(t,x)}(\mu)\rangle_{,t}, \nu^\delta_{(t,x)}(\lambda)\rangle \psi \, dx\, dt\\
&\quad +\int_{\mathbb{R}^{d+1}_+} \langle \mathbf{Q}(\lambda,\mu), \nu^\delta_{(t,x)}(\lambda)\otimes \sigma^\delta_{(t,x)}(\mu)\rangle \cdot \nabla\psi \, dx\, dt \ge 0.
\end{aligned} \tag{2.18}$$

Similarly to (2.15) it is not difficult to observe that

$$\begin{aligned}
\langle \zeta, \nu^\delta_{(t,x)} \otimes \sigma^\delta_{(t,x)}\rangle_{,t} &= \langle\langle \zeta, \nu^\delta_{(t,x)}\rangle, \sigma^\delta_{(t,x)}\rangle_{,t} = \langle \omega^\delta * \langle \zeta, \nu_{(t,x)}\rangle, \sigma^\delta_{(t,x)}\rangle_{,t}\\
&= \left\langle (\omega^\delta * \langle \zeta, \nu_{(t,x)}\rangle)_{,t}, \sigma^\delta_{(t,x)}\right\rangle + \left\langle (\omega^\delta * \langle \zeta, \sigma_{(t,x)}\rangle)_{,t}, \nu^\delta_{(t,x)}\right\rangle.
\end{aligned} \tag{2.19}$$

Thus, using (2.18), (2.19) and integrating by parts with respect to t, we find that

$$\begin{aligned}
&\int_{\mathbb{R}^{d+1}_+} \langle \zeta(x,\lambda,\mu), \nu^\delta_{(t,x)}(\lambda)\otimes \sigma^\delta_{(t,x)}(\mu)\rangle \psi_{,t} \, dx\, dt\\
&\quad +\int_{\mathbb{R}^{d+1}_+} \langle \mathbf{Q}(\lambda,\mu), \nu^\delta_{(t,x)}(\lambda)\otimes \sigma^\delta_{(t,x)}(\mu)\rangle \cdot \nabla\psi \, dx\, dt \ge 0.
\end{aligned}$$

Finally, letting $\delta \to 0_+$ we conclude (2.6) by the argument of weak* convergence of measures ν^δ and σ^δ to ν and σ, respectively. □

In Lemma 2.1 we showed that any two local entropy measure-valued solutions satisfy the contraction property, which is the main tool for proving uniqueness. However, in order to get such a result, we have to specify in which sense an initial condition is attained and what is the behavior of the solution for $|x| \to \infty$. Indeed, if (possibly) two different solutions have different initial and "boundary" value, one cannot expect that they are identical. The next theorem provides sufficient conditions[f] that provide the "uniqueness" of a solution.

[f] It is well known that in the case of Lipschitz continuous fluxes there is a finite speed of propagation which is bounded by the Lipschitz constant of $\mathbf{F}$. Therefore a solution at certain point depends only on the initial data in some compact subset (and not on some assumption on the behavior of a solution at infinity) as it is pointed out in Theorem 6.2.3 of Ref. 19. In our opinion, for the discontinuous fluxes there is no pure hyperbolic feature like finite speed of propagation, which imposes that some kind of behavior at infinity of a solution needs to be added to get uniqueness.

Theorem 2.1 (Uniqueness). *Let $\mathsf{F} = \mathsf{G} \circ \theta$ with G and θ satisfying **(A1)**–**(A3)** and let (A_1, U_1), (A_2, U_2) be two different admissible parametrization of G. Assume that ν^1, ν^2 are two local entropy measure-valued solutions to* (1.1) *corresponding to (A_1, U_1) and (A_2, U_2) respectively in the sense of Definition 2.1. Moreover, assume that*

- *There exists $u_0 \in L^1_{loc}(\mathbb{R}^d)$ (initial condition) such that for all compact $K \subset \mathbb{R}^d$ the following holds*

$$\begin{aligned} \operatorname*{ess-liminf}_{t\to 0_+} \int_K \langle |\eta(x, U_1(\lambda)) - u_0(x)|, \nu^1_{(t,x)}(\lambda)\rangle \, dx = 0, \\ \operatorname*{ess-liminf}_{t\to 0_+} \int_K \langle |\eta(x, U_2(\mu)) - u_0(x)|, \nu^2_{(t,x)}(\mu)\rangle \, dx = 0. \end{aligned} \tag{2.20}$$

- *There exists $1 \le p \le \frac{d}{d-1}$ such that for any $T > 0$ the following holds*

$$\begin{aligned} \int_0^T \left(\int_{\mathbb{R}^d} \left| \langle \mathsf{A}_1(\lambda), \nu^1_{(t,x)}(\lambda)\rangle \right|^p \, dx \right)^{\frac{1}{p}} \, dt < \infty \\ \int_0^T \left(\int_{\mathbb{R}^d} \left| \langle \mathsf{A}_2(\mu), \nu^2_{(t,x)}(\mu)\rangle \right|^p \, dx \right)^{\frac{1}{p}} \, dt < \infty. \end{aligned} \tag{2.21}$$

Then for a.a. $(t,x) \in \mathbb{R}^{d+1}_+$ there exists $\lambda_0 = \lambda_0(t,x)$ such that

$$\begin{aligned} \operatorname{supp} \nu^1_{(t,x)} \subset \{\lambda;\ U_1(\lambda) = U_1(\lambda_0)\}, \\ \operatorname{supp} \nu^2_{(t,x)} \subset \{\mu;\ U_2(\mu) = U_1(\lambda_0)\}. \end{aligned} \tag{2.22}$$

In particular, defining $u(t,x) := \int_{\mathbb{R}} \eta(x, U_1(\lambda)) \, d\nu^1_{(t,x)}(\lambda)$, we have for all $M \in \mathcal{C}(\mathbb{R})$ and almost all $(t,x) \in \mathbb{R}^{d+1}_+$ that

$$\begin{aligned} M(u(t,x)) &= \int_{\mathbb{R}} M(\eta(x, U_1(\lambda))) \, d\nu^1_{(t,x)}(\lambda) \\ &= \int_{\mathbb{R}} M(\eta(x, U_2(\mu))) \, d\nu^2_{(t,x)}(\mu). \end{aligned} \tag{2.23}$$

Proof. Here, we again proceed rigorously in order to avoid any doubts about the correctness of the procedure. Let $0 < \varepsilon < t_0 < T < \infty$ be arbitrary. We define an affine ψ_1 as follows

$$\psi_1(t) := \begin{cases} 0 & t \in [0, t_0 - \varepsilon) \cup [T, \infty), \\ \dfrac{t - t_0 + \varepsilon}{\varepsilon} & t \in (t_0 - \varepsilon, t_0), \\ \dfrac{T - t}{T - t_0} & t \in (t_0, T). \end{cases}$$

Let $\psi_2^n \in \mathcal{D}(\mathbb{R}^d)$ be arbitrary such that $\|\psi_2^n\|_\infty \le 1$. Then we set $\psi(t,x) := \psi_1(t)\psi_2^n(x)$ in (2.6) (it is a possible test function since we can mollify ψ_1 and then pass to the limit) to deduce that

$$\begin{aligned}
&\frac{1}{T-t_0}\int_{t_0}^{T}\int_{\mathbb{R}^d}\langle|\eta(x,U_1(\lambda))-\eta(x,U_2(\mu))|,\nu^1_{(t,x)}(\lambda)\otimes\nu^2_{(t,x)}(\mu)\rangle\psi^n_2(x)\,dx\,dt\\
&\le\frac{1}{\varepsilon}\int_{t_0-\varepsilon}^{t_0}\int_{\mathbb{R}^d}\langle|\eta(x,U_1(\lambda))-\eta(x,U_2(\mu))|,\nu^1_{(t,x)}(\lambda)\otimes\nu^2_{(t,x)}(\mu)\rangle\psi^n_2(x)\,dx\,dt\\
&+\int_{t_0-\varepsilon}^{T}\int_{\mathbb{R}^d}\langle\mathbf{Q}(\lambda,\mu),\nu^1_{(t,x)}(\lambda)\otimes\nu^2_{(t,x)}(\mu)\rangle\cdot\nabla\psi^n_2(x)\psi_1(t)\,dx\,dt\\
&=: I_1(\varepsilon,t_0,n)+I_2(\varepsilon,t_0,T,n).
\end{aligned}$$

Assume that $\operatorname{supp}\psi^n_2\subset K$, where K is a compact subset of $\mathbb{R}^d$ and let $\varepsilon\to 0_+$. Then one easily shows that

$$\lim_{\varepsilon\to 0_+} I_2(\varepsilon,t_0,T,n)=I_2(t_0,T,n),$$

where

$$\begin{aligned}
I_2(t_0,T,n)&:=\int_{t_0}^{T}\int_{\mathbb{R}^d}\langle\mathbf{Q}(\lambda,\mu),\nu^1_{(t,x)}(\lambda)\otimes\nu^2_{(t,x)}(\mu)\rangle\cdot\nabla\psi^n_2(x)\psi_1(t)\,dx\,dt\\
&\le\int_{0}^{T}\int_{\mathbb{R}^d}\langle|\mathbf{Q}(\lambda,\mu)|,\nu^1_{(t,x)}(\lambda)\otimes\nu^2_{(t,x)}(\mu)\rangle|\nabla\psi^n_2(x)|\,dx\,dt,
\end{aligned}\tag{2.24}$$

and where we used the Jensen inequality for estimating the term on the right hand side. By the mean value theorem we conclude that for almost all $t_0\in(0,T)$ there holds

$$\lim_{\varepsilon\to 0_+} I_1(\varepsilon,t_0,n)=I_1(t_0,n),$$

where

$$\begin{aligned}
I_1(t_0,n)&:=\int_{\mathbb{R}^d}\langle|\eta(x,U_1(\lambda))-\eta(x,U_2(\mu))|,\nu^1_{(t_0,x)}(\lambda)\otimes\nu^2_{(t_0,x)}(\mu)\rangle\psi^n_2(x)\,dx\\
&\le\int_K\langle|\eta(x,U_1(\lambda))-\eta(x,U_2(\mu))|,\nu^1_{(t_0,x)}(\lambda)\otimes\nu^2_{(t_0,x)}(\mu)\rangle\,dx.
\end{aligned}$$

We shall now show that I_1 tends to zero as $t_0\to 0$. For this purpose we notice that ν^1 and ν^2 are for almost all $(t,x)\in\mathbb{R}^{d+1}_+$ probabilistic measures with compact support and we use the triangle inequality to conclude that for almost all $t_0\in(0,T)$

$$\begin{aligned}
I_1(t_0,n)&\le\int_K\Big\langle|\eta(x,U_1(\lambda))-u_0(x)|+|\eta(x,U_2(\mu))-u_0(x)|,\\
&\qquad\nu^1_{(t_0,x)}(\lambda)\otimes\nu^2_{(t_0,x)}(\mu)\Big\rangle\,dx\\
&=\int_K\Big\langle|\eta(x,U_1(\lambda))-u_0(x)|,\nu^1_{(t_0,x)}(\lambda)\Big\rangle\,dx\\
&\qquad+C(\psi^n)\int_K\Big\langle|\eta(x,U_2(\mu))-u_0(x)|,\nu^2_{(t_0,x)}(\mu)\Big\rangle\,dx.
\end{aligned}$$

Hence, using the assumption (2.20), we find that

$$\operatorname*{ess\text{-}liminf}_{t_0\to 0_+} I_1(t_0,n) = 0. \tag{2.25}$$

Consequently, we first let $\varepsilon \to 0_+$ and then $t_0 \to 0_+$ in (2.6) and with help of (2.25) and (2.24) we find that for arbitrary $\psi_2^n \in \mathcal{D}(\mathbb{R}^d)$ such that $\|\psi_2^n\|_\infty \le 1$ and any $T > 0$ there holds

$$\begin{aligned}&\int_0^T \int_{\mathbb{R}^d} \langle |\eta(x,U_1(\lambda)) - \eta(x,U_2(\mu))|, \nu^1_{(t,x)}(\lambda)\otimes \nu^2_{(t,x)}(\mu)\rangle \psi_2^n(x)\,dx\,dt\\ &\le T\int_0^T \int_{\mathbb{R}^d} \langle |\mathbf{Q}(\lambda,\mu)|, \nu^1_{(t,x)}(\lambda)\otimes \nu^2_{(t,x)}(\mu)\rangle |\nabla\psi_2^n(x)|\,dx\,dt.\end{aligned} \tag{2.26}$$

Using the definition of $\mathbf{Q}$ and the triangle inequality we observe that for almost all $(t,x) \in \mathbb{R}^{d+1}_+$

$$\langle |\mathbf{Q}(\lambda,\mu)|, \nu^1_{(t,x)}(\lambda)\otimes \nu^2_{(t,x)}(\mu)\rangle \le \langle |\mathbf{A}_1(\lambda)|, \nu^1_{(t,x)}(\lambda)\rangle + \langle |\mathbf{A}_2(\mu)|, \nu^2_{(t,x)}(\mu)\rangle.$$

Hence, by (2.21) we conclude that

$$\langle |\mathbf{Q}(\lambda,\mu)|, \nu^1_{(t,x)}(\lambda)\otimes \nu^2_{(t,x)}(\mu)\rangle \in L^1(0,T;L^p(\mathbb{R}^d)). \tag{2.27}$$

Finally, we define a sequence $\psi_2^n \nearrow 1$ of smooth nonnegative compactly supported functions as $\psi_2^n(x) := 1$ in $B(0,n)$, $\psi_2^n(x) := 0$ for $x \in \mathbb{R}^d \setminus B(0,2n)$ such that $|\nabla\psi_2^n| \le \frac{c}{n}$. One immediately observes that

$$\int_{\mathbb{R}^d} |\nabla\psi_2^n|^q\,dx \le C \qquad \text{for all } q \ge d.$$

Consequently, we obtain that

$$|\nabla\psi^n| \rightharpoonup^* 0 \text{ weakly}^* \text{ in } L^\infty(0,T;L^q(\mathbb{R}^d)) \qquad \text{for all } q \ge d. \tag{2.28}$$

Hence, using (2.27) and the weak* convergence (2.28), we see that the right hand side of (2.26) tends to 0 as $n \to \infty$. With the monotone convergence theorem we conclude that

$$\int_0^T \int_{\mathbb{R}^d} \langle |\eta(x,U_1(\lambda)) - \eta(x,U_2(\mu))|, \nu^1_{(t,x)}(\lambda)\otimes \nu^2_{(t,x)}(\mu)\rangle\,dx\,dt \le 0, \tag{2.29}$$

which implies that for almost all $(t,x) \in \mathbb{R}^{d+1}_+$ we have

$$\langle |\eta(x,U_1(\lambda)) - \eta(x,U_2(\mu))|, \nu^1_{(t,x)}(\lambda)\otimes \nu^2_{(t,x)}(\mu)\rangle = 0. \tag{2.30}$$

As a conclusion of (2.30) we recover relations (2.22) in the same manner as in Ref. 10. The relation (2.23) then easily follows from (2.22). □

Theorem 2.2 (Existence). *Let $\mathbf{F} = \mathbf{G}\circ\theta$ with $\mathbf{G}$ and θ satisfying **(A1) – (A3)**. Assume $u_0 \in L^1(\mathbb{R}^d)\cap L^\infty(\mathbb{R}^d)$. Then there exists an admissible parametrization $(\mathbf{A},U)$ of $\mathbf{G}$ and there exists a local measure-valued solution ν to* (1.1) *in the sense of Definition 2.1 that in addition satisfies*

$$\operatorname*{ess\text{-}liminf}_{t\to 0_+} \int_K \int_{\mathbb{R}} |\eta(x,U(\lambda)) - u_0(x)|\,d\nu_{(t,x)}(\lambda)\,dx = 0 \tag{2.31}$$

*for any compact $K \subset \mathbb{R}^d$. Moreover, if G satisfies **(A4)**, then for any $T > 0$ the following holds*

$$\int_0^T \left(\int_{\mathbb{R}^d} \left| \langle \mathsf{A}(\lambda), \nu_{(t,x)}(\lambda) \rangle \right|^p \, dx \right)^{\frac{1}{p}} \, dt < \infty. \tag{2.32}$$

Proof. For function G there exist functions U and A as described in (1.16). Before we pass to constructing the approximative problem let us describe, for the readers' convenience, the change of variables[g].

$$\begin{aligned} u &= \eta(x, U(g)), & v &= \theta(x,u), & u &= \eta(x,v), \\ g &= U^{-1}(\theta(x,u)), & g &= U^{-1}(v), & v &= U(g). \end{aligned}$$

We construct an approximative problem by regularizing U, A, and θ. To be more precise, we define

$$U_n(s) := U(s) + \frac{s}{n}. \tag{2.33}$$

with a standard regularizing kernel $\omega^{\frac{1}{n}}$. Note that U_n is strictly increasing function and we denote by U_n^{-1} its inverse. It is easy to observe that U_n^{-1} is Lipschitz continuous with Lipschitz constant less or equal to n. For the purpose of applying Lemma A.1 we need to provide that the flux function is at least Lipschitz, so we find a sequence of continuously differentiable functions A^n such that for every compact set $K \subset \mathbb{R}$

$$\mathsf{A}^n \to \mathsf{A} \text{ strongly in } \mathcal{C}(K)^d. \tag{2.34}$$

Note that such a construction is always possible due to the continuity of A. Then we define

$$\eta^{(n)}(x,z) := \eta(x,z) + \frac{z}{n} \tag{2.35}$$

and by $\theta^{(n)}$ we mean the inverse to $\eta^{(n)}$ with respect to the second variable. Note that the construction is analogous to the construction of the function U_n^{-1}, hence again $\theta^{(n)}(x,s)$ is Lipschitz continuos w.r.t. the second variable. Finally, we use a standard mollification procedure and introduce

$$\theta^{\frac{1}{n}}(x,s) := \int_{\mathbb{R}}^d \omega^{\frac{1}{n}}(x-y)\theta^{(n)}(y,s) \, dy, \tag{2.36}$$

where $\omega^{\frac{1}{n}}$ is standard mollification kernel of radius $\frac{1}{n}$ and we denote by $\eta^{\frac{1}{n}}(x,z)$ the inverse function to $\theta^{\frac{1}{n}}(x,s)$, i.e., $\eta^{\frac{1}{n}}(x, \theta^{\frac{1}{n}}(x,s)) = s$. Note that the inverse surely exists since θ (and consequently $\theta^{(n)}$ and $\theta^{\frac{1}{n}}$) is for almost all x strictly increasing with respect to s.

[g]The function U^{-1} has a single-valued meaning when U is strictly monotone, otherwise we understand it as a maximal monotone, multi-valued operator.

The approximative problem we start with is the following

$$\begin{aligned} u^n_{,t} + \operatorname{div} \mathbf{A}^n(U_n^{-1}(\theta^{1/n}(x,u^n))) = 0 \qquad & \text{in } \mathbb{R}^{d+1}_+, \\ u^n(0,x) = u_0(x) \qquad & \text{in } \mathbb{R}^d. \end{aligned} \tag{2.37}$$

Due to the mollification of θ with respect to x, we can recall the results of Kružkov,[29] where the existence of a unique entropy solution is shown.

Due to the all introduced mollification we are now in position when we can apply Lemma A.1. Hence, defining

$$g^n(t,x) := U_n^{-1}(\theta^{\frac{1}{n}}(x,u(t,x))),$$

we see from (A.5) that it satisfies

$$\eta^{\frac{1}{n}}(x,U_n(g^n))_{,t} + \operatorname{div} \mathbf{A}^n(g^n) = 0 \qquad \text{in } \mathbb{R}^{d+1}_+, \tag{2.38}$$

$$\eta^{\frac{1}{n}}(x,U_n(g^n(0,x))) = u_0(x) \qquad \text{in } \mathbb{R}^d, \tag{2.39}$$

and in addition by Lemma A.1 the following entropy inequality holds

$$\begin{aligned} &\int_{\mathbb{R}^{d+1}_+} \operatorname{sgn}(U_n(g^n(t,x)) - U_n(k))(\mathbf{A}^n(g^n(t,x)) - \mathbf{A}^n(k)) \cdot \nabla\psi(t,x)\, dx\, dt \\ &\quad + \int_{\mathbb{R}^{d+1}_+} |\eta^{\frac{1}{n}}(x,U_n(g^n(t,x))) - \eta^{\frac{1}{n}}(x,U_n(k))|\psi_{,t}\, dx\, dt \\ &\quad + \int_{\mathbb{R}^d} |u_0(x) - \eta^{\frac{1}{n}}(x,U_n(k))|\psi(0,x)\, dx \geq 0. \end{aligned} \tag{2.40}$$

for any constant $k \in \mathbb{R}$ and for all nonnegative $\psi \in \mathcal{D}(\mathbb{R}^{d+1})$. Our goal now is to let $n \to \infty$ in (2.40).

By using a standard comparison argument (see Ref. 29 or 7, 33), we observe that

$$|\eta^{\frac{1}{n}}(x,U_n(g^n(t,x)))| \leq \|\eta^{\frac{1}{n}}(x,U_n(g^n(0,x)))\|_\infty = \|u_0\|_\infty \text{ a.e. in } \mathbb{R}^{d+1}_+. \tag{2.41}$$

Consequently, since $\eta^{\frac{1}{n}}(x,0) = U_n(0) = 0$ and both functions are strictly increasing, we see that

$$\|U_n(g^n)\|_\infty \leq \sup_{x\in\mathbb{R}^d} \theta^{\frac{1}{n}}(x,\|u_0\|_\infty) \leq h_2(\|u_0\|_\infty), \tag{2.42}$$

where for the second inequality we used **(A3)**. Thus, finally we get

$$\|g^n\|_\infty \leq U_n^{-1}(h_2(\|u_0\|_\infty)) \leq U^{-1}(h_2(\|u_0\|_\infty)) \leq C, \tag{2.43}$$

where U^{-1} is understood as a maximal monotone operator. Note that for the second inequality we used the fact that $|U(s)| \leq |U_n(s)|$ and for the last inequality we used the fact that U maps any bounded interval onto a bounded interval.

Hence having (2.43), we can find $g \in L^\infty(\mathbb{R}^{d+1}_+)$ and a Young measure $\nu_{(t,x)}$ corresponding to a (not relabeled) subsequence $\{g^n\}_{n=1}^\infty$, which is for almost all

(t, x) compactly supported in a a ball $B(0, \min(U^{-1}(h_2(R))))$, with $R := \|u_0\|_\infty$ such that for any continuous f

$$g^n \rightharpoonup^* g \qquad \text{weakly}^* \text{ in } L^\infty_{loc}(\mathbb{R}^{d+1}_+), \tag{2.44}$$

$$f(g^n) \rightharpoonup^* \overline{f} \qquad \text{weakly}^* \text{ in } L^\infty_{loc}(\mathbb{R}^{d+1}_+), \tag{2.45}$$

where

$$\overline{f}(t, x) = \int_{\mathbb{R}} f(\lambda)\, d\nu_{(t,x)}(\lambda). \tag{2.46}$$

Our goal is to show that the measure ν is an entropy measure valued solution in the sense of Definition 2.1. First, it directly follows from (2.43) that (2.4) holds. We let $n \to \infty$ in (2.40). For the first term we use the fact that U_n is strictly monotone and therefore

$$M^n(\xi) := \mathrm{sgn}(U_n(\xi) - U_n(k))(\mathbf{A}^n(\xi) - \mathbf{A}^n(k)) = \mathrm{sgn}(\xi - k)(\mathbf{A}^n(\xi) - \mathbf{A}^n(k)).$$

Moreover, it follows from (2.34) that for every compact set $K \subset \mathbb{R}$

$$M^n(\xi) \to M = \mathrm{sgn}(\xi - k)(\mathbf{A}(\xi) - \mathbf{A}(k)) \quad \text{strongly in } \mathcal{C}(K)^d.$$

Consequently, using (2.44) and (2.46) we conclude

$$\begin{aligned} M^n(g^n) &\rightharpoonup^* \overline{M} \quad \text{weakly}^* \text{ in } L^\infty_{loc}(\mathbb{R}^{d+1}_+). \\ \overline{M}(t, x) &= \int_{\mathbb{R}} \mathrm{sgn}(\lambda - k)(\mathbf{A}(\lambda) - \mathbf{A}(k))d\nu_{(t,x)}(\lambda). \end{aligned} \tag{2.47}$$

For the second and the the third term of (2.40), we recall the convergence properties of $\theta^{\frac{1}{n}}$ and $\eta^{\frac{1}{n}}$. Since θ is strictly increasing and continuous in u, then the same holds for $\theta^{\frac{1}{n}}$. The inverse function $\eta^{\frac{1}{n}}$ is also increasing and continuous with respect to u. The convergence of convolutions and the monotonicity of θ and $\theta^{\frac{1}{n}}$ provide that $\theta^{\frac{1}{n}}(x, u)$ converges point-wisely with respect to x and uniformly with respect to u on a bounded interval $[-R, R]$ to the function θ, see Proposition C.1. More precisely, we have for all $R > 0$

$$\theta^{\frac{1}{n}} \to \theta \qquad \text{strongly in } L^1_{\mathrm{loc}}(\mathbb{R}^d; \mathcal{C}([-R, R])). \tag{2.48}$$

Consequently, by using Proposition C.2, we obtain that the inverse functions $\eta^{\frac{1}{n}}$ have the same convergence properties, namely for all $R > 0$ there holds

$$\eta^{\frac{1}{n}} \to \eta \qquad \text{strongly in } L^1_{\mathrm{loc}}(\mathbb{R}^d; \mathcal{C}([-R, R])). \tag{2.49}$$

Moreover, using the definition of U_n we see that for all $R > 0$

$$U_n \to U \qquad \text{strongly in } \mathcal{C}([-R, R])). \tag{2.50}$$

Consequently, it follows from (2.46), (2.49) and (2.50) that defining $\zeta^n(x, r) := |\eta^{\frac{1}{n}}(x, U_n(r)) - \eta^{\frac{1}{n}}(x, U_n(k))|\psi_{,t}$, we observe

$$\begin{aligned} \lim_{n\to\infty} \int_{\mathbb{R}^{d+1}_+} \zeta^n(x, g^n)\, dx\, dt &= \lim_{n\to\infty} \langle \zeta^n, g^n \rangle \\ &= \int_{\mathbb{R}_+ \times \mathbb{R}^d} \int_{\mathbb{R}} |\eta(x, U(\lambda)) - \eta(x, U(k))| d\nu_{(t,x)}(\lambda)\psi_{,t}\, dx\, dt \end{aligned} \tag{2.51}$$

for all $\psi \in \mathcal{D}(\mathbb{R}^{d+1})$, where the duality pairing is understood between the spaces $L^1(\mathbb{R}^d; \mathcal{C}((-R,R);\mathbb{R})$ and $L^\infty_w(\mathbb{R}^d; \mathcal{M}([-R,R]))$. In the same manner one can also identify the limit in the last term in (2.40). Hence using all above established convergence results, we can easily let $n \to \infty$ in (2.40) to obtain (2.5). Note that in fact, we get a stronger result, since we do not require $\psi(0,x) = 0$. Moreover, it also implies (2.31), which can be proved by following the scheme used in the proof of Theorem 3.2 in Ref. 37.

To finish the proof we need to show that condition (2.32) holds and the proof relies on **(A4)**. Hence, we observe that **(A4)** also implies that for all $x \in \mathbb{R}^d$, $|x| \geq R_\infty$

$$|\mathsf{G}(U(s))|^p \leq C_\infty |\eta(x, U(s))|.$$

By (2.1) the function $\mathsf{A}(s)$ is either equal to $\mathsf{G}(U(s))$ or is linear. In the first case we obtain that

$$|\mathsf{A}(s)|^p \leq C_\infty |\eta(x, U(s))|$$

and in the second case, namely for any $s_k \in [\alpha_k, \beta_k]$ we observe that for $r \in [0,1]$

$$|\mathsf{A}(s_k)|^p = |r\mathsf{G}_+(z_k) + (1-r)\mathsf{G}_-(z)|^p \leq C_\infty |\eta(x, z_k)| \leq C_\infty C(s_k) |\eta(x, U(s_k))|,$$

where $C(s)$ is bounded on bounded intervals. Hence

$$\begin{aligned}
\int_0^T \left(\int_{\mathbb{R}^d} |\langle \mathsf{A}, \nu_{(t,x)} \rangle|^p \, dx \right)^{\frac{1}{p}} dt &\leq \int_0^T \left(\int_{\mathbb{R}^d} \langle |\mathsf{A}|, \nu_{(t,x)} \rangle^p \, dx \right)^{\frac{1}{p}} dt \\
&\leq \int_0^T \left(\int_{\mathbb{R}^d \setminus B_{R_\infty}(0)} \langle |\mathsf{A}|, \nu_{(t,x)} \rangle^p \, dx \right)^{\frac{1}{p}} dt \\
&\quad + \int_0^T \left(\int_{B_{R_\infty}(0)} \langle |\mathsf{A}|, \nu_{(t,x)} \rangle^p \, dx \right)^{\frac{1}{p}} dt \\
&\leq C_\infty \int_0^T \left(\int_{\mathbb{R}^d} \langle |\eta(x,U)|, \nu_{(t,x)} \rangle \, dx \right)^{\frac{1}{p}} dt + C \\
&\leq C_\infty \liminf_{n\to\infty} \int_0^T \left(\int_{\mathbb{R}^d} |\eta(x, U_n(g^n))| \, dx \right)^{\frac{1}{p}} dt + C,
\end{aligned} \tag{2.52}$$

Next, since $\theta^{\frac{1}{n}}(x,s)$ is a convolution of $\theta(x,s)$ with respect to x, then for almost all $x \in \mathbb{R}^d$ and all $s \in \mathbb{R}$ we have

$$\inf_{x\in\mathbb{R}^d} \theta(x,s) \leq \theta^{\frac{1}{n}}(x,s) \leq \sup_{x\in\mathbb{R}^d} \theta(x,s). \tag{2.53}$$

Consequently, using **(A3)**, we can deduce that for all $x \in \mathbb{R}^d$, all $R > 0$ and all $|s| \leq R$ that

$$C^1_R \theta(x,s) \leq \theta^{\frac{1}{n}}(x,s) \leq C^2_R \theta(x,s). \tag{2.54}$$

Hence, by the strict monotonicity of θ we may conclude that the inverse $\eta^{\frac{1}{n}}$ satisfies

$$C_R^2\eta(x,s) \le \eta^{\frac{1}{n}}(x,s) \le C_R^4\eta(x,s) \qquad \text{for all } |s| \le R. \tag{2.55}$$

Hence, since g^n is bounded, wee see from (2.55) that

$$I \le C \liminf_{n\to\infty} \int_0^T \left(\int_{\mathbb{R}^d} |\eta^{\frac{1}{n}}(x, U_n(g^n))| \, dx \right)^{\frac{1}{p}} dt = \int_0^T \|u^n\|_1^{\frac{1}{p}} \, dt, \tag{2.56}$$

where u^n is the Kružkov entropy solution to (2.37) or equivalently the kinetic solution. However, using the standard stability result for the Kružkov solution, we have that

$$\|u^n(t)\|_1 \le \|u_0\|_1 \le C \qquad \text{for all } t > 0,$$

see e.g. Proposition 2 in Ref. 20. Note that this stability result formally follows from setting $k = 0$ in entropy inequality (1.2) and integration over $\mathbb{R}^d$. Here, the term $\operatorname{sgn}(u^n - k) \operatorname{div} \mathbf{F}^n(x,0)$ vanishes, since $\mathbf{F}^n(x,0) = \mathbf{A}^n(U_n^{-1}(\theta^{\frac{1}{n}}(x,0)))$ and $\theta^{\frac{1}{n}}(x,0) = 0$. Consequently, we obtain

$$I \le C$$

which finishes the proof. □

3. Proof of Theorem 1.1

In this final section, we provide a short proof of Theorem 1.1 that is based on the results for entropy measure-valued solution. Having Theorem 2.1 and Theorem 2.2, the proof follows almost step by step the proof given in Ref. 10. Indeed, according to Theorem 2.2 we have the existence of a measure-valued solution. Next, we can use Theorem 2.1 to show that such a measure valued solution is in fact an entropy weak solution. The final part then concerns the uniqueness property. Hence, assume that u_1 and u_2 are two weak entropy solutions in the sense of Definition 1.1 and g_1 and g_2 are related to u_1, u_2 respectively. Then it is shown in Ref. 10 that there exist two Young measures ν^1, ν^2 (being entropy measure-valued solutions) such that $\nu^i_{(t,x)} = \delta_{g_i(t,x)}$ if U_i is strictly monotone at the point $g_i(t,x)$. Finally, we use the uniqueness Theorem 2.1 to get that $u_1 = u_2$. For this purpose we need to show the validity of assumptions of Theorem 2.1. The condition (2.20) is fulfilled trivially. Next, we show that (2.21) holds, which is a consequence of **(A4)**. Indeed, we have that for all $x \in \mathbb{R}^d \setminus B_{R_\infty}$

$$|\mathbf{A}_i(g_i(t,x))|^p \le |\mathbf{G}(U^i(g_i(t,x)))|^p \le C|\eta(x, U^i(g_i(t,x)))| = |u_i(t,x)|.$$

But since u_i, $i = 1,2$, are assumed to belong to $L^\infty(0,T;L^1(\mathbb{R}^d))$ we immediately deduce (2.21), which completes the proof of uniqueness.

A. Equivalent notions of entropy solutions

In this section we concentrate on relations between different notions of entropy weak solutions for the flux function F in a form $\mathsf{F}(x,u) = \mathsf{G}(\theta(x,u))$ with G, θ satisfying **(A1)**–**(A3)** with an additional condition that F is sufficiently regular in both variables. This relations play an important role on the level of approximations, namely after passing from discontinuous flux to sufficiently smooth one. We formulate the lemma collecting the relations between different notions of solutions.

Lemma A.1. *Let* F *satisfy* (1.8) *with* G, θ *satisfying* ***(A1)***–***(A3)*** *and assume that* $\mathsf{G} \in \mathcal{C}^1(\mathbb{R})$*,* θ *is continuous in* u *and continuously differentiable in* x*. In addition let* $U : \mathbb{R} \to \mathbb{R}$ *be smooth, strictly increasing one-to-one mapping with a smooth inverse* U^{-1} *and let* $u_0 \in L^\infty_{loc}(\mathbb{R}^d)$*. Assume that* $u \in L^\infty_{loc}(\mathbb{R}^{d+1})$ *be given and define*

$$v(t,x) := \theta(x, u(t,x)), \tag{A.1}$$

$$g(t,x) := U^{-1}(v(t,x)). \tag{A.2}$$

Then the following statements are equivalent.

(N1) *For all* $k \in \mathbb{R}$ *and all nonnegative* $\psi \in \mathcal{D}(\mathbb{R}^{d+1})$ *there holds*

$$\begin{aligned}\int_{\mathbb{R}^{d+1}_+} &|u(t,x) - k|\psi_{,t}(t,x) - \operatorname{sgn}(u(t,x) - k) \operatorname{div} \mathsf{F}(x,k)\psi(t,x)\, dx\, dt \\ &+ \int_{\mathbb{R}^{d+1}_+} \operatorname{sgn}(u(x,t) - k)(\mathsf{F}(x, u(x,t)) - \mathsf{F}(x,k)) \cdot \nabla\psi(x,t)\, dx\, dt \\ &+ \int_{\mathbb{R}^d} |u_0(x) - k|\psi(0,x)\, dx \geq 0.\end{aligned} \tag{A.3}$$

(N2) *For all* $k \in \mathbb{R}$ *and all nonnegative* $\psi \in \mathcal{D}(\mathbb{R}^{d+1})$ *there holds*

$$\begin{aligned}\int_{\mathbb{R}^{d+1}_+} &|\eta(x, v(t,x)) - \eta(x,k)|\psi_{,t}(t,x)\, dx\, dt \\ &+ \int_{\mathbb{R}^{d+1}_+} \operatorname{sgn}(v(t,x) - k)(\mathsf{G}(v(t,x)) - \mathsf{G}(k)) \cdot \nabla\psi(t,x)\, dx\, dt \\ &+ \int |u_0(x) - \eta(x,k)|\psi(0,x) dx \geq 0.\end{aligned} \tag{A.4}$$

(N3) *For all* $k \in \mathbb{R}$ *and all nonnegative* $\psi \in \mathcal{D}(\mathbb{R}^{d+1})$ *there holds*

$$\begin{aligned}\int_{\mathbb{R}^{d+1}_+} &\operatorname{sgn}(U(g(t,x)) - U(k))(\mathsf{G}(U(g(t,x))) - \mathsf{G}(U(k))) \cdot \nabla\psi(t,x)\, dx\, dt \\ &+ \int_{\mathbb{R}^{d+1}_+} |\eta(x, U(g(t,x))) - \eta(x, U(k))|\psi_{,t}\, dx\, dt \\ &+ \int_{\mathbb{R}^d} |u_0(x) - \eta(x, U(k))|\psi(0,x)\, dx \geq 0.\end{aligned} \tag{A.5}$$

Remark A.1. For the proof of the main result we only need the direction **(N1)**$\Rightarrow$ **(N3)**. For this direction in the proof of Lemma A.1 we only need that U is Lipschitz.

Proof. [Proof of Lemma A.1] Showing the equivalence of **(N2)** and **(N3)** is obvious. Indeed, if for any k we define $\tilde{k} := U(k)$, we can use $\tilde{k}$ in (A.4) and by using the definition of g we get exactly (A.5). The opposite implication is proved in the same way since U is one-to-one mapping.

To show **(N1)** $\Rightarrow$ **(N2)** consider the equation

$$(u_i)_{,t} + \operatorname{div} \mathsf{F}(x, u_i) = f_i, \quad i = 1, 2.$$

For any two entropy weak solutions u_1, u_2 the so-called Kato inequality holds

$$\begin{aligned} &|u_1 - u_2|_{,t} + \operatorname{div}\left(\operatorname{sgn}(u_1 - u_2)(\mathsf{F}(x, u_1) - \mathsf{F}(x, u_2))\right) \\ &\qquad \leq \operatorname{sgn}(u_1 - u_2)(f_1 - f_2) + |f_1 - f_2|\chi_{\{u_1 = u_2\}} \end{aligned} \tag{A.6}$$

in $\mathcal{D}'(\mathbb{R}^{d+1}_+)$, cf. Ref. 23. Choosing in (A.6) $u_1 = \eta(x, v_1)$ and $u_2 = \eta(x, k)$ with $f_1 = f_2 \equiv 0$. Note that the set of ks such that $|\{(t, x) : u_1(t, x) = \eta(x, k)\}| > 0$ is at most countable and hence it allows to pass from (A.6) to **(N2)**.

For showing the opposite direction let us consider the problem with $f_i : \mathbb{R}^{d+1}_+ \times \mathbb{R} \to \mathbb{R}$ satisfying Lipschitz condition with respect to the last variable.

$$\eta(x, v_i) + \operatorname{div} \mathsf{G}(v_i) = f_i, \quad i = 1, 2. \tag{A.7}$$

First we shall show that for any entropy weak solutions v_1, v_2 in the sense of **(N2)** it holds

$$\begin{aligned} &|\eta(x, v_1) - \eta(x, v_2)|_{,t} + \operatorname{div}\left(\operatorname{sgn}(v_1 - v_2)(\mathsf{G}(v_1) - \mathsf{G}(v_2))\right) \\ &\qquad\qquad \leq \operatorname{sgn}(v_1 - v_2)(f_1 - f_2) \end{aligned} \tag{A.8}$$

One method to prove (A.8) is the method of doubling the variables. The other approach allows to prove this inequality through the averaged contraction principle. We shall use the latter method, since it allows to obtain (A.8) only by extending the proof of Lemma 2.1. Assume that ν, σ are two local entropy measure-valued solutions to (A.7) with a right-hand side f_1 and f_2 respectively. Moreover, let $(E, \mathbf{Q})$ be an entropy-entropy flux pair[h], with $E \in \mathcal{C}^1(\mathbb{R})$ and even. Then following the steps of Lemma 2.1 one can show that

$$\begin{aligned} &\int_{\mathbb{R}^{d+1}_+} \langle E(\eta(x, \lambda) - \eta(x, \mu)), \nu_{(t,x)}(\lambda) \otimes \sigma_{(t,x)}(\mu) \rangle \psi_t(t, x)\, dx\, dt \\ &\quad + \int_{\mathbb{R}^{d+1}_+} \langle \mathbf{Q}(\lambda, \mu), \nu_{(t,x)}(\lambda) \otimes \sigma_{(t,x)}(\mu) \rangle \cdot \nabla \psi(t, x)\, dx\, dt \\ &\quad + \int_{\mathbb{R}^{d+1}_+} \langle E'(\lambda - \mu)(f_1(t, x, \lambda) - f_2(t, x, \mu)), \nu_{(t,x)}(\lambda) \otimes \sigma_{(t,x)}(\mu) \rangle \psi(t, x)\, dx\, dt \geq 0 \end{aligned} \tag{A.9}$$

[h]We say that $(E, \mathbf{Q})$ is an entropy-entropy flux if E is an arbitrary $\mathcal{C}^1$ function (entropy) and $\mathbf{Q}$ (flux) satisfies $\partial_u \mathbf{Q}(x, u) = E'(u) \partial_u \mathbf{F}(x, u)$.

for all nonnegative $\psi \in \mathcal{D}(\mathbb{R}^{d+1}_+)$. Let v_1, v_2 be two entropy weak solutions to (A.7), then obviously $\nu_{(t,x)} = \delta_{v_1(t,x)}$ and $\sigma_{(t,x)} = \delta_{v_2(t,x)}$ are corresponding measure-valued solutions, which we insert into (A.9) and get

$$\begin{aligned}\int_{\mathbb{R}^{d+1}_+} E(\eta(x,v_1)-\eta(x,v_2))\psi_t(t,x)\,dx\,dt + \int_{\mathbb{R}^{d+1}_+} \mathsf{Q}(v_1,v_2)\cdot\nabla\psi(t,x)\,dx\,dt \\ + \int_{\mathbb{R}^{d+1}_+} \langle E'(v_1-v_2)(f_1(t,x,v_1)-f_2(t,x,v_2))\psi(t,x)\,dx\,dt \geq 0\end{aligned} \tag{A.10}$$

In (A.10) we choose

$$E = E_\gamma(\xi) := \begin{cases} \xi^2/4\gamma, & |\xi| \leq 2\gamma, \\ |\xi|-\gamma, & |\xi| \geq 2\gamma, \end{cases} \tag{A.11}$$

and pass with $\gamma \to 0$ to obtain

$$\begin{aligned}&\int_{\mathbb{R}^{d+1}_+} |\eta(x,v_1)-\eta(x,v_2)|\psi_t(t,x)\,dx\,dt \\ &\quad + \int_{\mathbb{R}^{d+1}_+} \operatorname{sgn}(v_1-v_2)(\mathsf{G}(v_1)-\mathsf{G}(v_2))\cdot\nabla\psi(t,x)\,dx\,dt \\ &\quad + \int_{\mathbb{R}^{d+1}_+} \langle \operatorname{sgn}(v_1-v_2)(f_1(t,x,v_1)-f_2(t,x,v_2))\psi(t,x)\,dx\,dt \geq 0\end{aligned} \tag{A.12}$$

for all nonnegative $\psi \in \mathcal{D}(\mathbb{R}^{d+1}_+)$. For passing from (A.8) to **(N1)** we choose again $u_1 = \eta(x,v_1)$ and now $v_2 = \theta(x,k)$ with $f_1 \equiv 0$ and $f_2 = \operatorname{div}\mathsf{G}(v_2) = \operatorname{div}\mathsf{F}(x,k)$. □

B. Independence of solutions on the representation of the flux $\mathsf{F}(x,u) = \mathsf{G}(\theta(x,u))$

In the following section we shall consider whether different choice of G and θ can influence the solution notion that we consider.

Theorem B.1. *Let $\mathsf{F} = \mathsf{G}_1 \circ \theta_1 = \mathsf{G}_2 \circ \theta_2$ with G_i and θ_i satisfying **(A1)–(A4)** and let $u_0 \in L^1(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$. In addition, assume that for all $u \in \mathbb{R}$ the function $\theta_1(x, \eta_2(x,u))$ has a Darboux[i] property. Then the unique entropy weak solution to* (1.1) *related to* (G_1, θ_1) *is equal to the unique entropy weak solution to* (1.1) *related to* (G_2, θ_2).

We shall not provide the whole proof of this theorem, but focus on a key estimate formulated in Lemma B.1. Passing from the below stated contraction property to the uniqueness property follows the procedures developed in the present paper.

Remark B.1. Note that the example presented by Panov[33] does not fit into the framework of Theorem B.1. We recall the details. He considered the equation $u_t +$

[i]We say that $f(x)$ has Darboux property if for any $x, y \in \mathbb{R}^d$ such that $f(x) < f(y)$ and any $u \in (f(x), f(y))$ there exists $x_0 \in \mathbb{R}^d$ such that $f(x_0) = u$.

$F(u)_x = 0$ with the representation of the flux $F(u) = G(\theta(x,u))$ as follows $G_1(u) = F(u)$ and $\theta_1(x,u) = u$. If one assumes $F(0) = F(1) = 0$ the equation can be considered as the particular case of the presented there model equation

$$G_2(\theta) = \begin{cases} f(\theta), & \theta \le 1 \\ f(\theta-1), & \theta > 1 \end{cases}, \quad \theta_2(x,u) = u + H(x),$$

where H is the Heaviside function. Then if $F(c) > 0$ for some point $c \in (0,1)$ then the constant solution $u \equiv c$ is a en entropy solution in the sense defined by Panov, but it is not an entropy solution in the sense of Audusse-Perthame. The theorem presented below has an assumption that the function $\theta_1(x, \eta_2(x,u))$ has a Darboux property, whereas in the example one has $\theta_2(x, \eta_1(x,u)) = \theta_2(x,u)$.

Lemma B.1. *Let $\mathsf{F} = \mathsf{G}_1 \circ \theta_1 = \mathsf{G}_2 \circ \theta_2$ with (G_1, θ_1) and (G_2, θ_2) satisfying **(A1)**–**(A3)** and let (A_1, U_1), (A_2, U_2) be admissible parametrization of G_1 and G_2 respectively. Assume that ν and σ are two local entropy measure-valued solutions to* (1.1) *corresponding to (A_1, U_1), (A_2, U_2) respectively the in sense of Definition 2.1. Moreover, assume that for all $u \in \mathbb{R}$ the function $\theta_1(x, \eta_2(x,u))$ has a Darboux property. Then*

$$\begin{aligned} &\int_{\mathbb{R}^{d+1}_+} \langle |\eta_1(x, U_1(\lambda)) - \eta_2(x, U_2(\mu))|, \nu_{(t,x)}(\lambda) \otimes \sigma_{(t,x)}(\mu)\rangle \psi_{,t}(t,x)\, dx\, dt \\ &\quad + \int_{\mathbb{R}^{d+1}_+} \langle \mathsf{Q}(x,\lambda,\mu), \nu_{(t,x)}(\lambda) \otimes \sigma_{(t,x)}(\mu)\rangle \cdot \nabla \psi(t,x)\, dx\, dt \ge 0 \end{aligned} \tag{B.1}$$

for all nonnegative $\psi \in \mathcal{D}(\mathbb{R}^{d+1}_+)$. Here, we defined Q through

$$\mathsf{Q}(x,\lambda,\mu) := \begin{cases} (\mathsf{A}_1(\lambda) - \mathsf{A}_2(\mu)) \operatorname{sgn}\left(\dfrac{\lambda - \alpha^1_k}{\beta^1_k - \alpha^1_k} - \dfrac{\mu - \alpha^2_k}{\beta^2_k - \alpha^2_k}\right) \\ \qquad \text{if there is } k \text{ such that } \lambda \in [\alpha^1_k, \beta^1_k] \text{ or } \mu \in [\alpha^2_k, \beta^2_k], \\ (\mathsf{A}_1(\lambda) - \mathsf{A}_2(\mu)) \operatorname{sgn}(\eta_1(x, U_1(\lambda)) - \eta_2(x, U_2(\mu))) \text{ otherwise.} \end{cases} \tag{B.2}$$

The numbers α^i_k, β^i_k, $i = 1, 2$ are defined in (1.16) *correspond to U_1, U_2 respectively. Moreover, for fixed λ, μ the function $x \mapsto \mathsf{Q}(x,\lambda,\mu)$ is constant.*

Proof. The proof follows almost step by step the proof of Lemma 2.1 and is again inspired by Ref. 10. Therefore we skip here all details and focus only on possible differences.

Step 1: Independence of Q on x. The first key observation is a "curiosity" that Q does not depend on x. Hence let λ and μ be fixed. We show that then $\mathsf{Q}(x,\lambda,\mu)$ is constant. Clearly, it is a consequence of the definition that it may depend on x only if $\lambda \in (\beta^1_k, \alpha^1_{k+1})$ and $\mu \in (\beta^2_n, \alpha^2_{n+1})$. In case that $\mathsf{A}_1(\lambda) = \mathsf{A}_2(\mu)$, we see that Q is identically zero (hence x-independent). Thus, consider the case $\mathsf{A}_1(\lambda) \ne \mathsf{A}_2(\mu)$ and assume a contradiction, i.e., let there exist $x \in \mathbb{R}^d$ and $y \in \mathbb{R}^d$ such that

$$\begin{aligned} \eta_1(x, U_1(\lambda)) &> \eta_2(x, U_2(\mu)), \\ \eta_1(y, U_1(\lambda)) &< \eta_2(y, U_2(\mu)). \end{aligned} \tag{B.3}$$

Since, θ_1 is strictly increasing we deduce from (B.3) that

$$\begin{aligned} U_1(\lambda) > \theta_1(x, \eta_2(x, U_2(\mu))), \\ U_1(\lambda) < \theta_1(y, \eta_2(y, U_2(\mu))). \end{aligned} \tag{B.4}$$

As $\theta_1(x, \eta_2(x, U_2(\mu)))$ has a Darboux property, there surely exists $x_0 \in \mathbb{R}^d$ such that

$$U_1(\lambda) = \theta_1(x_0, \eta_2(x_0, U_2(\mu))). \tag{B.5}$$

But since λ and μ belong to the intervals where U_1 and U_2 are strictly increasing we can use (B.7) to deduce

$$\begin{aligned} \mathsf{A}_1(\lambda) = \mathsf{G}_1(U_1(\lambda)) &\overset{(B.5)}{=} \mathsf{G}_1(\theta_1(x_0, \eta_2(x_0, U_2(\mu)))) \\ &= \mathsf{G}_2(\theta_2(x_0, \eta_2(x_0, U_2(\mu)))) = \mathsf{G}_2(U_2(\mu)) = \mathsf{A}_2(\mu), \end{aligned} \tag{B.6}$$

which is a contradiction and finishes the prove of independence of Q on x. We shall use this fact in further parts of the proof and in what follows to avoid unclarity, we write $\mathsf{Q}(\lambda, \mu)$.

First, we recall some properties of a parametrization of G_i. Thus, recalling (2.1) we have

$$\mathsf{A}_i(s) = \begin{cases} \mathsf{G}_i(U_i(s)) & \text{for } s \in (\beta_k^i, \alpha_{k+1}^i), \\ \dfrac{(\mathsf{G}_i)_+(z_k^i) - (\mathsf{G}_i)_-(z_k^i)}{\beta_k^i - \alpha_k^i}(s - \alpha_k^i) + (\mathsf{G}_i)_-(z_k^i) & \text{for } s \in [\alpha_k^i, \beta_k^i]. \end{cases} \tag{B.7}$$

Moreover, here z_k^i denote the points of discontinuities of G_i and there is one to one relation between z_k^1 and z_k^2 since F is the same for both G_1 and G_2. In addition we know that U_i is strictly increasing on $(\beta_k^i, \alpha_{k+1}^i)$ and constant on $[\alpha_k^i, \beta_k^i]$. Next, we discuss more precise relation between intervals $(\beta_k^i, \alpha_{k+1}^i)$ for different i's. For this purpose we define Carethéodory mappings v_i as

$$\begin{aligned} v_1(x, s) &:= \theta_1(x, \eta_2(x, U_2(s))), \\ v_2(x, s) &:= \theta_2(x, \eta_1(x, U_1(s))). \end{aligned} \tag{B.8}$$

First, we show that for almost all fixed $x \in \mathbb{R}^d$ the mapping

$$v_1(x, \cdot) : [\beta_k^2, \alpha_{k+1}^2] \mapsto [U_1(\beta_k^1), U_1(\alpha_{k+1}^1)]$$

is bijection and also

$$v_2(x, \cdot) : [\beta_k^1, \alpha_{k+1}^1] \mapsto [U_2(\beta_k^2), U_2(\alpha_{k+1}^2)].$$

Note that we use convention $\beta_0 := -\infty$. Since, θ_1, η_1 are strictly increasing and U_1 is strictly increasing on each $(\beta_k^1, \alpha_{k+1}^1)$ and U_2 is strictly increasing on $(\beta_k^2, \alpha_{k+1}^2)$ as well, it is enough to show that for all k

$$\lim_{s \to (\alpha_k^2)_-} v_1(x, s) = \lim_{s \to (\alpha_k^1)_-} U_1(s), \qquad \lim_{s \to (\beta_k^2)_+} v_1(x, s) = \lim_{s \to (\beta_k^1)_+} U_1(s). \tag{B.9}$$

However, using the fact that $\mathsf{G}_1(v_1(x,s)) = \mathsf{G}_2(U_2(s))$, which follows from definition of v_i and the fact that F is the same for both G_1 and G_2, we see that (B.9) holds.

Consequently, using (B.9) and the fact that U_i is strictly increasing on each $(\beta_k^i, \alpha_{k+1}^i)$ and constant on $[\alpha_k^i, \beta_k^i]$, we can define Carathéodory mappings w_i that are for almost all fixed $x \in \mathbb{R}^d$ bijections by the relation

$$\begin{aligned} w_1(x,s) &:= \begin{cases} U_1^{-1}(v_1(x,s)) & \text{for } s \in (\beta_k^2, \alpha_{k+1}^2), \\ \dfrac{\alpha_k^1(s-\beta_k^2)}{\alpha_k^2-\beta_k^2} - \dfrac{\beta_k^1(s-\alpha_k^2)}{\alpha_k^2-\beta_k^2} & \text{for } s \in [\alpha_k^2, \beta_k^2]. \end{cases} \\ w_2(x,s) &:= \begin{cases} U_2^{-1}(v_2(x,s)) & \text{for } s \in (\beta_k^1, \alpha_{k+1}^1), \\ \dfrac{\alpha_k^2(s-\beta_k^1)}{\alpha_k^1-\beta_k^1} - \dfrac{\beta_k^2(s-\alpha_k^1)}{\alpha_k^1-\beta_k^1} & \text{for } s \in [\alpha_k^1, \beta_k^1]. \end{cases} \end{aligned} \tag{B.10}$$

Having these auxiliary results, we now prove the lemma. Following the proof of Lemma 2.1, we can deduce in the same manner that (we keep the notation for the mollification of measures ν and σ and use the inequalities (2.11)–(2.12)) that for any $\mu \in \mathbb{R}$ and all $(t,x) \in (\delta,\infty) \times \mathbb{R}^d$ there holds

$$\begin{aligned} &\Big(\omega_2^\varepsilon * \langle |\eta_1(x,U_1(\lambda)) - \eta_1(x,U_1(\mu))|, \nu_{(t,x)}^\delta(\lambda)\rangle\Big)_{,t} \\ &+ \operatorname{div}\Big(\langle (\mathsf{A}_1(\lambda) - \mathsf{A}_1(\mu)) \operatorname{sgn}(\lambda-\mu), \nu_{(t,x)}^{\delta,\varepsilon}(\lambda)\rangle\Big) \le 0 \end{aligned} \tag{B.11}$$

and for any $\lambda \in \mathbb{R}$ and all $(t,x) \in (\delta,\infty) \times \mathbb{R}^d$ we have

$$\begin{aligned} &\Big(\omega_2^\varepsilon * \langle |\eta_2(x,U_2(\lambda)) - \eta_2(x,U_2(\mu))|, \sigma_{(t,x)}^\delta(\mu)\rangle\Big)_{,t} \\ &+ \operatorname{div}\Big(\langle (\mathsf{A}_2(\lambda) - \mathsf{A}_2(\mu)) \operatorname{sgn}(\lambda-\mu), \sigma_{(t,x)}^{\delta,\varepsilon}(\mu)\rangle\Big) \le 0. \end{aligned} \tag{B.12}$$

Since w_i are Carathéodory functions we can use the push-forward formula and define Young measures $\tilde{\nu}$ and $\tilde{\sigma}$ as

$$\begin{aligned} \langle f(\mu), \tilde{\sigma}_{(t,x)}^{\delta,\varepsilon}(\mu)\rangle &:= \langle f(w_1(x,\mu)), \sigma_{(t,x)}^{\delta,\varepsilon}(\mu)\rangle, \\ \langle f(\mu), \tilde{\nu}_{(t,x)}^{\delta,\varepsilon}(\lambda)\rangle &:= \langle f(w_2(x,\lambda)), \nu_{(t,x)}^{\delta,\varepsilon}(\lambda)\rangle. \end{aligned} \tag{B.13}$$

For any fixed (t,x) we apply $\tilde{\sigma}_{(t,x)}$ onto (B.11) to get

$$\begin{aligned} &\Big\langle \Big(\omega_2^\varepsilon * \langle |\eta_1(x,U_1(\lambda)) - \eta_1(x,U_1(\mu))|, \nu_{(t,x)}^\delta(\lambda)\rangle\Big)_{,t}, \tilde{\sigma}_{(t,x)}^{\delta,\varepsilon}(\mu)\Big\rangle \\ &+ \Big\langle \operatorname{div}\Big(\langle (\mathsf{A}_1(\lambda) - \mathsf{A}_1(\mu)) \operatorname{sgn}(\lambda-\tilde{\mu}), \nu_{(t,x)}^{\delta,\varepsilon}(\lambda)\rangle\Big), \tilde{\sigma}_{(t,x)}^{\delta,\varepsilon}(\mu)\Big\rangle \le 0 \end{aligned} \tag{B.14}$$

and focus on the second term. The operator div is acting on the measure $\nu^{\delta,\varepsilon}$, hence using (B.13) we deduce that

$$\begin{aligned} &\Big\langle \operatorname{div}\Big(\langle (\mathsf{A}_1(\lambda) - \mathsf{A}_1(\mu)) \operatorname{sgn}(\lambda-\mu), \nu_{(t,x)}^{\delta,\varepsilon}(\lambda)\rangle\Big), \tilde{\sigma}_{(t,x)}^{\delta,\varepsilon}(\mu)\Big\rangle \\ &= \Big\langle \langle (\mathsf{A}_1(\lambda) - \mathsf{A}_1(\mu)) \operatorname{sgn}(\lambda-\mu), \nabla\nu_{(t,x)}^{\delta,\varepsilon}(\lambda)\rangle, \tilde{\sigma}_{(t,x)}^{\delta,\varepsilon}(\mu)\Big\rangle \\ &= \Big\langle \langle (\mathsf{A}_1(\lambda) - \mathsf{A}_1(w_1(x,\mu))) \operatorname{sgn}(\lambda-w_1(x,\mu)), \nabla\nu_{(t,x)}^{\delta,\varepsilon}(\lambda)\rangle, \sigma_{(t,x)}^{\delta,\varepsilon}(\mu)\Big\rangle \end{aligned} \tag{B.15}$$

Next, we evaluate the terms depending on w_1. Hence using (B.8), (B.10), (B.7) and strict monotonicity of θ_i, η_i and U_i we get that for all $\mu \in (\beta_k^2, \alpha_{k+1}^2)$

$$\begin{aligned} \mathsf{A}_1(w_1(x,\mu)) &= \mathsf{G}_1(v_1(x,\mu)) = \mathsf{G}_1(\theta_1(x,\eta_2(x,U_2(\mu)))) \\ &= \mathsf{G}_2(\theta_2(x,\eta_2(x,U_2(\mu)))) = \mathsf{G}_2(U_2(\mu)) = \mathsf{A}_2(\mu), \\ \operatorname{sgn}(\lambda - w_1(x,\mu)) &= \operatorname{sgn}(U_1(\lambda) - v_1(x,\mu)) \\ &= \operatorname{sgn}(\eta_1(x,U_1(\lambda)) - \eta_2(x,U_2(\mu))). \end{aligned} \tag{B.16}$$

Moreover, using (B.10), we see from (B.16) that if $\mu \in (\beta_k^2, \alpha_{k+1}^2)$ and $\lambda \in [\alpha_n^1, \beta_n^1]$ for some n then we directly obtain

$$\operatorname{sgn}(\lambda - w_1(x,\mu)) = \operatorname{sgn}\left(\frac{\lambda - \alpha_k^1}{\beta_k^1 - \alpha_k^1} - \frac{s - \alpha_k^2}{\beta_k^2 - \alpha_k^2}\right). \tag{B.17}$$

On the other hand, for $\mu \in [\alpha_k^2, \beta_k^2]$, we know from (B.10) that $w_1(x,\mu) \in [\alpha_k^1, \beta_k^1]$. Consequently, we can use (B.7) and (B.10) and by simple algebraic manipulation we deduce that

$$\begin{aligned} \mathsf{A}_1(w_1(x,\mu)) &= \frac{(\mathsf{G}_1)_+(z_k^1) - (\mathsf{G}_1)_-(z_k^1)}{\beta_k^1 - \alpha_k^1}(w_1(x,\mu) - \alpha_k^1) + (\mathsf{G}_1)_-(z_k^1) \\ &= \frac{(\mathsf{G}_2)_+(z_k^2) - (\mathsf{G}_2)_-(z_k^2)}{\beta_k^1 - \alpha_k^1}(w_1(x,\mu) - \alpha_k^1) + (\mathsf{G}_2)_-(z_k^2) \\ &= \frac{(\mathsf{G}_2)_+(z_k^2) - (\mathsf{G}_2)_-(z_k^2)}{\beta_k^2 - \alpha_k^2}(\mu - \alpha_k^2) + (\mathsf{G}_2)_-(z_k^2) \\ &= \mathsf{A}_2(\mu), \\ \operatorname{sgn}(\lambda - w_1(x,\mu)) &= \operatorname{sgn}\left(\lambda - \alpha_k^1 - \frac{(\alpha_k^1 - \beta_k^1)s - (\alpha_k^1 - \beta_k^1)\alpha_k^2}{\alpha_k^2 - \beta_k^2}\right) \\ &= \operatorname{sgn}\left(\frac{\lambda - \alpha_k^1}{\beta_k^1 - \alpha_k^1} - \frac{s - \alpha_k^2}{\beta_k^2 - \alpha_k^2}\right). \end{aligned} \tag{B.18}$$

Thus, inserting (B.16)–(B.18) into (B.15) and using the definition of Q we get

$$\begin{aligned} &\left\langle \operatorname{div}\left(\langle(\mathsf{A}_1(\lambda) - \mathsf{A}_1(\mu))\operatorname{sgn}(\lambda-\mu), \nu^{\delta,\varepsilon}_{(t,x)}(\lambda)\rangle\right), \tilde{\sigma}^{\delta,\varepsilon}_{(t,x)}(\mu)\right\rangle \\ &= \left\langle \langle(\mathsf{Q}(\lambda,\mu)), \nabla\nu^{\delta,\varepsilon}_{(t,x)}(\lambda)\rangle, \sigma^{\delta,\varepsilon}_{(t,x)}(\mu)\right\rangle \end{aligned} \tag{B.19}$$

Thus, since Q does not depend on x we can rewrite (B.19) as

$$\begin{aligned} &\left\langle \operatorname{div}\left(\langle(\mathsf{A}_1(\lambda) - \mathsf{A}_1(\mu))\operatorname{sgn}(\lambda-\mu), \nu^{\delta,\varepsilon}_{(t,x)}(\lambda)\rangle\right), \tilde{\sigma}^{\delta,\varepsilon}_{(t,x)}(\mu)\right\rangle \\ &= \left\langle \operatorname{div}\left(\langle\mathsf{Q}(\lambda,\mu), \nu^{\delta,\varepsilon}_{(t,x)}(\lambda)\rangle\right), \sigma^{\delta,\varepsilon}_{(t,x)}(\mu)\right\rangle \end{aligned} \tag{B.20}$$

Thus, substituting this into (B.14) we obtain

$$\begin{aligned} &\left\langle \left(\omega_2^\varepsilon * \langle|\eta_1(x,U_1(\lambda)) - \eta_1(x,U_1(\mu))|, \nu^{\delta}_{(t,x)}(\lambda)\rangle\right)_{,t}, \tilde{\sigma}^{\delta,\varepsilon}_{(t,x)}(\mu)\right\rangle \\ &+ \left\langle \operatorname{div}\left(\langle\mathsf{Q}(\lambda,\mu), \nu^{\delta,\varepsilon}_{(t,x)}(\lambda)\rangle\right), \sigma^{\delta,\varepsilon}_{(t,x)}(\mu)\right\rangle \le 0 \end{aligned} \tag{B.21}$$

and in the same spirit we can deduce that

$$\begin{aligned}&\left\langle \left(\omega_2^\varepsilon * \langle|\eta_2(x,U_2(\lambda))-\eta_2(x,U_2(\mu))|,\sigma^\delta_{(t,x)}(\mu)\rangle\right)_{,t},\tilde{\nu}^{\delta,\varepsilon}_{(t,x)}(\lambda)\right\rangle\\&+\left\langle \operatorname{div}\left(\langle \mathbf{Q}(\lambda,\mu),\sigma^{\delta,\varepsilon}_{(t,x)}(\mu)\rangle\right),\nu^{\delta,\varepsilon}_{(t,x)}(\lambda)\right\rangle\le 0\end{aligned} \tag{B.22}$$

Summing (B.21) and (B.22) and recalling the same procedure as in Lemma 2.1 we deduce that

$$\begin{aligned}&\left\langle \left(\omega_2^\varepsilon * \langle|\eta_1(x,U_1(\lambda))-\eta_1(x,U_1(\mu))|,\nu^\delta_{(t,x)}(\lambda)\rangle\right)_{,t},\tilde{\sigma}^{\delta,\varepsilon}_{(t,x)}(\mu)\right\rangle\\&+\left\langle \left(\omega_2^\varepsilon * \langle|\eta_2(x,U_2(\lambda))-\eta_2(x,U_2(\mu))|,\sigma^\delta_{(t,x)}(\mu)\rangle\right)_{,t},\tilde{\nu}^{\delta,\varepsilon}_{(t,x)}(\lambda)\right\rangle\\&+\operatorname{div}\left\langle \mathbf{Q}(\lambda,\mu),\nu^{\delta,\varepsilon}_{(t,x)}(\lambda)\otimes\sigma^{\delta,\varepsilon}_{(t,x)}(\mu)\right\rangle\le 0\end{aligned} \tag{B.23}$$

Hence, we can let $\varepsilon\to 0_+$ to conclude that (see the proof of Lemma 2.1 for details) for all $t>\delta$ and all nonnegative $\psi\in\mathcal{D}(\mathbb{R}^d)$

$$\begin{aligned}&\int_{\mathbb{R}^d}\left\langle \left(\langle|\eta_1(x,U_1(\lambda))-\eta_1(x,U_1(\mu))|,\nu^\delta_{(t,x)}(\lambda)\rangle\right)_{,t},\tilde{\sigma}^{\delta}_{(t,x)}(\mu)\right\rangle\psi(x)\,dx\\&+\int_{\mathbb{R}^d}\left\langle \left(\langle|\eta_2(x,U_2(\lambda))-\eta_2(x,U_2(\mu))|,\sigma^\delta_{(t,x)}(\mu)\rangle\right)_{,t},\tilde{\nu}^{\delta}_{(t,x)}(\lambda)\right\rangle\psi(x)\,dx\\&\le\int_{\mathbb{R}^d}\left\langle \mathbf{Q}(\lambda,\mu),\nu^{\delta}_{(t,x)}(\lambda)\otimes\sigma^{\delta}_{(t,x)}(\mu)\right\rangle\cdot\nabla\psi(x)\,dx,\end{aligned} \tag{B.24}$$

where $\tilde{\nu}^\delta$ and $\tilde{\sigma}^\delta$ are defined through

$$\begin{aligned}\langle f(\mu),\tilde{\sigma}^\delta_{(t,x)}(\mu)\rangle &:= \langle f(w_1(x,\mu)),\sigma^\delta_{(t,x)}(\mu)\rangle,\\ \langle f(\mu),\tilde{\nu}^\delta_{(t,x)}(\lambda)\rangle &:= \langle f(w_2(x,\lambda)),\nu^\delta_{(t,x)}(\lambda)\rangle.\end{aligned} \tag{B.25}$$

Finally, we evaluate $\eta_1(x,U_1(w_1(x,\mu)))$. First, for $\mu\in(\beta_k^2,\alpha_{k+1}^2)$, by (B.10) we get

$$\eta_1(x,U_1(w_1(x,\mu)))=\eta_1(x,\theta_1(x,\eta_2(x,U_2(\mu))))=\eta_2(x,U_2(\mu)). \tag{B.26}$$

On the other hand, if $\mu\in[\alpha_k^2,\beta_k^2]$ we can again use (B.8) and (B.10) to conclude that for all μ

$$\eta_1(x,U_1(w_1(x,\mu)))=\eta_2(x,U_2(\mu)). \tag{B.27}$$

Therefore (B.24) reduces to

$$\begin{aligned}&\int_{\mathbb{R}^d}\left\langle \left(\langle|\eta_1(x,U_1(\lambda))-\eta_2(x,U_2(\mu))|,\nu^\delta_{(t,x)}(\lambda)\rangle\right)_{,t},\sigma^{\delta}_{(t,x)}(\mu)\right\rangle\psi(x)\,dx\\&+\int_{\mathbb{R}^d}\left\langle \left(\langle|\eta_1(x,U_1(\lambda))-\eta_2(x,U_2(\mu))|,\sigma^\delta_{(t,x)}(\mu)\rangle\right)_{,t},\nu^{\delta}_{(t,x)}(\lambda)\right\rangle\psi(x)\,dx\\&\le\int_{\mathbb{R}^d}\left\langle \mathbf{Q}(\lambda,\mu),\nu^{\delta}_{(t,x)}(\lambda)\otimes\sigma^{\delta}_{(t,x)}(\mu)\right\rangle\cdot\nabla\psi(x)\,dx,\end{aligned} \tag{B.28}$$

which implies

$$\begin{aligned}&\frac{d}{dt}\int_{\mathbb{R}^d}\left\langle|\eta_1(x,U_1(\lambda))-\eta_2(x,U_2(\mu))|,\nu^\delta_{(t,x)}(\lambda)\otimes\sigma^\delta_{(t,x)}(\mu)\right\rangle\psi(x)\;dx\\&\le\int_{\mathbb{R}^d}\left\langle\mathbf{Q}(\lambda,\mu),\nu^\delta_{(t,x)}(\lambda)\otimes\sigma^\delta_{(t,x)}(\mu)\right\rangle\cdot\nabla\psi(x)\;dx.\end{aligned}\tag{B.29}$$

Consequently, letting $\delta\to 0_+$ and following again the proof of Lemma 2.1, we observe (B.1). Thus, the proof is complete. □

C. Auxiliary results

Proposition C.1. *Let $[a,b]\subset\mathbb{R}$ and let f be continuous, f,f_n be monotone functions such that $f_n\to f$ pointwisely. Then $f_n\to f$ uniformly on $[a,b]$.*

The above fact in an elementary exercise. For the proof see e.g. Ref. 3.

Proposition C.2. *Let $[a,b]\subset\mathbb{R}$ and let $f,f_n:\mathbb{R}\to\mathbb{R}$, $\mathrm{Im}\,(f)=\mathbb{R}$, $\mathrm{Im}\,(f_n)=\mathbb{R}$, f,f_n be strictly monotone functions such that $f_n\to f$ point-wisely. Then the inverse functions converge locally uniformly to the inverse of the limit, namely $(f_n)^{-1}\to f^{-1}$ uniformly on every compact subset of $\mathbb{R}$.*

Proof. We provide the proof by contradiction. Assume that f_n converges uniformly to f and that $(f_n)^{-1}$ does not converge point-wisely to f^{-1}. Hence there exists $y,\varepsilon>0$ and a subsequence $(f_{n_k})^{-1}$ such that

$$(f_{n_k})^{-1}(y)\notin[f^{-1}(y)-\varepsilon,f^{-1}(y)+\varepsilon].\tag{C.1}$$

We only prove the case $(f_{n_k})^{-1}(y)>f^{-1}(y)+\varepsilon$. The second case follows analogously. Let $\bar{y}_{n_k}:=f^{-1}_{n_k}(f(y))$. By (C.1) we have the estimate

$$\bar{y}_{n_k}>y+\varepsilon.$$

Using the strict monotonicity of f, monotonicity of f_{n_k} and the definition of $\bar{y}_{n_k}$ we conclude an existence of δ such that

$$0<\delta\le f(y+\varepsilon)-f(y)=f(y+\varepsilon)-f_{n_k}(\bar{y}_{n_k})\le f(y+\varepsilon)-f_{n_k}(y+\varepsilon)\tag{C.2}$$

which contradicts the uniform convergence of f_n. Hence $(f_n)^{-1}$ converges point-wisely to f^{-1}. The uniform convergence of $(f_n)^{-1}$ can be concluded by Proposition C.1. □

Acknowledgment

The authors are grateful to the referee for numerous suggestions which influenced the final version of the paper, and in particular indicating different and shorter method of proving Lemma A.1.

References

1. Adimurthi, J. Jaffré, and G. Veerappa Gowda. Godunov-type methods for conservation laws with a flux function discontinuous in space. *SIAM J. Numer. Anal.*, 42(1):179–208, 2004.
2. Adimurthi, S. Mishra, and G. D. V. Gowda. Optimal entropy solutions for conservation laws with discontinuous flux-functions. *J. Hyperbolic Differ. Equ.*, 2(4):783–837, 2005.
3. H. Amann and J. Escher. *Analysis I.* Basel: Birkhäuser. xv, 426 p., 2005.
4. K. Ammar and P. Wittbold. On a degenerate scalar conservation law with general boundary condition. *Differential Integral Equations*, 21(3-4):363–386, 2008.
5. B. Andreianov, K. H. Karlsen, and N. H. Risebro. On vanishing viscosity approximation of conservation laws with discontinuous flux. *Netw. Heterog. Media*, 5(3):617–633, 2010.
6. B. Andreianov, K. H. Karlsen, and N. H. Risebro. A theory of L^1-dissipative solvers for scalar conservation laws with discontinuous flux. *Arch. Ration. Mech. Anal.*, 201(1):27–86, 2011.
7. E. Audusse and B. Perthame. Uniqueness for scalar conservation laws with discontinuous flux via adapted entropies. *Proc. Roy. Soc. Edinburgh Sect. A*, 135(2):253–265, 2005.
8. F. Bachmann and J. Vovelle. Existence and uniqueness of entropy solution of scalar conservation laws with a flux function involving discontinuous coefficients. *Commun. Partial Differ. Equations*, 31(1-3):371–395, 2006.
9. P. Baiti and H. K. Jenssen. Well-posedness for a class of 2×2 conservation laws with L^∞ data. *J. Differential Equations*, 140(1):161–185, 1997.
10. M. Bulíček, P. Gwiazda, J. Málek, and A. Świerczewska Gwiazda. On scalar hyperbolic conservation laws with a discontinuous flux. *Math. Models Methods Appl. Sci.*, 21(1):89–113, 2011.
11. R. Bürger, K. H. Karlsen, S. Mishra, and J. D. Towers. On conservation laws with discontinuous flux. In *Trends in applications of mathematics to mechanics*, Ber. Math., pages 75–84. Shaker Verlag, Aachen, 2005.
12. R. Bürger, K. H. Karlsen, and J. D. Towers. An Engquist-Osher-type scheme for conservation laws with discontinuous flux adapted to flux connections. *SIAM J. Numer. Anal.*, 47(3):1684–1712, 2009.
13. C. Cancès. Asymptotic behavior of two-phase flows in heterogeneous porous media for capillarity depending only on space. I. Convergence to the optimal entropy solution. *SIAM J. Math. Anal.*, 42(2):946–971, 2010.
14. C. Cancès. Asymptotic behavior of two-phase flows in heterogeneous porous media for capillarity depending only on space. II. Nonclassical shocks to model oil-trapping. *SIAM J. Math. Anal.*, 42(2):972–995, 2010.
15. C. Cancès. On the effects of discontinuous capillarities for immiscible two-phase flows in porous media made of several rock-types. *Netw. Heterog. Media*, 5(3):635–647, 2010.
16. J. Carrillo. Entropy solutions for nonlinear degenerate problems. *Arch. Ration. Mech. Anal.*, 147(4):269–361, 1999.
17. J. Carrillo. Conservation laws with discontinuous flux functions and boundary condition. *J. Evol. Equ.*, 3(2):283–301, 2003.
18. G.-Q. Chen, N. Even, and C. Klingenberg. Hyperbolic conservation laws with discontinuous fluxes and hydrodynamic limit for particle systems. *J. Differential Equations*, 245(11):3095–3126, 2008.
19. C. M. Dafermos. *Hyperbolic conservation laws in continuum physics. 3rd ed.* Grundlehren der Mathematischen Wissenschaften 325. Berlin: Springer., 2010.
20. A.-L. Dalibard. Kinetic formulation for heterogeneous scalar conservation laws. *Ann.*

Inst. H. Poincaré Anal. Non Linéaire, 23(4):475–498, 2006.
21. S. Diehl. A uniqueness condition for nonlinear convection-diffusion equations with discontinuous coefficients. *J. Hyperbolic Differ. Equ.*, 6(1):127–159, 2009.
22. R. J. DiPerna. Measure-valued solutions to conservation laws. *Arch. Rational Mech. Anal.*, 88(3):223–270, 1985.
23. G. Gagneux and M. Madaune-Tort. *Analyse mathématique de modèles non linéaires de l'ingénierie pétrolière*, volume 22 of *Mathématiques & Applications (Berlin)*. Springer-Verlag, Berlin, 1996.
24. M. Garavello, R. Natalini, B. Piccoli, and A. Terracina. Conservation laws with discontinuous flux. *Netw. Heterog. Media*, 2(1):159–179, 2007.
25. T. Gimse. Conservation laws with discontinuous flux functions. *SIAM J. Math. Anal.*, 24(2):279–289, 1993.
26. J. Jimenez. Mathematical analysis of a scalar multidimensional conservation law with discontinuous flux. *J. Evol. Equ.*, 11(3):553–576, 2011.
27. K. H. Karlsen, M. Rascle, and E. Tadmor. On the existence and compactness of a two-dimensional resonant system of conservation laws. *Commun. Math. Sci.*, 5(2):253–265, 2007.
28. K. H. Karlsen, N. H. Risebro, and J. D. Towers. L^1 stability for entropy solutions of nonlinear degenerate parabolic convection-diffusion equations with discontinuous coefficients. *Skr., K. Nor. Vidensk. Selsk.*, 2003(3):49 p., 2003.
29. S. N. Kružkov. First order quasilinear equations with several independent variables. *Mat. Sb. (N.S.)*, 81 (123):228–255, 1970.
30. S. Martin and J. Vovelle. Convergence of implicit finite volume methods for scalar conservation laws with discontinuous flux function. *M2AN Math. Model. Numer. Anal.*, 42(5):699–727, 2008.
31. D. Mitrovic. Existence and stability of a multidimensional scalar conservation law with discontinuous flux. *Netw. Heterog. Media*, 5(1):163–188, 2010.
32. E. Y. Panov. Existence of strong traces for quasi-solutions of multidimensional conservation laws. *J. Hyperbolic Differ. Equ.*, 4(4):729–770, 2007.
33. E. Y. Panov. On existence and uniqueness of entropy solutions to the Cauchy problem for a conservation law with discontinuous flux. *J. Hyperbolic Differ. Equ.*, 6(3):525–548, 2009.
34. E. Y. Panov. Existence and strong pre-compactness properties for entropy solutions of a first-order quasilinear equation with discontinuous flux. *Arch. Ration. Mech. Anal.*, 195(2):643–673, 2010.
35. K. R. Rajagopal. On implicit constitutive theories. *Appl. Math.*, 48(4):279–319, 2003.
36. N. H. Risebro. An introduction to the theory of scalar conservation laws with spatially discontinuous flux functions. Quak, Ewald (ed.) et al., Applied wave mathematics. Selected topics in solids, fluids, and mathematical methods. Berlin: Springer. 395-464., 2009.
37. A. Szepessy. An existence result for scalar conservation laws using measure valued solutions. *Comm. Partial Differential Equations*, 14(10):1329–1350, 1989.