A vanishing dynamic capillarity limit equation with discontinuous flux

M. Graf, M. Kunzinger, D. Mitrovic and D. Vujadinovic

Abstract. We prove existence and uniqueness of a solution to the Cauchy problem corresponding to the **dynamics capillarity** equation

$$\begin{cases} \partial_t u_{\varepsilon, \delta} + \operatorname{div} f_{\varepsilon, \delta}(\mathbf{x}, u_{\varepsilon, \delta}) = \varepsilon \Delta u_{\varepsilon, \delta} + \delta(\varepsilon) \partial_t \Delta u_{\varepsilon, \delta}, & \mathbf{x} \in M, \quad t \geq 0 \\ u|_{t=0} = u_0(\mathbf{x}). \end{cases}$$

Here, $f_{\varepsilon, \delta}$ and u_0 are smooth functions while ε and $\delta = \delta(\varepsilon)$ are fixed constants. Assuming $f_{\varepsilon, \delta} \rightarrow f \in L^p(\mathbb{R}^d \times \mathbb{R}; \mathbb{R}^d)$ for some $1 < p < \infty$, strongly as $\varepsilon \rightarrow 0$, we prove that, under an appropriate relationship between ε and $\delta(\varepsilon)$ depending on the regularity of the flux f , the sequence of solutions $(u_{\varepsilon, \delta})$ strongly converges in $L^1_{loc}(\mathbb{R}^+ \times \mathbb{R}^d)$ toward a solution to the conservation law

$$\partial_t u + \operatorname{div} f(\mathbf{x}, u) = 0.$$

The main tools employed in the proof are the Leray–Schauder fixed point theorem for the first part and reduction to the kinetic formulation combined with recent results in the velocity averaging theory for the second. These results have the potential to generate a stable semigroup of solutions to the underlying scalar conservation laws different from the Kruzhkov entropy solutions concept.

Mathematics Subject Classification. 35K65, 42B37, 76S99.

Keywords. Pseudo-parabolic equations, Vanishing dynamic capillarity, Discontinuous flux, Conservation laws.

1. Introduction and notation

Flow in a two-phase porous medium is governed by the Darcy law [7]

$$q = -K(S) (\nabla p + \rho g e_d), \quad (1.1)$$

where $e_d = (0, \dots, 0, 1)$, is the direction of gravity. The quantity S is the saturation, p is the pressure, and (the vector) q is the flow velocity of the wetting phase (usually water, while the non-wetting one is oil or a gas).

The Darcy law represents conservation of momentum and, in order to close the system, we also need the conservation of mass

$$\partial_t S + \operatorname{div} q = 0. \quad (1.2)$$

In the two-dimensional situation, we have three equations, given by (1.1) and (1.2), while we have four unknowns (two velocity components, saturation, and pressure). Therefore, usually one assumes a constitutive relation between the pressure p , the capillary pressure P_c (equal to differences of pressures between wetting and non-wetting phases), and the saturation S . If it is assumed that the capillary pressure is

(almost) constant, one can derive the Buckley–Leverett equation (a scalar conservation law derived in [8]):

$$\partial_t S + \partial_x f(S) = 0, \quad (1.3)$$

for $f(S) = \frac{S^2}{S^2 + A(1-S)^2}$ and an appropriate constant A .

If we assume that P_c is “static” (independent of the t -derivative of S) in the sense that $p_c = p_c(S)$, then we arrive at a parabolic perturbation of the Buckley–Leverett equation [7]. However, both of the models (the standard Buckley–Leverett or the one perturbed by a parabolic term) appear to give results inconsistent with certain fairly simple experiments [14]. Namely, if we take a thin tube filled with dry sand and dip it in water at a constant rate at one side of the tube and then measure the concentration of the water as a function of time, for certain dipping rates the concentration will not be monotonic. This phenomenon, called overshoot, had been noticed long ago and is a subeffect of the fingering effect [25]. Namely, the tips of the fingers that appear when water is penetrating into dry land have larger water concentrations than the body of the finger and, interestingly, the tips of the fingers almost do not change their shape or size. On the other hand, an entropy solution to the Buckley–Leverett equation (in the sense of Kruzhkov [33] or more precisely Oleinik [40] since we are dealing with a Riemann problem) or the parabolic perturbation of the Buckley–Leverett equation with the Riemann initial data

$$S(0, x) = \begin{cases} S_L, & x < 0 \\ 0, & x > 0 \end{cases} \quad (1.4)$$

is monotonic (see, e.g., the introduction of [53]). Therefore, many attempts to explain the gap between the standard theory (provided by the Buckley–Leverett or Richards [45] equations) have been put forth recently. A purely mathematical approach can be found in [27], where δ -type solutions to (1.3) are constructed and δ -distributions appearing as a part of the solution are explained as an inadequacy of the model. In [13], a model is suggested that involves the notion of non-local energy, which eventually leads to a fourth-order equation whose solution (with the Riemann initial data given above) has a shape corresponding to the experimental results from [14].

In [53] one can find an interesting approach explaining the discrepancy between theory and experiment described above, based on the dynamics capillarity concept introduced in [22, 23], which has drawn a lot of interest (especially after the publication of [53]; see also [1]). Namely, in [22, 23] it was supposed that the capillary pressure depends not only on the saturation S , but also on the time derivative of the saturation (of the wetting phase):

$$P_c = p_c(S) - \phi\tau \frac{\partial S}{\partial t}.$$

Taking this into account and proceeding along the lines of deriving the Buckley–Leverett equation, one reaches a nonlinear pseudo-parabolic equation ([53, (1.16)]) which, after linearization of higher-order terms reduces to

$$\partial_t S + \partial_x f(S) = \varepsilon \partial_{xx} S + \tau \delta \partial_{xx} \partial_t S, \quad (1.5)$$

where ε and δ ($\delta = \varepsilon^2$ in [53]) are small parameters, while τ is a fixed constant. By analyzing possible traveling wave solutions $S(\frac{x-ct}{\varepsilon})$, $S(-\infty) = S_L$, $S(+\infty) = S_R$, where the constant c is given by the Rankine–Hugoniot condition $c = \frac{f(S_L) - f(S_R)}{S_L - S_R}$, the authors arrive at solutions constructed from elementary waves (i.e., solutions consisting of shock waves and rarefaction waves) to (1.3), (1.4) that are non-standard (i.e., non-admissible in the sense of Oleinik). Moreover, for some values S_L , such a solution exhibits an overshoot-type phenomenon (see [53, Figure 7]).

In this paper, we shall make a step forward in the sense that we shall consider a multi-dimensional generalization to (1.5) with a flux $f = f(\mathbf{x}, \lambda)$ explicitly depending on the space variable and we shall analyze the dynamics capillarity limit to (1.5) as $\varepsilon \rightarrow 0$ for arbitrary initial data.

The $\mathbf{x}$ -dependence of the flux means that we assume that the medium in which we consider the phenomenon is heterogeneous, i.e., that it has different properties at different points (e.g., since in some parts of the medium we have sand while in some others clay). Moreover, we shall assume that the flux is discontinuous with respect to the space variable $\mathbf{x}$, which means that the medium experiences abrupt (discontinuous) changes in its properties (one can imagine that we have sand which is highly permeable adjacent to the clay layer, which is weakly permeable—permeability is discontinuous in such a medium). Let us remark here that, stipulated by different applications, evolutionary equations with discontinuous flux have attracted considerable attention recently. For a (non-exhaustive) selection of recent results, cf. [2, 3, 12, 29, 43] and references therein.

Explicitly, we consider the conservation law

$$\partial_t u + \operatorname{div} \mathbf{f}(\mathbf{x}, u) = 0, \quad (1.6)$$

where

(C1) $\mathbf{f} = (f_i)_{i=1}^d = (\mathbf{x}, \lambda) \mapsto \mathbf{f}(\mathbf{x}, \lambda)$, $f_i \in L^1(\mathbb{R}^d \times \mathbb{R}) \cap L^p(\mathbb{R}^d \times \mathbb{R})$ and $\partial_\lambda f_i \in L^p(\mathbb{R}^d \times \mathbb{R})$ for some $1 < p < \infty$ and all $i = 1, \dots, d$;

(C2) $\operatorname{div}_{\mathbf{x}} \mathbf{f}(\mathbf{x}, \xi) \in \mathcal{M}(\mathbb{R}^d \times \mathbb{R})$, where $\mathcal{M}(\mathbb{R}^d \times \mathbb{R})$ is the space of Radon measures, and there exists a constant $\beta > 0$ and a finite Radon measure $\mu \in \mathcal{M}(\mathbb{R}^d)$ such that (in the sense of measures)

$$|\operatorname{div}_{\mathbf{x}} \mathbf{f}(\mathbf{x}, \xi)| \leq \frac{\mu(\mathbf{x})}{1 + |\lambda|^{1+\beta}};$$

(C3) $\|\sup_{\lambda \in \mathbb{R}} |\partial_\lambda f_i(\cdot, \lambda)|\|_{L^1(\mathbb{R}^d)} < \infty$ for all $i = 1, \dots, d$;

We then perturb (1.6) by (a linearized) vanishing dynamic capillarity limit:

$$\partial_t u_\varepsilon + \operatorname{div} \mathbf{f}_\varepsilon(\mathbf{x}, u_\varepsilon) = \varepsilon \Delta u_\varepsilon + \delta(\varepsilon) \Delta \partial_t u_\varepsilon. \quad (1.7)$$

Here, $\mathbf{f}_\varepsilon := K_\varepsilon \cdot \mathbf{f} \star \omega_{n(\varepsilon)}$, where componentwise convolution $\mathbf{f} \star \omega_{n(\varepsilon)}$ is a regularization of $\mathbf{f}$ and $K_\varepsilon = K_\varepsilon(\mathbf{x}, \lambda)$ is a bounded family of compactly supported functions equal to one on the ball $B(0, 1/\varepsilon) \subset \mathbb{R}^{d+1}$ and such that $\|\nabla K_\varepsilon\|_{L^\infty(\mathbb{R}^{d+1})} \leq 1$. We remark here that (1.7) reminds on the diffusion–dispersion regularization (see, e.g., [4, 31]), but it contains t -derivative in the third-order term which makes significant difference between the two situations. We will comment on this in more details later.

More precisely, we suppose that $\omega \in C_c^\infty(\mathbb{R}^{d+1})$ is of the form $\omega(\mathbf{x}, \lambda) = \omega^{(1)}(\mathbf{x}) \omega^{(2)}(\lambda)$ with $\omega^{(1)}, \omega^{(2)}$ test functions with unit integral on $\mathbb{R}^d$ and $\mathbb{R}$, respectively, and

$$\omega_{n(\varepsilon)}(\mathbf{x}, \lambda) \equiv \omega_{n(\varepsilon)}^{(1)}(\mathbf{x}) \omega_{n(\varepsilon)}^{(2)}(\lambda) = n(\varepsilon)^{-d} \omega^{(1)}(\mathbf{x}/n(\varepsilon)) n(\varepsilon)^{-1} \omega^{(2)}(\lambda/n(\varepsilon)) \quad (1.8)$$

with an appropriate $n(\varepsilon)$ such that $n(\varepsilon) \rightarrow 0$ as $\varepsilon \rightarrow 0$ (the form of $n(\varepsilon)$ will be precisely determined later). Note that

$$K_\varepsilon \mathbf{f}_\varepsilon \rightarrow \mathbf{f} \text{ strongly in } L^p(\mathbb{R}^d \times \mathbb{R}) \text{ as } \varepsilon \rightarrow 0.$$

We supplement (1.6) with the initial data

$$u|_{t=0} = u_0(\mathbf{x}) \in L^1(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d), \quad (1.9)$$

while we take

$$u_\varepsilon|_{t=0} = u_0^\varepsilon(\mathbf{x}) \rightarrow u_0(\mathbf{x}) \text{ strongly in } L^p(\mathbb{R}^d), \quad (1.10)$$

where, for every $i = 1, \dots, d$,

$$\|u_0^\varepsilon\|_{L^2(\mathbb{R}^d)} + n(\varepsilon) \|\nabla u_0^\varepsilon\|_{L^2(\mathbb{R}^d)} + n(\varepsilon)^2 \|\nabla u_{0x_i}^\varepsilon\|_{L^2(\mathbb{R}^d)} \leq C_0. \quad (1.11)$$

It is not difficult to see that $u_0^\varepsilon = u_0 \star \omega_{n(\varepsilon)}^{(1)}$, where $\omega_{n(\varepsilon)}^{(1)}$ is as above, satisfies (1.11). We shall, however, use (1.11) as an assumption.

With regard to pseudo-parabolic equations, we have already explained their importance in the porous media theory. Besides a possible justification of the experimental results from [14], there is a confirmed

application to the seepage of homogeneous fluids through a fissured rock [5]. Also, they describe the unidirectional propagation of nonlinear, dispersive long waves [6, 51] (where u is typically the amplitude or velocity), or population dynamics [41] (where u represents the population density).

In [48, 52], the authors investigated the initial-boundary value problem and the Cauchy problem for linear pseudo-parabolic equations and established the existence and uniqueness of solutions. As for the nonlinear variants, one can find numerous results even for singular pseudo-parabolic equations and degenerate pseudo-parabolic equations (see, e.g., [9, 28, 30, 34, 39, 44] and the references therein). Together with the existence and uniqueness results, among the given references, one can also find properties of solutions, such as asymptotic behavior and regularity. However, we are not aware of any corresponding results for the Cauchy problem for an equation of type (1.5).

This problem will be considered in Sect. 2. We shall use an approach that is characteristic for the theory of wave equations [47]: we shall apply the Fourier transform with respect to $\mathbf{x}$, solve the ordinary differential equation so obtained with respect to t and prove the following theorem (the constants ε and δ are omitted for simplicity):

Theorem 1.1. *There exists a unique solution to*

$$u_t + \operatorname{div}(\mathfrak{f}(\mathbf{x}, u)) = \Delta u + \partial_t \Delta u, \quad (t, x) \in [0, T) \times \mathbb{R}^d, \quad (1.12)$$

where $T > 0$, $\mathfrak{f} \in C_c^\infty(\mathbb{R}^d \times \mathbb{R})$ is a bounded function, supplemented with (merely) the initial condition

$$u|_{t=0} = u_0(\mathbf{x}) \in C_c^\infty(\mathbb{R}^d), \quad (1.13)$$

which belongs to $L^2((0, T) \times \mathbb{R}^d) \cap C^\infty((0, T) \times \mathbb{R}^d)$.

In the next step (Sect. 3), we are going to let the perturbation parameter $\varepsilon \rightarrow 0$ in (1.7). This type of problem for conservation laws (when there is not only vanishing viscosity, but also third- or higher-order perturbation) was first addressed in [46] and received considerable attention after that. A thorough analysis of this kind of limit *in the one-dimensional situation with a regular flux* can be found in [38], where the theory of non-classical shocks for conservation laws was essentially initiated. The standard approach here is to rewrite the equation under consideration in the kinetic formulation or using Young measures (which is essentially equivalent), and then applying velocity averaging results [4, 26], compensated compactness (in one-dimensional situations) [10, 11, 46], or Di Perna techniques involving Young measures [15], as done in [31] and many others. Again we note that this list of citations is far from complete. The problem in our case is the low regularity of the flux function $\mathfrak{f}$ (it can be discontinuous with respect to $\mathbf{x} \in \mathbb{R}^d$) as well as the multidimensional character of our problem, which prevents us from using the Young measures and compensated compactness approach here (see item (iii) below and Remark 3.9 for the case of a regular flux). We also remark that, to the best of our knowledge, the only diffusion–dispersion type result for equations with discontinuous coefficients is [24].

Moreover, unlike the situation that is typical in the case of the diffusion–dispersion limit ([4, 31, 46] etc.), where the existence of a solution to the perturbed equation is assumed together with all necessary properties of the solution, we have proved in the previous section the existence of the functional sequence whose convergence we analyze.

As for the velocity averaging theory, most of the results are given in the case of a homogeneous flux [16, 21, 42, 50] or a flux for which $p \geq 2$ [19, 35]. We have recently proved [36] the velocity averaging lemma in the case $p > 1$ and this will enable us to prove the strong convergence of the sequence (u_ε) of solutions to (1.7), (1.10) along a subsequence toward a weak solution to (1.6). The following statements are the main results of the paper:

(i) If $\delta = o(\varepsilon^2)$ and

$$\frac{\sqrt{\delta}}{\varepsilon n(\varepsilon)^{d+2}} \rightarrow 0 \quad (1.14)$$

as $\varepsilon \rightarrow 0$, then the family (u_ε) contains a strongly converging subsequence in $L^1_{loc}(\mathbb{R}^+ \times \mathbb{R}^d)$ under a suitable non-degeneracy condition (see Definition 3.2);

- (ii) If $\delta = \mathcal{O}(\varepsilon^2)$ and $\frac{\sqrt{\delta}}{\sqrt{\varepsilon n(\varepsilon)^{d/2+2}}} \rightarrow 0$ as $\varepsilon \rightarrow 0$, and the λ -derivative of the flux $\partial_\lambda \mathbf{f}$ is bounded in addition to condition (C3), then the family (u_ε) is strongly precompact in $L^1_{loc}(\mathbb{R}^+ \times \mathbb{R}^d)$ under the non-degeneracy condition;
- (iii) If $\delta = o(\varepsilon^2)$ and $\frac{\sqrt{\delta}}{\sqrt{\varepsilon n(\varepsilon)^{d/2+2}}} \rightarrow 0$ as $\varepsilon \rightarrow 0$, and if the flux $\mathbf{f} \in C^1(\mathbb{R}^d \times \mathbb{R})$, then the family (u_ε) converges strongly in $L^1_{loc}(\mathbb{R}^+ \times \mathbb{R}^d)$ toward the entropy solution to (1.6), (1.9) (no non-degeneracy condition is needed).

We conclude the introduction by noting that the subsequence of (u_ε) given in (i) does not converge toward an entropy admissible solution to the underlying conservation law since its flux is not smooth and for such fluxes the well-posedness theory is not developed yet in full generality (compare with item (iii) above).

Also, in the case (ii), even when the flux is smooth, we do not necessarily have convergence toward the Kruzhkov entropy solution (see [53] for stepwise initial data). On the other hand, the approximating procedure actually models a physical situation in which diffusion and capillary limits have equal effect on the process and, as any physical phenomenon in the macro-world is well-posed, it is therefore expected to generate a stable semigroup of solutions to the underlying conservation law. We will deal with this question in a future research.

The paper is organized as follows: In Sect. 2, we prove existence and uniqueness for the pseudo-parabolic problem (1.12), (1.13). In Sect. 3, we derive necessary estimates from (1.7), (1.10) and use it to prove the strong convergence toward a weak solution to the underlying conservation law.

2. Existence and uniqueness of the solution to the pseudo-parabolic equation (1.12) with (1.13)

Throughout this section, we suppose that $\mathbf{f}$ is **smooth and compactly supported**. Under this assumption, we want to show that (1.12) with the initial condition (1.13) has a unique solution. The strategy for solving this problem is to define a mapping $\mathcal{T} : L^2([0, T) \times \mathbb{R}^d) \rightarrow L^2([0, T) \times \mathbb{R}^d)$ such that for every $v \in L^2([0, T) \times \mathbb{R}^d)$ the function

$$u = \mathcal{T}(v) \quad (2.1)$$

represents a solution to

$$u_t + \operatorname{div}(\mathbf{f}(\mathbf{x}, v)) = \Delta u + \partial_t \Delta u, \quad (t, x) \in [0, T) \times \mathbb{R}^d, \quad (2.2)$$

with the initial conditions (1.13). Generally, for $u_0 \in L^2(\mathbb{R}^d)$, we are seeking weak solutions u of this initial value problem in the sense that, for any smooth test function φ with compact support in $[0, T)$, we require

$$\int_0^T \int_{\mathbb{R}^d} u \varphi_t + f(x, v) \operatorname{div} \varphi - u \Delta \varphi + u \partial_t \Delta \varphi \, dx dt + \int_{\mathbb{R}^d} u_0 (\varphi(0, x) + \Delta \varphi(0, x)) \, dx = 0.$$

Then, we shall prove that the mapping $\mathcal{T}$ possesses a fixed point, which will turn out to be the solution to (1.12), (1.13).

To this end, we need the following consequence of the Leray–Schauder fixed point theorem (cf. [20, Th. 11.3]):

Theorem 2.1. *Let $\mathcal{T}$ be a compact mapping of a Banach space $\mathcal{B}$ into itself and suppose that there exists a constant C such that*

$$\|u\|_{\mathcal{B}} \leq C \quad (2.3)$$

for all $u \in \mathcal{B}$ and $\sigma \in [0, 1]$ satisfying $u = \sigma \mathcal{T}u$. Then, $\mathcal{T}$ has a fixed point, that is, $\mathcal{T}u = u$ for some $u \in \mathcal{B}$.

Finally, let us fix our conventions for Fourier transform and inverse Fourier transform. For $u \in L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$, the Fourier transform of u is

$$\mathcal{F}(u)(\xi) = \hat{u}(\xi) = \int_{\mathbb{R}^d} e^{-i\mathbf{x} \cdot \xi} u(\mathbf{x}) \, d\mathbf{x}$$

and the inverse Fourier transform is given by

$$\mathcal{F}^{-1}(u)(\mathbf{x}) = \check{u}(x) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} e^{i\mathbf{x} \cdot \xi} u(\xi) \, d\xi,$$

where $\cdot$ denotes the scalar product and i is the imaginary unit.

We begin by proving existence and uniqueness of solutions to (2.2), (1.13).

Lemma 2.2. *Let $v \in L^2([0, T] \times \mathbb{R}^d)$. Then, for any $u_0 \in L^2(\mathbb{R}^d)$ there exists a unique solution $u \in L^2([0, T] \times \mathbb{R}^d)$ to (2.2) with $u|_{t=0} = u_0$.*

Proof. After applying the Fourier transform with respect to $\mathbf{x} \in \mathbb{R}^d$ to (2.2), we obtain

$$\partial_t \hat{u}(t, \xi) + i\xi \cdot \widehat{\mathbf{f}(\cdot, v(t, \cdot))}(\xi) = -|\xi|^2 \hat{u}(t, \xi) - |\xi|^2 \partial_t \hat{u}(t, \xi),$$

where $|\xi| = (\xi_1^2 + \dots + \xi_d^2)^{\frac{1}{2}}$. We rewrite the last equation in the form

$$\frac{d\hat{u}(t, \xi)}{dt} + \frac{|\xi|^2}{1 + |\xi|^2} \hat{u}(t, \xi) = -\frac{i\xi \cdot \widehat{\mathbf{f}(\cdot, v(t, \cdot))}(\xi)}{1 + |\xi|^2}$$

It follows from this that any weak solution to (2.2) has t -derivative in L^2 as well, hence in particular is continuous with respect to t and therefore possesses a classical trace on $t = 0$. Solving the ODE with the initial data $\hat{u}(0, \xi) = \hat{u}_0(\xi)$, we arrive at

$$\hat{u}(t, \xi) = e^{-\frac{|\xi|^2 t}{1 + |\xi|^2}} \left(\hat{u}_0(\xi) - \int_0^t \frac{i\xi \cdot \hat{\mathbf{f}}}{1 + |\xi|^2} e^{\frac{|\xi|^2 t'}{1 + |\xi|^2}} dt' \right), \quad (2.4)$$

where $\hat{\mathbf{f}} = \widehat{\mathbf{f}(\cdot, v(t, \cdot))}(\xi)$. By finding the inverse Fourier transform here with respect to $\xi \in \mathbb{R}^d$, we indeed obtain a weak solution, so combined with the above considerations both existence and uniqueness follow. $\square$

Theorem 2.3. *For any function $v \in L^2([0, T] \times \mathbb{R}^d)$, the solution to (2.2), (1.13) belongs to $H^1((0, T) \times \mathbb{R}^d)$.*

Proof. Let u be the solution to (2.2), (1.13) defined in Lemma 2.2. We need to estimate $\partial_{x_j} u$, $j = 1, \dots, d$ and $\partial_t u$, where ∂_{x_j} , $j = 1, \dots, d$ and ∂_t are the weak derivatives of u . We have according to the Plancherel theorem (below, we take the Fourier transform with respect to $\mathbf{x} \in \mathbb{R}^d$) and the Cauchy-Schwartz

inequality

$$\begin{aligned}
\|\partial_{x_j} u\|_{L^2([0,T] \times \mathbb{R}^d)} &= \|i\xi_j \hat{u}\|_{L^2([0,T] \times \mathbb{R}^d)} \\
&\leq \|\xi_j \hat{u}_0\|_{L^2([0,T] \times \mathbb{R}^d)} + \left\| \int_0^t \frac{i\xi_j \xi \cdot \hat{f}}{1 + |\xi|^2} e^{\frac{|\xi|^2 t'}{1 + |\xi|^2}} dt' \right\|_{L^2([0,T] \times \mathbb{R}^d)} \\
&\leq \|\xi_j \hat{u}_0\|_{L^2([0,T] \times \mathbb{R}^d)} + \left(\int_0^T \int_{\mathbb{R}^d} \left| \int_0^t \frac{i\xi_j \xi \cdot \hat{f}}{1 + |\xi|^2} e^{\frac{|\xi|^2 t'}{1 + |\xi|^2}} dt' \right|^2 d\xi dt \right)^{1/2} \\
&\leq \|\xi_j \hat{u}_0\|_{L^2([0,T] \times \mathbb{R}^d)} + \left(\int_0^T \int_{\mathbb{R}^d} t \int_0^t \frac{\xi_j^2 |\xi \cdot \hat{f}|^2}{(1 + |\xi|^2)^2} e^{\frac{2|\xi|^2 t'}{1 + |\xi|^2}} dt' d\xi dt \right)^{1/2} \quad (2.5) \\
&\leq \|\xi_j \hat{u}_0\|_{L^2([0,T] \times \mathbb{R}^d)} + Te^T \left(\int_0^T \int_{\mathbb{R}^d} \frac{\xi_j^2 |\xi|^2 |\hat{f}|^2}{(1 + |\xi|^2)^2} d\xi dt \right)^{1/2} \\
&\leq \|\xi_j \hat{u}_0\|_{L^2([0,T] \times \mathbb{R}^d)} + Te^T \left(\int_0^T \int_{\mathbb{R}^d} |f(\mathbf{x}, v(t, \mathbf{x}))|^2 d\mathbf{x} dt \right)^{1/2},
\end{aligned}$$

where in the last step we used the Plancherel theorem again. From here, since f is compactly supported with respect to $\mathbf{x} \in \mathbb{R}^d$ and bounded, we conclude that $\partial_{x_j} u \in L^2([0, T] \times \mathbb{R}^d)$.

As for the t -derivative, we have

$$\partial_t \hat{u} = -\frac{|\xi|^2}{1 + |\xi|^2} e^{\frac{-|\xi|^2 t}{1 + |\xi|^2}} \left(\hat{u}_0(\xi) - \int_0^t \frac{i\xi \cdot \hat{f}}{1 + |\xi|^2} dt' \right) + \frac{i\xi \cdot \hat{f}}{1 + |\xi|^2}, \quad (2.6)$$

and therefore, repeating the procedure giving us the estimates on $\partial_{x_j} u$, we get

$$\|\partial_t u\|_{L^2([0,T] \times \mathbb{R}^d)} = \|\partial_t \hat{u}\|_{L^2([0,T] \times \mathbb{R}^d)} \leq C$$

for a constant $C \in \mathbb{R}$. □

Based on this, we can proceed to applying Theorem 2.1 to prove existence of a solution to (1.12), (1.13).

Theorem 2.4. *There exists a unique solution to (1.12), (1.13). This solution belongs to $L^2([0, T] \times \mathbb{R}^d) \cap C^\infty([0, T] \times \mathbb{R}^d)$.*

Proof. Let us first take a sequence of balls $B_n = B(0, n)$ centered at 0 and of radius n . Set $u_0^n(\mathbf{x}) = u_0(\mathbf{x})\chi_{B_n}(\mathbf{x})$, where $\chi_{B_n} : \mathbb{R}^d \rightarrow [0, 1]$ is a smooth regularization of the characteristic function of the set B_n such that $\chi_{B_n}(\mathbf{x}) = 1$ for $\mathbf{x} \in B(0, n-1)$ and $\chi_{B_n}(\mathbf{x}) = 0$ for $\mathbf{x} \notin B(0, n)$. Then, we define the mapping $\mathcal{T}_n : L^2([0, T] \times B_n) \rightarrow L^2([0, T] \times B_n)$ by

$$\mathcal{T}_n(v) := \chi_{B_n} \mathcal{T}(v\chi_{B_n}),$$

where $\mathcal{T}$ is the operator from (2.1) with the initial data $u_0\chi_{B_n}$ and we extend $v\chi_{B_n}$ by zero so it is defined on $[0, T] \times \mathbb{R}^d$. This map is continuous since $\mathcal{T}$ is continuous. In fact, by the proof of Theorem 2.3, $\mathcal{T}$ is in fact continuous as a map from $L^2([0, T] \times \mathbb{R}^d)$ to $H^1((0, T) \times \mathbb{R}^d)$. So, we can use Theorem 2.3 and the Rellich theorem (keeping in mind that B_n is a bounded set) to conclude that $\mathcal{T}_n$ is a compact mapping.

Let us check condition (2.3) from Theorem 2.1. If a function $u \in L^2([0, T] \times B_n)$ satisfies $u = \sigma \mathcal{T}_n u$, then the function $\tilde{u} := \sigma \mathcal{T}(u \chi_{B_n})$ satisfies $\chi_{B_n} \tilde{u}(t, \mathbf{x}) = u(t, \mathbf{x})$ for $\mathbf{x} \in B_n$ and is a solution of the Cauchy problem

$$\begin{aligned} \tilde{u}_t + \sigma \operatorname{div}(\mathbf{f}(\mathbf{x}, u \chi_{B_n})) &= \Delta \tilde{u} + \partial_t \Delta \tilde{u}, \quad (t, x) \in [0, T] \times \mathbb{R}^d \\ \tilde{u}(0, \mathbf{x}) &= \sigma u_0(\mathbf{x}) \chi_{B_n}(\mathbf{x}) \end{aligned}$$

If we apply Fourier transform with respect to $\mathbf{x} \in \mathbb{R}^d$ here and solve the ODE so obtained, we arrive at an expression analogous to (2.4):

$$\hat{\tilde{u}}(t, \boldsymbol{\xi}) = \sigma e^{-\frac{t|\boldsymbol{\xi}|^2}{1+|\boldsymbol{\xi}|^2}} \left(\widehat{u_0 \chi_{B_n}}(\boldsymbol{\xi}) - \int_0^t \frac{i \boldsymbol{\xi} \cdot \hat{\mathbf{f}}}{1+|\boldsymbol{\xi}|^2} e^{\frac{t'|\boldsymbol{\xi}|^2}{1+|\boldsymbol{\xi}|^2}} dt' \right),$$

where $\hat{\mathbf{f}} = \mathcal{F}(\mathbf{f}(\mathbf{x}, u(t, \mathbf{x}) \chi_{B_n}(t, \mathbf{x})))$ is the Fourier transform with respect to $\mathbf{x}$. Now $\|u\|_{L^2([0, T] \times B_n)} \leq \|\tilde{u}\|_{L^2([0, T] \times \mathbb{R}^d)}$ and repeating the procedure from (2.5) and (2.6), we conclude that (2.3) holds for the mapping $\mathcal{T}_n$ (keeping in mind that n is fixed).

Thus, for every $n \in \mathbb{N}$ there exists a function $u_n \in L^2([0, T] \times B_n)$ solving

$$\begin{aligned} \partial_t u_n + \operatorname{div}(\mathbf{f}(\mathbf{x}, u_n)) &= \Delta u_n + \partial_t \Delta u_n, \quad (t, x) \in [0, T] \times B_{n-1} \\ u_n(0, \mathbf{x}) &= u_0(\mathbf{x}), \quad \mathbf{x} \in B_{n-1} \end{aligned}$$

Extending u_n by zero outside of B_n , we thereby obtain a sequence (u_n) that is bounded in $L^2([0, T] \times \mathbb{R}^d)$, and locally bounded in $H^1((0, T) \times \mathbb{R}^d)$. Thus, (u_n) admits a subsequence that locally converges toward $u \in L^2([0, T] \times \mathbb{R}^d)$. It is clear that the function u is the weak solution to (1.12), (1.13) and thus it must belong at least to $H^1((0, T) \times \mathbb{R}^d)$.

Moreover, the proof of Theorem 2.3 with $v = u$ is a classical example of a bootstrapping procedure, which enables us to conclude that u actually belongs to each $H^k((0, T) \times \mathbb{R}^d)$ ($k \in \mathbb{N}$), hence is smooth (due to smoothness of u_0 and $\mathbf{f}$).

Now we turn to the proof of uniqueness. Assume that u_1 and u_2 are solutions to (1.12), supplemented with the initial conditions $u_1|_{t=0} = u_{10}$ and $u_2|_{t=0} = u_{20}$. After applying the procedure from Theorem 2.3 we conclude that

$$\hat{u}_1(t, \boldsymbol{\xi}) - \hat{u}_2(t, \boldsymbol{\xi}) = e^{-\frac{t|\boldsymbol{\xi}|^2}{1+|\boldsymbol{\xi}|^2}} \left(\hat{u}_{10}(\boldsymbol{\xi}) - \hat{u}_{20}(\boldsymbol{\xi}) - \int_0^t \frac{i \boldsymbol{\xi} \cdot (\hat{\mathbf{f}}_1 - \hat{\mathbf{f}}_2)}{1+|\boldsymbol{\xi}|^2} e^{\frac{t'|\boldsymbol{\xi}|^2}{1+|\boldsymbol{\xi}|^2}} dt' \right), \quad (2.7)$$

where

$$\hat{\mathbf{f}}_1 - \hat{\mathbf{f}}_2 = \mathcal{F}(\mathbf{f}(\mathbf{x}, u_1(t, \mathbf{x})) - \mathbf{f}(\mathbf{x}, u_2(t, \mathbf{x}))).$$

Let us first estimate

$$\begin{aligned} \left\| \frac{\boldsymbol{\xi} \cdot (\hat{\mathbf{f}}_1 - \hat{\mathbf{f}}_2)}{1+|\boldsymbol{\xi}|^2} \right\|_{L^2(\mathbb{R}^d)} &\leq \|\hat{\mathbf{f}}_1 - \hat{\mathbf{f}}_2\|_{L^2(\mathbb{R}^d)} = \|\mathbf{f}_1 - \mathbf{f}_2\|_{L^2(\mathbb{R}^d)} \\ &\leq \|\partial_\lambda \mathbf{f}\|_\infty \|\hat{u}_1 - \hat{u}_2\|_{L^2(\mathbb{R}^d)}. \end{aligned} \quad (2.8)$$

Taking the $L^2(\mathbb{R}^d)$ -norm of (2.7), we get, due to (2.8):

$$\|\hat{u}_1(t, \cdot) - \hat{u}_2(t, \cdot)\|_{L^2(\mathbb{R}^d)} \leq \|\hat{u}_{10} - \hat{u}_{20}\|_{L^2(\mathbb{R}^d)} + \|\partial_\lambda \mathbf{f}\|_\infty \int_0^t \|\hat{u}_1(t', \cdot) - \hat{u}_2(t', \cdot)\|_{L^2(\mathbb{R}^d)} e^{t'} dt'. \quad (2.9)$$

Applying Gronwall's inequality, we arrive at

$$\|\hat{u}_1(t, \cdot) - \hat{u}_2(t, \cdot)\|_{L^2(\mathbb{R}^d)} \leq \|\hat{u}_{10} - \hat{u}_{20}\|_{L^2(\mathbb{R}^d)} e^{(e^t - 1)\|\partial_\lambda \mathbf{f}\|_\infty}.$$

From here, we see that $u_{10} = u_{20}$ entails $u_1 = u_2$. This concludes the proof. $\square$

3. Vanishing capillarity limit

In this section, we inspect the vanishing capillarity limit of (1.7). By Theorem 1.1, under the assumptions (C1)–(C3) from Sect. 1, the Eq. (1.7) with initial data (1.10) (satisfying (1.11)) possesses a unique solution u_ε in $L^2((0, T) \times \mathbb{R}^d) \cap C^\infty((0, T) \times \mathbb{R}^d)$. As announced in Sect. 1, our main result then is as follows:

Theorem 3.1. *Under the above assumptions, if $(F_1, \dots, F_d) = \partial_\lambda \mathbf{f}$ lies in $L^{\bar{p}'}(\mathbb{R}^{d+1})$ for some $\bar{p}' \in (1, 2 + \frac{4}{d})$ then:*

- (i) *If $\delta = o(\varepsilon^2)$ and $n(\varepsilon)$ satisfies (1.14), then the family (u_ε) contains a strongly convergent subsequence in $L^1_{loc}(\mathbb{R}^+ \times \mathbb{R}^d)$ under the non-degeneracy condition given in Definition 3.2 below.*
- (ii) *If $\delta = \mathcal{O}(\varepsilon^2)$ and $\frac{\sqrt{\delta}}{\sqrt{\varepsilon n(\varepsilon)^{d/2+2}}} \rightarrow 0$ as $\varepsilon \rightarrow 0$, and the λ -derivative of the flux $\partial_\lambda \mathbf{f}$ is bounded in addition to condition (C3), then the family (u_ε) contains a strongly convergent subsequence in $L^1_{loc}(\mathbb{R}^+ \times \mathbb{R}^d)$ under the non-degeneracy condition;*
- (iii) *If $\delta = o(\varepsilon^2)$ and $\frac{\sqrt{\delta}}{\sqrt{\varepsilon n(\varepsilon)^{d/2+2}}} \rightarrow 0$ as $\varepsilon \rightarrow 0$, and if the flux $\mathbf{f} \in C^1(\mathbb{R}^d \times \mathbb{R})$, then the family (u_ε) converges strongly in $L^1_{loc}(\mathbb{R}^+ \times \mathbb{R}^d)$ toward the entropy solution to (1.6), (1.9).*

Below we give a complete proof of (i). Although non-trivial, the other two claims are more standard and we shall only comment on the necessary modifications of the argument for establishing (i).

To proceed, let us first introduce the non-degeneracy condition. Such a condition is standard in many results using velocity averaging lemmas (cf., e.g., [36, 37, 42, 43, 50]).

Definition 3.2. The flux appearing in (1.6) is called non-degenerate if for the function $\mathbf{F} = (F_1, \dots, F_d) = \partial_\lambda \mathbf{f}$, where $\mathbf{f} = \mathbf{f}(\mathbf{x}, \lambda)$ we have: for almost every $(t, \mathbf{x}) \in \mathbb{R}_+^d$ and every $\boldsymbol{\xi} \in S^d$, where S^d is the unit sphere in $\mathbb{R}^{d+1}$ the mapping

$$\lambda \mapsto \left(\xi_0 + \sum_{k=1}^d F_k(\mathbf{x}, \lambda) \xi_k \right), \quad (3.1)$$

is not zero on any set of positive measure.

Under these assumptions, we proved the following assertion in [36, Theorem 3.4].

Theorem 3.3. *Assume that $h_n \rightharpoonup h$ weakly in $L^s_{loc}(\mathbb{R}^{d+1})$ for some $s \geq 2$, where h_n are weak solutions to*

$$\sum_{k=1}^d \partial_{x_k} (F_k(\mathbf{x}, \lambda) h_n(\mathbf{x}, \lambda)) = \partial_\lambda^\kappa G_n(\mathbf{x}, \lambda). \quad (3.2)$$

Here, $\kappa \in \mathbb{N}_0$ and $\lambda \in \mathbb{R}$. Let $\bar{p} \in (1, s)$ and let $\frac{1}{\bar{p}} + \frac{1}{\bar{p}'} = 1$. We assume

- (a) $F_k \in L^{\bar{p}'}(\mathbb{R}^{d+1})$ for $k = 1, \dots, d$,
- (b) *The sequence (G_n) is strongly precompact in the space $W_{loc}^{-1,r}(\mathbb{R}^{d+1})$, where $r > 1$ satisfies the relation $1 + \frac{1}{s} = \frac{1}{\bar{p}} + \frac{1}{r}$.*

Finally, assume that the non-degeneracy condition (3.1) is satisfied.

Then, for any $\rho \in L^2_c(\mathbb{R})$, there exists a subsequence (h_{n_k}) of (h_n) such that

$$\int_{\mathbb{R}} \rho(\lambda) h_{n_k}(\mathbf{x}, \lambda) d\lambda \rightarrow \int_{\mathbb{R}} \rho(\lambda) h(\mathbf{x}, \lambda) d\lambda \quad \text{strongly in } L^1_{loc}(\mathbb{R}^d) \quad (3.3)$$

as $k \rightarrow \infty$.

Now, we proceed to the estimates sufficient to apply the mentioned functional analytic tools.

Lemma 3.4. *The solution u_ε of (1.7) satisfies the following a priori estimates:*

$$\|u_\varepsilon(t, \cdot)\|_{L^2(\mathbb{R}^d)} \leq \|u_0^\varepsilon\|_{L^2(\mathbb{R}^d)} + \sqrt{\delta}\|\nabla u_0^\varepsilon\|_{L^2(\mathbb{R}^d)} + \sqrt{2t}\|\operatorname{div} \mathbf{f}_\varepsilon(\mathbf{x}, \lambda)\|_{L^1(\mathbb{R}^d \times \mathbb{R})}^{\frac{1}{2}}, \quad (3.4)$$

$$\sqrt{2\varepsilon}\|\nabla u_\varepsilon\|_{L^2([0,T], L^2(\mathbb{R}^d))} \leq \|u_0^\varepsilon\|_{L^2(\mathbb{R}^d)} + \sqrt{\delta}\|\nabla u_0^\varepsilon\|_{L^2(\mathbb{R}^d)} + \sqrt{2T}\|\operatorname{div} \mathbf{f}_\varepsilon(\mathbf{x}, \lambda)\|_{L^1(\mathbb{R}^d \times \mathbb{R})}^{\frac{1}{2}}, \quad (3.5)$$

$$\sqrt{\delta}\|\nabla u_\varepsilon(t, \cdot)\|_{L^2(\mathbb{R}^d)} \leq \|u_0^\varepsilon\|_{L^2(\mathbb{R}^d)} + \sqrt{\delta}\|\nabla u_0^\varepsilon\|_{L^2(\mathbb{R}^d)} + \sqrt{2t}\|\operatorname{div} \mathbf{f}_\varepsilon(\mathbf{x}, \lambda)\|_{L^1(\mathbb{R}^d \times \mathbb{R})}^{\frac{1}{2}} \quad (3.6)$$

$$\begin{aligned} \varepsilon\delta\|\nabla \partial_{x_i} u_\varepsilon\|_{L^2([0,T], L^2(\mathbb{R}^d))}^2 &\leq 2\delta\|\partial_{x_i} u_0^\varepsilon\|_{L^2(\mathbb{R}^d)}^2 + 2\delta^2\|\nabla \partial_{x_i} u_0^\varepsilon\|_{L^2(\mathbb{R}^d)}^2 \\ &\quad + \frac{d\delta}{\varepsilon^2}\|\partial_\lambda \mathbf{f}_\varepsilon\|_\infty^2 (\|u_0^\varepsilon\|_{L^2(\mathbb{R}^d)}^2 + \delta\|\nabla u_0^\varepsilon\|_{L^2(\mathbb{R}^d)}^2) \\ &\quad + 2T\|\operatorname{div} \mathbf{f}_\varepsilon(\mathbf{x}, \lambda)\|_{L^1(\mathbb{R}^d \times \mathbb{R})} + \frac{4T\delta}{\varepsilon}\|\operatorname{div} \mathbf{f}_\varepsilon(\mathbf{x}, \lambda)\|_{L^2(\mathbb{R}^d, L^\infty(\mathbb{R}))}^2 \end{aligned} \quad (3.7)$$

$$\begin{aligned} \delta\|\nabla \partial_t u_\varepsilon\|_{L^2([0,T], L^2(\mathbb{R}^d))} &\leq \frac{\sqrt{\delta}\sqrt{d}}{2\sqrt{\varepsilon}}\|\partial_\lambda \mathbf{f}_\varepsilon\|_\infty\|u_0^\varepsilon\|_{L^2(\mathbb{R}^d)} \\ &\quad + \sqrt{\delta}\left(\frac{\sqrt{\delta}\sqrt{d}}{2\sqrt{\varepsilon}}\|\partial_\lambda \mathbf{f}_\varepsilon\|_\infty + \frac{\sqrt{\varepsilon}}{\sqrt{2}}\right)\|\nabla u_0^\varepsilon\|_{L^2(\mathbb{R}^d)} \\ &\quad + \frac{\sqrt{\delta}\sqrt{dT}}{\sqrt{2\varepsilon}}\|\partial_\lambda \mathbf{f}_\varepsilon\|_\infty\|\operatorname{div} \mathbf{f}_\varepsilon\|_{L^1(\mathbb{R}^d \times \mathbb{R})}^{\frac{1}{2}} + \sqrt{\delta}\sqrt{T}\|\operatorname{div} \mathbf{f}_\varepsilon\|_{L^2(\mathbb{R}^d, L^\infty(\mathbb{R}))}. \end{aligned} \quad (3.8)$$

Proof. Let us consider a family (u_ε) of smooth solutions to (1.7). Multiplying the equation by $\eta'(u_\varepsilon)$ (with $\eta(u_\varepsilon) = u_\varepsilon^2$), we have

$$\begin{aligned} \partial_t(u_\varepsilon^2) + 2\sum_{j=1}^d q_j^\varepsilon(\mathbf{x}, u_\varepsilon)_{x_j} + 2\int_0^{u_\varepsilon} \operatorname{div} \mathbf{f}_\varepsilon(\mathbf{x}, \lambda) d\lambda \\ = 2\varepsilon\sum_{j=1}^n (u_\varepsilon \partial_{x_j} u_\varepsilon)_{x_j} - 2\varepsilon|\nabla u_\varepsilon|^2 + \delta(\Delta \eta(u_\varepsilon))_t - 2\delta\sum_{j=1}^n \eta(\partial_{x_j} u_\varepsilon)_t - 2\delta\Delta u_\varepsilon \partial_t u_\varepsilon, \end{aligned} \quad (3.9)$$

where $\mathbf{q}_\varepsilon$ is defined by (recall that $\mathbf{f}_\varepsilon = (f_1^\varepsilon, \dots, f_d^\varepsilon)$)

$$\partial_\lambda \mathbf{q}_\varepsilon(\mathbf{x}, \lambda) = (\partial_\lambda F_1^\varepsilon(\mathbf{x}, \lambda), \dots, \partial_\lambda F_d^\varepsilon(\mathbf{x}, \lambda)) = (\lambda \partial_\lambda f_1^\varepsilon(\mathbf{x}, \lambda), \dots, \lambda \partial_\lambda f_d^\varepsilon(\mathbf{x}, \lambda)),$$

and normalized by the condition $q_j^\varepsilon(\mathbf{x}, 0) = 0$, $j = 1, \dots, d$. Integrating over $\mathbf{x} \in \mathbb{R}^d$ and $0 \leq t \leq T$ and using integration by parts, we get

$$\begin{aligned} \int_{\mathbb{R}^d} (|u_\varepsilon(t)|^2 - |u_0^\varepsilon|^2) dx + 2\int_0^t \int_{\mathbb{R}^d} \int_0^{u_\varepsilon(\mathbf{x}, t')} \operatorname{div} \mathbf{f}_\varepsilon(\mathbf{x}, \lambda) d\lambda d\mathbf{x} dt' \\ = -2\varepsilon\int_0^t \int_{\mathbb{R}^d} |\nabla u_\varepsilon|^2 d\mathbf{x} dt' + \delta\int_{\mathbb{R}^d} \Delta(\eta(u_\varepsilon(t, \mathbf{x})) - \eta(u_0^\varepsilon(\mathbf{x}))) d\mathbf{x} \\ - 2\delta\int_{\mathbb{R}^d} \sum_{j=1}^d (|\partial_{x_j} u_\varepsilon(t, \mathbf{x})|^2 - |\partial_{x_j} u_0^\varepsilon(\mathbf{x})|^2) d\mathbf{x} - 2\delta\int_0^t \int_{\mathbb{R}^d} \Delta u_\varepsilon \partial_t u_\varepsilon d\mathbf{x} dt' \\ = -2\varepsilon\int_0^t \int_{\mathbb{R}^d} |\nabla u_\varepsilon|^2 d\mathbf{x} dt' - 2\delta\int_{\mathbb{R}^d} \sum_{j=1}^d (|\partial_{x_j} u_\varepsilon(t, \mathbf{x})|^2 - |\partial_{x_j} u_0^\varepsilon(\mathbf{x})|^2) d\mathbf{x} \end{aligned}$$

$$\begin{aligned}
& + \delta \sum_{j=1}^d \int_{\mathbb{R}^d} \int_0^t \eta(\partial_{x_j} u_\varepsilon)_t dt dx = -2\varepsilon \int_0^t \int_{\mathbb{R}^d} |\nabla u_\varepsilon|^2 dx dt' \\
& - \delta \int_{\mathbb{R}^d} \sum_{j=1}^d |\partial_{x_j} u_\varepsilon(t, \mathbf{x})|^2 - |\partial_{x_j} u_0^\varepsilon(\mathbf{x})|^2(\mathbf{x}) d\mathbf{x}
\end{aligned}$$

i.e.,

$$\begin{aligned}
& \int_{\mathbb{R}^d} |u_\varepsilon(t)|^2 dx + 2\varepsilon \int_0^t \int_{\mathbb{R}^d} |\nabla u_\varepsilon|^2 dx dt' + \delta \sum_{j=1}^d \int_{\mathbb{R}^d} |\partial_{x_j} u_\varepsilon|^2 d\mathbf{x} \\
& = \int_{\mathbb{R}^d} |u_0^\varepsilon|^2 dx + \delta \sum_{j=1}^d \int_{\mathbb{R}^d} |\partial_{x_j} u_0^\varepsilon(\mathbf{x})|^2(x) dx \\
& \quad - 2 \int_0^t \int_{\mathbb{R}^d} \int_0^{u_\varepsilon(\mathbf{x}, t')} \operatorname{div} f_\varepsilon(\mathbf{x}, \lambda) d\lambda dx dt'. \tag{3.10}
\end{aligned}$$

The relations (3.4), (3.5) and (3.6) are easy consequences of (3.10).

To show (3.7), let us now differentiate equation (1.7) with respect to x_i and then multiply it by $2\partial_{x_i} u_\varepsilon$. Since the right-hand side of (1.7) is linear in u_ε , the calculation for that side will remain the same and after integrating the equation over $[0, t] \times \mathbb{R}^d$ and applying integration by parts, we get

$$\begin{aligned}
& \int_{\mathbb{R}^d} (|\partial_{x_i} u_\varepsilon(t, \mathbf{x})|^2 - |\partial_{x_i} u_0^\varepsilon(\mathbf{x})|^2) d\mathbf{x} - 2 \sum_{k=1}^d \int_0^t \int_{\mathbb{R}^d} \partial_\lambda f_k^\varepsilon(\mathbf{x}, u_\varepsilon) \partial_{x_k} u_\varepsilon \partial_{x_i}^2 u_\varepsilon dx dt \\
& - 2 \int_0^t \int_{\mathbb{R}^d} \partial_{x_i}^2 u_\varepsilon \operatorname{div} f^\varepsilon(\mathbf{x}, \lambda) \Big|_{\lambda=u_\varepsilon} dx dt = -2\varepsilon \int_0^t \int_{\mathbb{R}^d} |\nabla \partial_{x_i} u_\varepsilon|^2 dx ds \\
& - \delta \int_{\mathbb{R}^d} \sum_{j=1}^d (|\partial_{x_i x_j} u_\varepsilon(t, \mathbf{x})|^2 - |\partial_{x_i x_j} u_0^\varepsilon|^2) d\mathbf{x},
\end{aligned}$$

i.e.,

$$\begin{aligned}
& \int_{\mathbb{R}^d} |\partial_{x_i} u_\varepsilon(t, \mathbf{x})|^2 d\mathbf{x} + 2\varepsilon \int_0^t \int_{\mathbb{R}^d} |\nabla \partial_{x_i} u_\varepsilon|^2 dx dt' + \delta \int_{\mathbb{R}^d} |\nabla \partial_{x_i} u_\varepsilon(t, \mathbf{x})|^2 d\mathbf{x} \\
& = \int_{\mathbb{R}^d} |\partial_{x_i} u_0^\varepsilon(\mathbf{x})|^2 d\mathbf{x} + \delta \int_{\mathbb{R}^d} |\nabla \partial_{x_i} u_0^\varepsilon(\mathbf{x})|^2 d\mathbf{x} + 2 \int_0^t \int_{\mathbb{R}^d} \partial_{x_i}^2 u_\varepsilon \operatorname{div} f^\varepsilon(\mathbf{x}, \lambda) \Big|_{\lambda=u_\varepsilon} dx dt' \\
& \quad + 2 \sum_{k=1}^d \int_0^t \int_{\mathbb{R}^d} \partial_\lambda f_k^\varepsilon(\mathbf{x}, u_\varepsilon) \partial_{x_k} u_\varepsilon \partial_{x_i}^2 u_\varepsilon dx dt' \\
& \leq \|\partial_{x_i} u_0^\varepsilon\|_{L^2(\mathbb{R}^d)}^2 + \delta \|\nabla \partial_{x_i} u_0^\varepsilon\|_{L^2(\mathbb{R}^d)}^2 + \frac{\varepsilon}{2} \int_0^t \int_{\mathbb{R}^d} |\partial_{x_i}^2 u_\varepsilon|^2 dx dt'
\end{aligned}$$

$$\begin{aligned}
& + \frac{2}{\varepsilon} \|\operatorname{div} f^\varepsilon(\mathbf{x}, \lambda)|_{\lambda=u_\varepsilon}\|_{L^2([0,T] \times \mathbb{R}^d)}^2 + \frac{d \|\partial_\lambda f_\varepsilon\|_\infty^2}{\varepsilon} \int_0^t \int_{\mathbb{R}^d} |\nabla u_\varepsilon|^2 d\mathbf{x} dt' \\
& + \varepsilon \int_0^t \int_{\mathbb{R}^d} |\partial_{x_i}^2 u_\varepsilon|^2 d\mathbf{x} dt' \\
& \leq \|\partial_{x_i} u_0^\varepsilon\|_{L^2(\mathbb{R}^d)}^2 + \delta \|\nabla \partial_{x_i} u_0^\varepsilon\|_{L^2(\mathbb{R}^d)}^2 + \frac{d \|\partial_\lambda f_\varepsilon\|_\infty^2}{\varepsilon} \|\nabla u_\varepsilon\|_{L^2([0,T], L^2(\mathbb{R}^d))}^2 \\
& + \frac{3\varepsilon}{2} \int_0^t \int_{\mathbb{R}^d} |\nabla \partial_{x_i} u_\varepsilon|^2 d\mathbf{x} dt' + \frac{2}{\varepsilon} \|\operatorname{div} f^\varepsilon(\mathbf{x}, \lambda)|_{\lambda=u_\varepsilon}\|_{L^2([0,T] \times \mathbb{R}^d)}^2
\end{aligned} \tag{3.11}$$

where $\|\partial_\lambda f_\varepsilon\|_\infty = \max_{1 \leq j \leq n} \|\partial_\lambda f_j^\varepsilon\|_{L^\infty(\mathbb{R}^{d+1})}$. Multiplying the above inequality by δ , we have

$$\begin{aligned}
& \delta \int_{\mathbb{R}^d} |\partial_{x_i} u_\varepsilon(t, \mathbf{x})|^2 d\mathbf{x} + \frac{\varepsilon \delta}{2} \int_0^t \int_{\mathbb{R}^d} |\nabla \partial_{x_i} u_\varepsilon|^2 d\mathbf{x} dt + \delta^2 \int_{\mathbb{R}^d} |\nabla \partial_{x_i} u_\varepsilon(t, \mathbf{x})|^2 d\mathbf{x} \\
& \leq \delta \|\partial_{x_i} u_0^\varepsilon\|_{L^2(\mathbb{R}^d)}^2 + \delta^2 \|\nabla \partial_{x_i} u_0^\varepsilon\|_{L^2(\mathbb{R}^d)}^2 + \frac{d\delta}{\varepsilon} \|\partial_\lambda f_\varepsilon\|_\infty^2 \|\nabla u_\varepsilon\|_{L^2([0,T], L^2(\mathbb{R}^d))}^2 \\
& + \frac{2\delta}{\varepsilon} \|\operatorname{div} f^\varepsilon(\mathbf{x}, \lambda)|_{\lambda=u_\varepsilon}\|_{L^2([0,T] \times \mathbb{R}^d)}^2 \\
& \leq \delta \|\partial_{x_i} u_0^\varepsilon\|_{L^2(\mathbb{R}^d)}^2 + \delta^2 \|\nabla \partial_{x_i} u_0^\varepsilon\|_{L^2(\mathbb{R}^d)}^2 + \frac{2T\delta}{\varepsilon} \|\operatorname{div} f^\varepsilon(\mathbf{x}, \lambda)\|_{L^2(\mathbb{R}^d, L^\infty(\mathbb{R}))}^2 \\
& + \frac{d\delta}{2\varepsilon^2} \|\partial_\lambda f_\varepsilon\|_\infty^2 (\|u_0^\varepsilon\|_{L^2(\mathbb{R}^d)}^2 + \delta \|\nabla u_0^\varepsilon\|_{L^2(\mathbb{R}^d)}^2 + 2T \|\operatorname{div}_\varepsilon(\mathbf{x}, \lambda)\|_{L^1(\mathbb{R}^d \times \mathbb{R})}),
\end{aligned}$$

where in the last inequality we used Eq. (3.10). From here, (3.7) immediately follows.

We next aim to prove inequality (3.8). To this end, we multiply equation (1.7) by $\partial_t u_\varepsilon$ and then proceed similarly to the above. Namely,

$$\begin{aligned}
& (\partial_t u_\varepsilon)^2 + \sum_{j=1}^d \partial_t u_\varepsilon \partial_\lambda f_j^\varepsilon(\mathbf{x}, u_\varepsilon) \partial_{x_j} u_\varepsilon + \partial_t u_\varepsilon \operatorname{div} f^\varepsilon(\mathbf{x}, \lambda)|_{\lambda=u_\varepsilon} \\
& = \varepsilon \partial_t u_\varepsilon \Delta u_\varepsilon + \delta \partial_t u_\varepsilon \Delta \partial_t u_\varepsilon,
\end{aligned} \tag{3.12}$$

and we integrate equation (3.12) with respect to time and space.

Integration by part gives

$$\begin{aligned}
& \int_0^t \int_{\mathbb{R}^d} |\partial_t u_\varepsilon|^2 d\mathbf{x} dt' + \sum_{j=1}^d \int_0^t \int_{\mathbb{R}^d} \partial_t u_\varepsilon \partial_\lambda f_j^\varepsilon(\mathbf{x}, u_\varepsilon) \partial_{x_j} u_\varepsilon d\mathbf{x} dt' \\
& + \int_0^t \int_{\mathbb{R}^d} \partial_t u_\varepsilon \operatorname{div} f^\varepsilon(\mathbf{x}, \lambda)|_{\lambda=u_\varepsilon} d\mathbf{x} d\lambda \\
& = \varepsilon \int_0^t \int_{\mathbb{R}^d} \Delta u_\varepsilon \partial_t u_\varepsilon d\mathbf{x} dt' + \delta \int_0^t \int_{\mathbb{R}^d} \partial_t u_\varepsilon \Delta \partial_t u_\varepsilon d\mathbf{x} dt' \\
& = - \sum_{j=1}^d \int_0^t \int_{\mathbb{R}^d} \varepsilon \partial_{x_j} u_\varepsilon \partial_{tx_j} u_\varepsilon d\mathbf{x} dt' - \delta \int_0^t \int_{\mathbb{R}^d} |\nabla \partial_t u_\varepsilon|^2 d\mathbf{x} dt',
\end{aligned}$$

i.e.,

$$\begin{aligned}
& \int_0^t \int_{\mathbb{R}^d} |\partial_t u_\varepsilon|^2 d\mathbf{x} dt' + \delta \int_0^t \int_{\mathbb{R}^d} |\nabla \partial_t u_\varepsilon|^2 d\mathbf{x} dt' \\
&= -\varepsilon \sum_{j=1}^d \int_0^t \int_{\mathbb{R}^d} \partial_{x_j} u_\varepsilon \partial_{tx_j} u_\varepsilon d\mathbf{x} dt' \\
&\quad - \sum_{j=1}^d \int_0^t \int_{\mathbb{R}^d} (\partial_t u_\varepsilon \partial_\lambda f_j^\varepsilon(\mathbf{x}, u_\varepsilon) \partial_{x_j} u_\varepsilon - \partial_t u_\varepsilon \partial_{x_j} f_j^\varepsilon(\mathbf{x}, \lambda)|_{\lambda=u_\varepsilon}) d\mathbf{x} dt' \\
&\leq \|\partial_\lambda f\|_\infty \sum_{j=1}^d \int_0^t \int_{\mathbb{R}^d} |\partial_t u_\varepsilon| |\partial_{x_j} u_\varepsilon| d\mathbf{x} dt' \\
&\quad + \|\partial_t u_\varepsilon\|_{L^2([0,T] \times \mathbb{R}^d)} \sqrt{T} \|\operatorname{div} f_\varepsilon\|_{L^2(\mathbb{R}^d, L^\infty(\mathbb{R}))} - \frac{\varepsilon}{2} \sum_{j=1}^d \int_0^t \int_{\mathbb{R}^d} (|\partial_{x_j} u_\varepsilon|^2)_t d\mathbf{x} dt' \\
&\leq \frac{3}{4} \int_0^t \int_{\mathbb{R}^d} |\partial_t u_\varepsilon|^2 d\mathbf{x} dt' + \frac{d \|\partial_\lambda f_\varepsilon\|_\infty^2}{2} \int_0^t \int_{\mathbb{R}^d} |\nabla u_\varepsilon|^2 d\mathbf{x} dt' \\
&\quad + T \|\operatorname{div} f_\varepsilon\|_{L^2(\mathbb{R}^d, L^\infty(\mathbb{R}))}^2 - \frac{\varepsilon}{2} \int_{\mathbb{R}^d} |\nabla u_\varepsilon(t)|^2 d\mathbf{x} + \frac{\varepsilon}{2} \int_{\mathbb{R}^d} |\nabla u_0^\varepsilon|^2 d\mathbf{x}.
\end{aligned} \tag{3.13}$$

By multiplying (3.13) by δ and taking into account (3.5), we conclude

$$\begin{aligned}
& \frac{\delta}{4} \int_0^t \int_{\mathbb{R}^d} |\partial_t u_\varepsilon|^2 d\mathbf{x} dt' + \frac{\varepsilon \delta}{2} \int_{\mathbb{R}^d} |\nabla u_\varepsilon(t)|^2 d\mathbf{x} + \delta^2 \int_0^t \int_{\mathbb{R}^d} |\nabla \partial_t u_\varepsilon|^2 d\mathbf{x} dt' \\
&\leq \frac{\delta d \|\partial_\lambda f_\varepsilon\|_\infty^2}{2} \int_0^t \int_{\mathbb{R}^d} |\nabla u_\varepsilon|^2 d\mathbf{x} dt' + \frac{\delta \varepsilon}{2} \int_{\mathbb{R}^d} |\nabla u_0^\varepsilon|^2 d\mathbf{x} + \delta T \|\operatorname{div} f_\varepsilon\|_{L^2(\mathbb{R}^d, L^\infty(\mathbb{R}))}^2 \\
&\leq \frac{\delta d \|\partial_\lambda f_\varepsilon\|_\infty^2}{4\varepsilon} \|u_0^\varepsilon\|_{L^2(\mathbb{R}^d)}^2 + \left(\frac{\delta^2 d \|\partial_\lambda f_\varepsilon\|_\infty^2}{4\varepsilon} + \frac{\delta \varepsilon}{2} \right) \|\nabla u_0^\varepsilon\|_{L^2(\mathbb{R}^d)}^2 \\
&\quad + \frac{\delta d \|\partial_\lambda f_\varepsilon\|_\infty^2 T}{2\varepsilon} \|\operatorname{div} f_\varepsilon\|_{L^1(\mathbb{R}^d \times \mathbb{R})} + \delta T \|\operatorname{div} f_\varepsilon\|_{L^2(\mathbb{R}^d, L^\infty(\mathbb{R}))}^2,
\end{aligned} \tag{3.14}$$

which gives the inequality (3.8). $\square$

We shall now derive the kinetic formulation for (1.7) and then show that we can apply the velocity averaging result on this form of the equation.

Lemma 3.5. (Kinetic formulation of (1.7)) *Assume that the function u_ε satisfies (1.7). Then, the function $h_\varepsilon(t, \mathbf{x}, \lambda) = \operatorname{sgn}(u_\varepsilon(t, \mathbf{x}) - \lambda)$ solves a linear PDE of the form*

$$\partial_t h_\varepsilon + \operatorname{div}(h_\varepsilon \partial_\lambda f_\varepsilon(\mathbf{x}, \lambda)) - \partial_\lambda (h_\varepsilon \operatorname{div} f_\varepsilon(\mathbf{x}, \lambda)) = \partial_\lambda \operatorname{div}(G_1^\varepsilon) + \partial_\lambda^2 G_2^\varepsilon, \tag{3.15}$$

where

- (i) $\operatorname{div} G_1^\varepsilon \rightarrow 0$ in $W_{loc}^{-1,r}(\mathbb{R}^+ \times \mathbb{R}^d \times \mathbb{R})$ for any $r \in [1, 2)$;

(ii) (G_2^ε) is bounded in the space of measures $\mathcal{M}_{loc}(\mathbb{R}^+ \times \mathbb{R}^d \times \mathbb{R})$ and thus strongly precompact in $W_{loc}^{-1,r}(\mathbb{R}^+ \times \mathbb{R}^d \times \mathbb{R})$ for any $r \in [1, \frac{d+2}{d+1})$.

Proof. Before we start, notice that we require merely local convergence (see conditions (i) and (ii) of the lemma), which means that we can always choose ε small enough so that $K_\varepsilon \equiv 1$ on the fixed relatively compact set on which we derive the estimates.

Therefore, it is enough to prove the lemma under the simplifying assumption

$$\mathbf{f}_\varepsilon = \mathbf{f} \star \omega_{n(\varepsilon)}.$$

Also, in order to avoid proliferation of symbols, we will notationally adhere to integration over the entire space $\mathbb{R}^d \times \mathbb{R}$, with the implicit understanding we are in fact integrating over the support of the corresponding test function (see (3.19)).

Then, note that for any $\eta \in C^3(\mathbb{R})$ such that $\eta' \in C_c^2(\mathbb{R})$ we have (recall that $h_\varepsilon(t, \mathbf{x}, \lambda) = \text{sgn}(u_\varepsilon(t, \mathbf{x}) - \lambda)$)

$$\eta(u_\varepsilon) = \frac{1}{2} \int_{\mathbb{R}} \eta'(\lambda) h_\varepsilon(t, \mathbf{x}, \lambda) d\lambda + \frac{\eta(-R) + \eta(R)}{2} \quad (3.16)$$

for $R > 0$ such that $\text{supp}(\eta') \subset [-R, R]$. Moreover, for any function $f \in L^1(\mathbb{R})$

$$\int_{-\infty}^{u_\varepsilon} f(\lambda) d\lambda = \frac{1}{2} \int_{\mathbb{R}} f(\lambda) h_\varepsilon(t, \mathbf{x}, \lambda) d\lambda + \frac{1}{2} \int_{\mathbb{R}} f(\lambda) d\lambda. \quad (3.17)$$

To proceed, we multiply equation (1.7) by $\eta'(u_\varepsilon)$. We get

$$\begin{aligned} & \partial_t \eta(u_\varepsilon) + \text{div} \int_{-\infty}^{u_\varepsilon} \eta'(\lambda) \partial_\lambda \mathbf{f}_\varepsilon(\mathbf{x}, \lambda) d\lambda + \int_{-\infty}^{u_\varepsilon} \eta''(\lambda) \text{div} \mathbf{f}_\varepsilon(\mathbf{x}, \lambda) d\lambda \\ &= \sum_{j=1}^d \partial_{x_j} (\varepsilon \eta'(u_\varepsilon) \partial_{x_j} u_\varepsilon + \delta \eta'(u_\varepsilon) \partial_{t, x_j} u_\varepsilon) \\ & \quad - \varepsilon \eta''(u_\varepsilon) \sum_{j=1}^d |\partial_{x_j} u_\varepsilon|^2 - \delta \eta''(u_\varepsilon) \sum_{j=1}^d \partial_{x_j} u_\varepsilon \partial_{t, x_j} u_\varepsilon. \end{aligned} \quad (3.18)$$

Now, we apply (3.16) and (3.17) and rewrite the latter equation in the form

$$\begin{aligned} & \frac{1}{2} \partial_t \int_{\mathbb{R}} \eta'(\lambda) h_\varepsilon d\lambda + \text{div} \int_{\mathbb{R}} \eta'(\lambda) \partial_\lambda \mathbf{f}_\varepsilon(\mathbf{x}, \lambda) h_\varepsilon d\lambda + \int_{\mathbb{R}} \eta''(\lambda) \text{div} \mathbf{f}_\varepsilon(\mathbf{x}, \lambda) h_\varepsilon d\lambda \\ &= \sum_{j=1}^d \partial_{x_j} \left(\varepsilon \int_{\mathbb{R}} \eta''(\lambda) h_\varepsilon d\lambda \partial_{x_j} u_\varepsilon + \delta \int_{\mathbb{R}} \eta''(\lambda) h_\varepsilon d\lambda \partial_{t, x_j} u_\varepsilon \right) \\ & \quad - \varepsilon \int_{\mathbb{R}} \eta'''(\lambda) h_\varepsilon d\lambda \sum_{j=1}^d |\partial_{x_j} u_\varepsilon|^2 - \delta \int_{\mathbb{R}} \eta'''(\lambda) h_\varepsilon d\lambda \sum_{j=1}^d \partial_{x_j} u_\varepsilon \partial_{t, x_j} u_\varepsilon. \end{aligned}$$

If we rewrite the latter in the variational formulation, we get for every test function $\phi \in C_c^2(\mathbb{R}^+ \times \mathbb{R}^d)$

$$- \int_{\mathbb{R}^+ \times \mathbb{R}^d \times \mathbb{R}} \left(\frac{1}{2} \partial_t \phi(t, \mathbf{x}) \eta'(\lambda) + \nabla \phi'(\lambda) \cdot \partial_\lambda \mathbf{f}_\varepsilon(\mathbf{x}, \lambda) \right) h_\varepsilon dt d\mathbf{x} d\lambda \quad (3.19)$$

$$\begin{aligned}
& + \int_{\mathbb{R}^+ \times \mathbb{R}^d \times \mathbb{R}} \eta''(\lambda) \phi \operatorname{div} f_\varepsilon(\mathbf{x}, \lambda) h_\varepsilon dt d\mathbf{x} d\lambda \\
& = - \sum_{j=1}^d \int_{\mathbb{R}^+ \times \mathbb{R}^d \times \mathbb{R}} h_\varepsilon \partial_{x_j} \phi \eta''(\lambda) (\varepsilon \partial_{x_j} u_\varepsilon + \delta \partial_{t, x_j} u_\varepsilon) dt d\mathbf{x} d\lambda \\
& \quad - \int_{\mathbb{R}^+ \times \mathbb{R}^d \times \mathbb{R}} \phi(t, \mathbf{x}) \eta'''(\lambda) h_\varepsilon \left(\varepsilon \sum_{j=1}^d |\partial_{x_j} u_\varepsilon|^2 - \delta \sum_{j=1}^d \partial_{x_j} u_\varepsilon \partial_{t, x_j} u_\varepsilon \right) dt d\mathbf{x} d\lambda \\
& =: - \int_{\mathbb{R}^+ \times \mathbb{R}^d \times \mathbb{R}} (\Gamma_1^\varepsilon + \Gamma_2^\varepsilon) \cdot \nabla \phi(t, \mathbf{x}) \eta''(\lambda) dt d\mathbf{x} d\lambda \\
& \quad - \int_{\mathbb{R}^+ \times \mathbb{R}^d \times \mathbb{R}} (\Gamma_3^\varepsilon + \Gamma_4^\varepsilon) \phi(t, \mathbf{x}) \eta'''(\lambda) dt d\mathbf{x} d\lambda.
\end{aligned}$$

Now, since $\eta' \in C_c^2(\mathbb{R})$ is arbitrary, if we put

$$G_1^\varepsilon = \Gamma_1^\varepsilon + \Gamma_2^\varepsilon \quad \text{and} \quad G_2^\varepsilon = \Gamma_3^\varepsilon + \Gamma_4^\varepsilon \quad (3.20)$$

we arrive at (3.15).

Next, we prove (i) and (ii). To begin with, we show that $\operatorname{div} \Gamma_1^\varepsilon \rightarrow 0$ as $\varepsilon \rightarrow 0^+$ in $H_{loc}^{-1, r}(\mathbb{R}^+ \times \mathbb{R}^d \times \mathbb{R})$, $r \leq 2$.

We have for any $\phi \in C_c^1(\mathbb{R}^+ \times \mathbb{R}^d \times \mathbb{R})$ supported in the hypercube with side length M centered at zero:

$$\begin{aligned}
|\langle \operatorname{div} \Gamma_1^\varepsilon, \phi \rangle| & = \left| \sum_{j=1}^d \int_0^T \int_{\mathbb{R}^{d+1}} \varepsilon h_\varepsilon u_{x_j}^\varepsilon \phi_{x_j}(t, \mathbf{x}, \lambda) d\mathbf{x} d\lambda dt \right| \\
& \leq \varepsilon M^{1/2} \|\phi\|_{H^1(\mathbb{R}^+ \times \mathbb{R}^d \times \mathbb{R})} \|\nabla u^\varepsilon\|_{L^2([0, T] \times \mathbb{R}^d)}.
\end{aligned} \quad (3.21)$$

To proceed further, we use (3.5). For estimating the term $\|\operatorname{div}(f_\varepsilon)\|_{L^1(\mathbb{R}^{d+1})}$ appearing in (3.5), we use approximations of the form $f_\varepsilon := f \star \omega_{n(\varepsilon)}$ as in (1.8) with $n(\varepsilon)$ satisfying (1.14). Since $\delta = o(\varepsilon^2)$, the f_ε will converge to f as $\varepsilon \rightarrow 0$ in $L_{loc}^r(\mathbb{R}^d \times \mathbb{R})$ for any $1 \leq r \leq p$. Using (C2) and (C3), we get

$$\begin{aligned}
\|\operatorname{div}(f_\varepsilon)\|_{L^1(\mathbb{R}^{d+1})} & = \|\operatorname{div}(f) \star \omega_{n(\varepsilon)}\|_{L^1(\mathbb{R}^{d+1})} \\
& \leq \left\| \frac{\mu(\mathbf{x})}{1 + |\lambda|^{1+\beta}} \star \omega_{n(\varepsilon)}(\mathbf{x}, \lambda) \right\|_{L^1(\mathbb{R}^{d+1})} \\
& \leq C \mu(\mathbb{R}^d) \|\omega_{n(\varepsilon)}\|_{L^1(\mathbb{R}^{d+1})} \leq C,
\end{aligned} \quad (3.22)$$

where here and below, C denotes a generic constant that may alter from line to line. Thus (keeping in mind (3.5) and (1.11)), from (3.22) we conclude

$$|\langle \operatorname{div} \Gamma_1^\varepsilon, \phi \rangle| \leq \sqrt{\varepsilon} (n(\varepsilon)^{-1} \sqrt{\delta} + 1) C \|\phi\|_{H^1(\mathbb{R}^+ \times \mathbb{R}^d \times \mathbb{R})} \rightarrow 0^+, \quad \text{as } \varepsilon \rightarrow 0^+ \quad (3.23)$$

by (1.14).

For the next estimate, we have

$$|\langle \operatorname{div} \Gamma_2^\varepsilon, \phi \rangle| \leq 2M^{1/2} \|\phi\|_{H_0^1([0, T] \times \mathbb{R}^{d+1})} \delta \|\nabla u_t^\varepsilon\|_{L^2([0, T], L^2(\mathbb{R}^d))}. \quad (3.24)$$

Comparing with (3.8), we see that it remains to estimate the terms $\|\partial_\lambda f_\varepsilon\|_\infty$ and $\|\operatorname{div}(f_\varepsilon)\|_{L^2(\mathbb{R}^d, L^\infty(\mathbb{R}))}$. For the first one, we have

$$\begin{aligned}\|\partial_\lambda f_\varepsilon\|_\infty &= \sup_i \|(\partial_\lambda f_i) \star \omega_{n(\varepsilon)}\|_{L^\infty(\mathbb{R}^{d+1})} \\ &= \sup_i \|(\partial_\lambda f_i) \star \omega_{n(\varepsilon)}\|_{L^\infty(\mathbb{R}^d, L^\infty(\mathbb{R}))} \\ &\leq \sup_i \|\partial_\lambda f_i\|_{L^1(\mathbb{R}^d, L^\infty(\mathbb{R}))} \|\omega_{n(\varepsilon)}\|_{L^\infty(\mathbb{R}^{d+1})} \leq Cn(\varepsilon)^{-d-1}\end{aligned}\quad (3.25)$$

by (C3). For the second one, we obtain from (C2) and the Young inequality

$$\begin{aligned}\|\operatorname{div}(f_\varepsilon)\|_{L^2(\mathbb{R}^d, L^\infty(\mathbb{R}))} &\leq \left\| \frac{\mu(\mathbf{x})}{1 + |\lambda|^{1+\beta}} \star \omega_{n(\varepsilon)}(\mathbf{x}, \lambda) \right\|_{L^2(\mathbb{R}^d, L^\infty(\mathbb{R}))} \\ &= \|\mu \star \omega_{n(\varepsilon)}^{(1)}(\mathbf{x}) \frac{1}{1 + |\lambda|^{1+\beta}} \star \omega_{n(\varepsilon)}^{(2)}(\lambda)\|_{L^2(\mathbb{R}^d, L^\infty(\mathbb{R}))} \\ &\leq Cn(\varepsilon)^{-1} \|\mu \star \omega_{n(\varepsilon)}^{(1)}\|_{L^2(\mathbb{R}^d)} \\ &\leq Cn(\varepsilon)^{-1} \mu(\mathbb{R}^d) \|\omega_{n(\varepsilon)}^{(1)}\|_{L^2(\mathbb{R}^d)} \leq Cn(\varepsilon)^{-d/2-1}.\end{aligned}\quad (3.26)$$

Combining (3.24), (3.25), (3.8), (3.22), and (1.11), we see that

$$\begin{aligned}&|\langle \operatorname{div} \Gamma_2^\varepsilon, \phi \rangle| \\ &\leq C\delta^{1/2} \left(\frac{1}{\sqrt{\varepsilon}n(\varepsilon)^{d+1}} + \left(\frac{\sqrt{\delta}}{\sqrt{\varepsilon}n(\varepsilon)^{d+1}} + \sqrt{\varepsilon} \right) \frac{1}{n(\varepsilon)} + \frac{1}{\sqrt{\varepsilon}n(\varepsilon)^{d+1}} + \frac{1}{n(\varepsilon)^{d/2+1}} \right) \\ &\quad \cdot \|\phi\|_{H_0^1([0, T] \times \mathbb{R}^{d+1})} \rightarrow 0^+ \text{ as } \varepsilon \rightarrow 0^+.\end{aligned}\quad (3.27)$$

by (1.14). This implies claim (i) of the current lemma. Next, using (3.5), (3.8), (1.11), (3.22), and (3.25), we estimate

$$\begin{aligned}|\langle \Gamma_4^\varepsilon, \phi \rangle| &\leq \delta \left| \sum_{j=1}^d \int_0^T \int_{[-M, M]^{d+1}} |u_{x_j}^\varepsilon| |u_{t, x_j}^\varepsilon| |\phi(t, \mathbf{x}, \lambda)| d\mathbf{x} dt d\lambda \right| \\ &\leq M\delta \|\phi\|_\infty \sum_{j=1}^d \int_0^T \int_{\mathbb{R}^d} |u_{x_j}^\varepsilon| |u_{t, x_j}^\varepsilon| d\mathbf{x} dt \\ &\leq M\delta \|\phi\|_\infty \|\nabla u^\varepsilon\|_{L^2([0, T] \times \mathbb{R}^d)} \|\nabla u_t^\varepsilon\|_{L^2([0, T] \times \mathbb{R}^d)} \\ &\leq C\sqrt{\delta} \|\phi\|_\infty \left(\frac{1}{\sqrt{\varepsilon}} + \frac{\sqrt{\delta}}{\sqrt{\varepsilon}n(\varepsilon)} \right) \left[\frac{1}{\sqrt{\varepsilon}n(\varepsilon)^{d+1}} + \right. \\ &\quad \left. + \left(\frac{\sqrt{\delta}}{\sqrt{\varepsilon}n(\varepsilon)^{d+1}} + \sqrt{\varepsilon} \right) \frac{1}{n(\varepsilon)} + \frac{1}{\sqrt{\varepsilon}n(\varepsilon)^{d+1}} + \frac{1}{n(\varepsilon)^{d/2+1}} \right],\end{aligned}\quad (3.28)$$

which goes to 0 as $\varepsilon \rightarrow 0+$ by (1.14).

Finally, it remains to estimate Γ_3^ε . Using (3.5), (1.11), and (3.22), we obtain

$$\begin{aligned}|\langle \Gamma_3^\varepsilon, \phi \rangle| &\leq \varepsilon \left| \sum_{j=1}^d \int_0^T \int_{[-M, M]^{d+1}} |u_{x_j}^\varepsilon(t, \mathbf{x})|^2 \phi(t, \mathbf{x}, \lambda) d\mathbf{x} dt d\lambda \right| \\ &\leq \varepsilon M \|\phi\|_\infty \|\nabla u^\varepsilon\|_{L^2([0, T] \times \mathbb{R}^d)}^2 \\ &\leq C \|\phi\|_\infty \left(1 + \frac{\delta}{n(\varepsilon)^2} \right) \leq C \|\phi\|_\infty.\end{aligned}\quad (3.29)$$

From the estimates given above, we see that $\operatorname{div}\Gamma_1^\varepsilon$ and $\operatorname{div}\Gamma_2^\varepsilon$ converge strongly to zero in $H_{loc}^{-1}(\mathbb{R}^+ \times \mathbb{R}^d \times \mathbb{R})$ implying that they converge strongly in $W_{loc}^{-1,r}(\mathbb{R}^+ \times \mathbb{R}^d \times \mathbb{R})$, $r \in [1, 2)$, as well.

As for Γ_3^ε and Γ_4^ε , they are locally bounded in the space of Radon measures (which we denote it by $\mathcal{M}(\mathbb{R}^+ \times \mathbb{R}^d \times \mathbb{R})$) and thus, they are precompact in $W_{loc}^{-1,r}(\mathbb{R}^+ \times \mathbb{R}^d \times \mathbb{R})$, $r \in [1, \frac{d+2}{d+1})$ [18, Th. 1.3.2]. This implies condition (ii) from the Lemma. $\square$

According to the previous theorem, we see that we can apply the velocity averaging lemma Theorem 3.3. Indeed, we have:

Lemma 3.6. *Any solution h_ε of (3.15) satisfies*

$$\partial_t h_\varepsilon + \operatorname{div}(h_\varepsilon \partial_\lambda f(\mathbf{x}, \lambda)) = \partial_\lambda \operatorname{div}(G_1^\varepsilon) + \partial_\lambda^2 G_2^\varepsilon + \operatorname{div}(G_3^\varepsilon) + \partial_\lambda(h_\varepsilon \operatorname{div} f_\varepsilon(\mathbf{x}, \lambda)), \quad (3.30)$$

where (G_j^ε) , $j = 1, 3$, are strongly precompact in $L_{loc}^r(\mathbb{R}^+ \times \mathbb{R}^d \times \mathbb{R})$ while (G_2^ε) and $(h_\varepsilon \operatorname{div} f_\varepsilon(\mathbf{x}, \lambda))$ are strongly precompact in $W_{loc}^{-1,r}(\mathbb{R}^+ \times \mathbb{R}^d \times \mathbb{R})$, $r \in [1, \frac{d+2}{d+1})$.

Proof. It is enough to rewrite (3.15) in the form

$$\begin{aligned} \frac{1}{2} \partial_t h_\varepsilon + \operatorname{div}(h_\varepsilon \partial_\lambda f(\mathbf{x}, \lambda)) &= \partial_\lambda \operatorname{div}(G_1^\varepsilon) + \partial_\lambda^2 G_2^\varepsilon + \operatorname{div}(h_\varepsilon (\partial_\lambda f(\mathbf{x}, \lambda) - \partial_\lambda f_\varepsilon(\mathbf{x}, \lambda))) \\ &\quad + \partial_\lambda(h_\varepsilon \operatorname{div} f_\varepsilon(\mathbf{x}, \lambda)), \end{aligned}$$

and denote

$$G_3^\varepsilon = h_\varepsilon (\partial_\lambda f(\mathbf{x}, \lambda) - \partial_\lambda f_\varepsilon(\mathbf{x}, \lambda)).$$

Clearly, since $f_\varepsilon = K_\varepsilon \cdot f \star \omega_{n(\varepsilon)}$, we have $G_3^\varepsilon \rightarrow 0$ as $\varepsilon \rightarrow 0$ in $L_{loc}^p(\mathbb{R}^d \times \mathbb{R})$ (were p is given in (C1)). On the other hand, according to condition (C2), we see that $(h_\varepsilon \operatorname{div} f_\varepsilon(\mathbf{x}, \lambda))$ is bounded in the space of measures and thus strongly precompact in $W_{loc}^{-1,r}(\mathbb{R}^+ \times \mathbb{R}^d \times \mathbb{R})$. This concludes the proof. $\square$

Now, we are ready to use Theorem 3.3. In fact, we may choose $r \in (1, \frac{d+2}{d-1})$ and $s \geq 2$ in such a way that

$$1 + \frac{1}{s} - \frac{1}{r} < 1$$

and then set $\bar{p} := (1 + \frac{1}{s} - \frac{1}{r})^{-1}$. Then, fixing $\bar{p}'$ such that $\frac{1}{\bar{p}} + \frac{1}{\bar{p}'} = 1$, the assumptions of Theorem 3.1 precisely allow to apply Theorem 3.3 (noting that (h_ε) is bounded in any L_{loc}^s , hence contains a weakly convergent subsequence). More precisely, one easily checks that any $\bar{p} \in (\frac{2d+4}{d+4}, \infty)$, and thereby any $\bar{p}' \in (1, 2 + \frac{4}{d})$ can be obtained in this way. This is the reason for the specific assumption in Theorem 3.1.

Consequently, for any sequence $\varepsilon_n \searrow 0$ and any $\rho \in L^2(\mathbb{R})$, setting $h_n := h_{\varepsilon_n}$ and $u_n := u_{\varepsilon_n}$, the sequence $\left(\int_{\mathbb{R}} \rho(\lambda) h_n(t, \mathbf{x}, \lambda) d\lambda \right)$ is strongly precompact in $L_{loc}^1(\mathbb{R}^+ \times \mathbb{R}^d)$. As we shall see in the next theorem, this implies strong convergence of the sequence (u_n) . We first need the following auxiliary result.

Lemma 3.7. *Assume that the sequence (u_n) is bounded in $L^p(\Omega)$, $\Omega \Subset \mathbb{R}^d$, $d \in \mathbb{N}$, for some $p \geq 1$. Define*

$$\Omega_n^l = \{\mathbf{x} \in \Omega : |u_n(\mathbf{x})| > l\}.$$

Then,

$$\lim_{l \rightarrow \infty} \sup_{n \in \mathbb{N}} \operatorname{meas}(\Omega_n^l) = 0. \quad (3.31)$$

Proof. Since (u_n) is bounded in $L^p(\Omega)$ and thus in $L^1(\Omega)$ as well (since $\Omega \Subset \mathbb{R}^m$), we have

$$\sup_{n \in \mathbb{N}} \int_{\Omega} |u_n(\mathbf{x})| d\mathbf{x} \geq \sup_{n \in \mathbb{N}} \int_{\Omega_n^l} l d\mathbf{x} \implies \frac{1}{l} \sup_{k \in \mathbb{N}} \int_{\Omega} |u_n(\mathbf{x})| d\mathbf{x} \geq \sup_{n \in \mathbb{N}} \operatorname{meas}(\Omega_n^l),$$

implying (3.31) after letting $l \rightarrow \infty$ here. $\square$

With the above notations, we finally arrive at:

Theorem 3.8. *Under the non-degeneracy conditions (3.1), the sequence $(u_n = u_{\varepsilon_n})$ of solutions to (1.7), (1.10) strongly converges along a subsequence in $L^1_{loc}(\mathbb{R}^+ \times \mathbb{R}^d)$.*

Proof. For $l \in \mathbb{N}$ we set $\rho(\lambda) = \chi_{(-l,l)}(\lambda)$, where $\chi_{(-l,l)}$ is the characteristic function of the interval $(-l, l)$. Then, by Theorem 3.3 combined with a diagonalization argument there exists a common subsequence, again denoted by (u_n) , such that for any $l \in \mathbb{N}$:

$$\begin{aligned} \frac{1}{2} \int_{-l}^l h_n(t, \mathbf{x}, \lambda) d\lambda &= u_n \chi_{\{|u_n| \leq l\}}(t, \mathbf{x}) + l \chi_{\{u_n > l\}}(t, \mathbf{x}) - l \chi_{\{u_n < -l\}}(t, \mathbf{x}) \\ &=: T_l(u_n) \rightarrow u^l \end{aligned} \quad (3.32)$$

as $n \rightarrow \infty$ in $L^1_{loc}(\mathbb{R}^+ \times \mathbb{R})$. The operators T_l are known as truncation operators [17].

It is not difficult to prove that from here we can conclude about the convergence of (u_n) . First, we show that the sequence (u^l) converges strongly in $L^1_{loc}(\mathbb{R}^+ \times \mathbb{R})$ as $l \rightarrow \infty$.

To this end, let $\Omega \Subset \mathbb{R}^+ \times \mathbb{R}^d$. We claim that

$$\limsup_l \sup_n \|T_l(u_n) - u_n\|_{L^1(\Omega)} \rightarrow 0. \quad (3.33)$$

In fact, let

$$\Omega_n^l = \{(t, \mathbf{x}) \in \Omega : |u_n(t, \mathbf{x})| > l\}.$$

Then, since (by (3.4), (1.11), (1.14), and (3.22)) (u_n) is bounded in $L^2(\Omega)$, we have

$$\int_{\Omega} |u_n - T_l(u_n)| dt d\mathbf{x} \leq \int_{\Omega_n^l} |u_n| dt d\mathbf{x} \leq \text{meas}(\Omega_n^l)^{1/2} \|u_n\|_{L^2(\Omega)} \rightarrow 0$$

as $l \rightarrow \infty$, uniformly with respect to n according to (3.31). This gives (3.33).

Next, we estimate

$$\begin{aligned} \|u^{l_1} - u^{l_2}\|_{L^1(\Omega)} &\leq \|u^{l_1} - T_{l_1}(u_n)\|_{L^1(\Omega)} + \|T_{l_1}(u_n) - u_n\|_{L^1(\Omega)} \\ &\quad + \|T_{l_2}(u_n) - u_n\|_{L^1(\Omega)} + \|T_{l_2}(u_n) - u^{l_2}\|_{L^1(\Omega)}, \end{aligned}$$

which together with (3.32) and (3.33) implies that (u^l) is a Cauchy sequence. Thus, there exists $u \in L^1(\Omega)$ such that

$$u^l \rightarrow u \quad \text{in } L^1(\Omega). \quad (3.34)$$

Now it is not difficult to see that the entire sequence (u_n) converges toward u in $L^1(\Omega)$ as well. Namely,

$$\|u_n - u\|_{L^1(\Omega)} \leq \|u_n - T_l(u_n)\|_{L^1(\Omega)} + \|T_l(u_n) - u^l\|_{L^1(\Omega)} + \|u^l - u\|_{L^1(\Omega)},$$

which by the definition of the functions u^l , in conjunction with (3.33) and (3.34) gives the claim. $\square$

We actually proved only item (i) of Theorem 3.1. The other two items (item (ii) and item (iii)) can be proven by an adaptation of the proof for item (i). We provide a more precise explanation in the following remark.

Remark 3.9. Derivation of (ii) and (iii) from Theorem 3.1.

If we additionally assume that $\|\partial_\lambda f\| \leq C < \infty$, then we can use (3.8) in the case $\delta = \mathcal{O}(\varepsilon^2)$ and $\frac{\sqrt{\delta}}{\sqrt{\varepsilon n(\varepsilon)^{d/2+2}}} \rightarrow 0$ and the considerations thereafter will remain the same. This gives (ii). Indeed, we have to check whether the sequence of equations (3.30) satisfies the conditions of Theorem 3.3. To this end,

we need to estimate Γ_j^ε , $j = 1, \dots, 4$, appearing in (3.20). If we assume $\|\partial_\lambda f\| \leq C < \infty$, then we can omit (3.25) and get instead of (3.27)

$$\begin{aligned} |\langle \operatorname{div} \Gamma_2^\varepsilon, \phi \rangle| &\leq C \delta^{1/2} \left(\frac{1}{\sqrt{\varepsilon}} + \left(\frac{\sqrt{\delta}}{\sqrt{\varepsilon}} + \sqrt{\varepsilon} \right) \frac{1}{n(\varepsilon)} + \frac{1}{\sqrt{\varepsilon}} + \frac{1}{n(\varepsilon)^{d/2+1}} \right) \\ &\quad \cdot \|\phi\|_{H_0^1([0,T] \times \mathbb{R}^{d+1})} \rightarrow 0^+, \quad \text{as } \varepsilon \rightarrow 0^+. \end{aligned} \quad (3.35)$$

Similarly, instead of (3.28), we have

$$\begin{aligned} |\langle \Gamma_4^\varepsilon, \phi \rangle| &\leq \delta \left| \sum_{j=1}^d \int_0^T \int_{[-M,M]^{d+1}} |u_{x_j}^\varepsilon| |u_{t,x_j}^\varepsilon| |\phi(t, \mathbf{x}, \lambda)| d\mathbf{x} dt d\lambda \right| \\ &\leq C \sqrt{\delta} \|\phi\|_\infty \left(\frac{1}{\sqrt{\varepsilon}} + \frac{\sqrt{\delta}}{\sqrt{\varepsilon} n(\varepsilon)} \right) \left[\frac{1}{\sqrt{\varepsilon}} + \left(\frac{\sqrt{\delta}}{\sqrt{\varepsilon}} + \sqrt{\varepsilon} \right) \frac{1}{n(\varepsilon)} + \frac{1}{\sqrt{\varepsilon}} + \frac{1}{n(\varepsilon)^{d/2+1}} \right] \leq C \|\phi\|_\infty \end{aligned} \quad (3.36)$$

If we have a regular flux, i.e., $f \in C^1(\mathbb{R}^d \times \mathbb{R})$ and the conditions $\delta = o(\varepsilon^2)$ and $\frac{\sqrt{\delta}}{\sqrt{\varepsilon} n(\varepsilon)^{d/2+2}} \rightarrow 0$ as $\varepsilon \rightarrow 0$, we can use Young measures in the way given in [31] to derive the convergence. Actually, in this case the diffusion given by the second-order term in (1.7) will dominate over the dynamic capillarity given by the third-order term in (1.7), and therefore, we will end up with the unique Kruzhkov admissible solution to (1.6) with the corresponding initial data.

Indeed, it is not difficult to see that under the regularity assumptions on f and $\delta = o(\varepsilon^2)$, we have

$$\begin{aligned} & - \int_{\mathbb{R}^+ \times \mathbb{R}^d \times \mathbb{R}} (\Gamma_1^\varepsilon + \Gamma_2^\varepsilon) \cdot \nabla \phi(t, \mathbf{x}) \eta''(\lambda) dt d\mathbf{x} d\lambda \rightarrow 0 \\ & - \int_{\mathbb{R}^+ \times \mathbb{R}^d \times \mathbb{R}} \Gamma_4^\varepsilon(t, \mathbf{x}, \lambda) \phi(t, \mathbf{x}) \eta'''(\lambda) dt d\mathbf{x} d\lambda \rightarrow 0. \end{aligned} \quad (3.37)$$

Now, denote by $\nu_{(t,\mathbf{x})}$ the Young measure (e.g., [18]) corresponding to a subsequence of the family (u^ε) . If we let $\varepsilon \rightarrow 0$ in (3.18) along the subsequence for a convex entropy η , and take (3.37) into account, we get in the sense of distributions (below, $q(\mathbf{x}, \lambda) = \int_{-\infty}^{\lambda} \eta'(v) \partial_v f(\mathbf{x}, v) dv$ is the entropy flux)

$$\partial_t \int \eta(\lambda) d\nu_{(t,\mathbf{x})}(\lambda) + \operatorname{div} \int q(\mathbf{x}, \lambda) d\nu_{(t,\mathbf{x})}(\lambda) + \int \int_{-\infty}^{\lambda} \eta''(v) \operatorname{div} f(\mathbf{x}, v) dv d\nu_{(t,\mathbf{x})}(\lambda) \leq 0.$$

Now, we simply rely on the result from [15] (see also [32, 49]) to conclude that the Young measure is unique and atomic, i.e., of the form $\nu_{(t,\mathbf{x})}(\lambda) = \delta(\lambda - u(t, \mathbf{x}))$ for the unique entropy admissible solution to the underlying conservation law (1.6).

Acknowledgements

The work is supported in part by Project P30233 of the Austrian Science Fund (FWF) and by the Lise Meitner Project M 2669 of the Austrian Science Fund (FWF). It is also supported in part by the Croatian Science Foundation under Project MiTPDE (Number IP-2018-01-2449).

Funding Open access funding provided by Austrian Science Fund (FWF).

Open Access. This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

- [1] Andreianov, B., Cancès, C.: Vanishing capillarity solutions of Buckley–Leverett equation with gravity in two-rocks medium. *Comput. Geosci.* **17**(3), 551–572 (2013)
- [2] Andreianov, B., Karlsen, K.H., Risebro, N.H.: A theory of L^1 -dissipative solvers for scalar conservation laws with discontinuous flux. *Arch. Ration. Mech. Anal.* **201**, 27–86 (2011)
- [3] Andreianov, B., Mitrović, D.: Entropy conditions for scalar conservation laws with discontinuous flux revisited. *Ann. Inst. H. Poincaré Anal. Non Linéaire* **32**, 1307–1335 (2015)
- [4] Aleksić, J., Mitrović, D., Pilipović, S.: Hyperbolic conservation laws with vanishing nonlinear diffusion and linear dispersion in heterogeneous media. *J. Evol. Equ.* **9**, 809–828 (2009)
- [5] Barenblatt, G., Zheltov, I., Kochiva, I.: Basic concepts in the theory of seepage of homogeneous liquids in fissured rocks. *J. Appl. Math. Mech.* **24**(5), 1286–1303 (1960)
- [6] Benjamin, T.B., Bona, J.L., Mahony, J.J.: Model equations for long waves in nonlinear dispersive systems. *Philos. Trans. R. Soc. Lond. Ser. A* **272**(1220), 47–78 (1972)
- [7] Bear, J.: *Dynamics of Fluids in Porous Media*. American Elsevier Publishing Company, Inc., New York (1972)
- [8] Buckley, S.E., Leverett, M.C.: Mechanism of fluid displacements in sands. *Trans. AIME* **146**, 107–116 (1942)
- [9] Cao, Y., Yin, J.X., Wang, C.P.: Cauchy problems of semilinear pseudo-parabolic equations. *J. Differ. Equ.* **246**, 4568–4590 (2009)
- [10] Coclite, G.M., di Ruvo, L.: A note on convergence of the solutions of Benjamin–Bona–Mahony type equations. *Nonlinear Anal. Real World Appl.* **40**, 64–81 (2018)
- [11] Coclite, G.M., di Ruvo, L.: A singular limit problem for conservation laws related to the Rosenau–Korteweg–de Vries equation. *J. Math. Pures Appl.* (9) **107**, 315–335 (2017)
- [12] Crasta, G., De Cicco, V., De Philippis, G., Ghiraldin, F.: Structure of solutions of multidimensional conservation laws with discontinuous flux and applications to uniqueness. *Arch. Ration. Mech. Anal.* **221**, 961–985 (2016)
- [13] Cueto-Felgueroso, L.C., Juanes, R.: Stability analysis of a phase-field model of gravity-driven unsaturated flow through porous media. *Phys. Rev. E* **79**, 13 (2009)
- [14] DiCarlo, D.A.: Modeling observed saturation overshoot with continuum additions to standard unsaturated theory. *Adv. Water Resour.* **28**, 1021–1027 (2005)
- [15] DiPerna, R.J.: Measure-valued solutions to conservation laws. *Arch. Ration. Mech. Anal.* **88**, 223–270 (1985)
- [16] DiPerna, R.J., Lions, P.L., Meyer, Y.: L^p -regularity of velocity averages. *Ann. Inst. H. Poincaré Anal. Non Linéaire* **8**, 271–287 (1991)
- [17] Dolzmann, G., Hungerbühler, N., Müller, S.: Nonlinear elliptic systems with measure valued right-hand side. *Math. Z.* **226**, 545–574 (1997)
- [18] Evans, L.C.: Weak convergence methods for nonlinear partial differential equations. In: *Regional Conference Series in Mathematics*, No. 74. Conference Board of the Mathematical Sciences (1990)
- [19] Gérard, P.: Microlocal defect measures. *Commun. Partial Differ. Equ.* **16**, 1761–1794 (1991)
- [20] Gilbarg, D., Trudinger, N.S.: *Elliptic Partial Differential Equations of Second Order*. Springer, Berlin (2001)
- [21] Golse, F., Saint-Raymond, L.: Velocity averaging in L^1 for the transport equation. *C. R. Acad. Sci. Paris Ser. I Math.* **334**, 557–562 (2002)
- [22] Hassanizadeh, S.M., Gray, W.G.: Thermodynamics basis of capillary pressure in porous media. *Water Resour. Res.* **29**, 3389–3405 (1993)
- [23] Hassanizadeh, S.M., Gray, W.G.: Mechanics and thermodynamics of multiphase flow in porous media including inter-phase boundaries. *Adv. Water Resour.* **13**, 179–186 (1990)
- [24] Holden, H., Karlsen, K.H., Mitrovic, D.: Zero diffusion-dispersion-smoothing limits for a scalar conservation law with discontinuous flux function. *Int. J. Differ. Equ.* (2009). Article ID 279818
- [25] Homsy, G.M.: Viscous fingering in porous media. *Annu. Rev. Fluid Mech.* **19**, 271–311 (1987)

- [26] Hwang, S., Tzavaras, A.E.: Kinetic decomposition of approximate solutions to conservation laws: application to relaxation and diffusion-dispersion approximations. *Commun. Partial Differ. Equ.* **27**, 1229–1254 (2002)
- [27] Kalisch, H., Mitrović, D., Nordbotten, J.M.: Rayleigh–Taylor instability of immiscible fluids in porous media. *Contin. Mech. Thermodyn.* **28**, 721–731 (2016)
- [28] Karch, G.: Asymptotic behaviour of solutions to some pseudoparabolic equations. *Math. Methods Appl. Sci.* **20**, 271–289 (1997)
- [29] Karlsen, K.H., Risebro, N.H., Towers, J.: L^1 -stability for entropy solutions of nonlinear degenerate parabolic convection-diffusion equations with discontinuous coefficients. *Skr. K. Nor. Vid. Selsk.* **3**, 1–49 (2003)
- [30] Kaikina, E.I., Naumkin, P.I., Shishmarev, I.A.: The Cauchy problem for a Sobolev type equation with power like nonlinearity. *Izv. Math.* **69**, 59–111 (2005)
- [31] Kondo, C.I., LeFloch, P.G.: Zero diffusion-dispersion limits for scalar conservation laws. *SIAM J. Math. Anal.* **33**, 1320–1329 (2002)
- [32] Kondo, C., LeFloch, P.G.: Measure-valued solutions and well-posedness of multi-dimensional conservation laws in a bounded domain. *Port. Math.* **58**, 171–194 (2001)
- [33] Kruzhkov, S.N.: First order quasilinear equations in several independent variables. *Matorsz. Sb.* **81**, 217–243 (1970)
- [34] Kwek, K.H., Qu, C.C.: Alternative principle for pseudo-parabolic equations. *Dyn. Syst. Appl.* **5**, 211–217 (1996)
- [35] Lazar, M., Mitrović, D.: Velocity averaging: a general framework. *Dyn. PDEs* **9**, 239–260 (2012)
- [36] Lazar, M., Mitrović, D.: Existence of solutions to a scalar conservation laws with a flux of low regularity. *Electron. J. Differ. Equ.* **2016**(325), 1–18 (2016)
- [37] Lions, P.-L., Perthame, B., Tadmor, E.: A kinetic formulation of multidimensional scalar conservation laws and related equations. *J. Am. Math. Soc.* **7**, 169–191 (1994)
- [38] LeFloch, P.G.: *Hyperbolic Systems of Conservation Laws. The Theory of Classical and Nonclassical Shock Waves*, Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel (2002). ISBN: 3-7643-6687-7
- [39] Mikelić, A.: A global existence result for the equations describing unsaturated flow in porous media with dynamic capillary pressure. *J. Differ. Equ.* **248**, 1561–1577 (2010)
- [40] Oleinik, O.A.: Discontinuous solutions of nonlinear partial differential equations. *Dokl. Akad. Nauk SSSR* **95**, 451–455 (1954)
- [41] Padron, V.: Effect of aggregation on population recovery modeled by a forward–backward pseudoparabolic equation. *Trans. Am. Math. Soc.* **356**, 2739–2756 (2004)
- [42] Perthame, B., Souganidis, P.: A limiting case for velocity averaging. *Ann. Sci. Ec. Norm. Sup.* **4**, 591–598 (1998)
- [43] Panov, E.Y.: Existence and strong pre-compactness properties for entropy solutions of a first-order quasilinear equation with discontinuous flux. *Arch. Ration. Mech. Anal.* **195**, 643–673 (2010)
- [44] Ptashnyk, M.: Degenerate quasilinear pseudoparabolic equations with memory terms and variational inequalities. *Nonlinear Anal. TMA* **66**, 2653–2675 (2007)
- [45] Richards, L.A.: Capillary conduction of liquids through porous mediums. *Physics* **1**, 318–333 (1931)
- [46] Schonbek, M.E.: Convergence of solutions to nonlinear dispersive equations. *Commun. Partial Differ. Equ.* **7**, 959–1000 (1982)
- [47] Selberg, S., Tesfahun, A.: Low regularity well-posedness for some nonlinear Dirac equations in one space dimension. *Differ. Integral Equ.* **23**(3–4), 265–278 (2010)
- [48] Showalter, R.E., Ting, T.W.: Pseudoparabolic partial differential equations. *SIAM J. Math. Anal.* **1**, 1–26 (1970)
- [49] Szepessy, A.: An existence result for scalar conservation laws using measure-valued solutions. *Commun. Partial Differ. Equ.* **14**, 1329–1350 (1989)
- [50] Tao, T., Tadmor, E.: Velocity averaging, kinetic formulations, and regularizing effects in quasi-linear partial differential equations. *Commun. Pure Appl. Math.* **60**, 1488–1521 (2007)
- [51] Ting, T.W.: Certain non-steady flows of second-order fluids. *Arch. Ration. Mech. Anal.* **14**, 1–26 (1963)
- [52] Ting, T.W.: Parabolic and pseudo-parabolic partial differential equations. *J. Math. Soc. Jpn.* **21**, 440–453 (1969)
- [53] van Duijn, C.J., Peletier, L.A., Pop, I.S.: A new class of entropy solutions of the Buckley–Leverett equation *SIAM. J. Math. Anal.* **39**, 507–536 (2007)

M. Graf
 Department of Mathematics
 University of Washington
 C-138 Padelford Hall
 Seattle WA 98195-4350
 USA
 e-mail: mgraf2@uw.edu

M. Kunzinger and D. Mitrovic
Faculty of Mathematics
University of Vienna
Oskar-Morgenstern-Platz 1
1090 Vienna
Austria
e-mail: darko.mitrovic@univie.ac.at

M. Kunzinger
e-mail: michael.kunzinger@univie.ac.at

D. Mitrovic and D. Vujadinovic
Faculty of Mathematics
University of Montenegro
Cetinjski put bb
81000 Podgorica
Montenegro
e-mail: djordjijjevuj@ucg.ac.me

(Received: April 28, 2020; revised: November 3, 2020; accepted: November 10, 2020)