UNIVERSITY OF
BATH

University of Bath

CURVES OF STEEPEST DESCENT ARE ENTROPY SOLUTIONS FOR A CLASS OF DEGENERATE CONVECTION-DIFFUSION EQUATIONS

MARCO DI FRANCESCO AND DANIEL MATTHES

ABSTRACT. We consider a nonlinear degenerate convection-diffusion equation with inhomogeneous convection and prove that its entropy solutions in the sense of Kružkov are obtained as the — a posteriori unique — limit points of the JKO variational approximation scheme for an associated gradient flow in the L^2-Wasserstein space. The equation lacks the necessary convexity properties which would allow to deduce well-posedness of the initial value problem by the abstract theory of metric gradient flows. Instead, we prove the entropy inequality directly by variational methods and conclude uniqueness by doubling of the variables.

1. INTRODUCTION

The goal of this paper is to show on the example of the degenerate parabolic equation

$$u_t = (u^m)_{yy} + (b(y)u^m)_y, \quad y \in \mathbb{R}, \ \ t \geq 0, \tag{1}$$

how the solution concepts of *metric gradient flows* and *entropy solutions* can be combined to obtain global in time well-posedness of the initial value problem. Specifically, we define an energy functional $\mathcal{F}$, construct time-discrete curves of steepest descent in the landscape of $\mathcal{F}$ with respect to the L^2-Wasserstein metric by means of the JKO variational scheme [16], and show that these curves converge in the limit of continuous time to entropy solutions à la Kružkov [19, 20] (or, more precisely, in the sense of Carrillo [3]) for (1). This way, existence of solutions to (1) is obtained by a variational method, and their uniqueness follows from the theory of entropy solutions.

In (1), the exponent $m > 1$ determines the degeneracy of the diffusion for vanishing densities, and

$$b \in L^1(\mathbb{R}) \cap W^{1,\infty}(\mathbb{R}) \tag{2}$$

is a given coefficient modelling heterogeneity in the convection part. As initial condition for (1), we prescribe a non-negative datum $u^0 \in L^1(\mathbb{R}) \cap L^\infty(\mathbb{R})$ of finite second moment.

The link to metric gradient flows is established as follows. There is a spatial coordinate change $y = \mathfrak{T}(x)$ and an according transformation $\rho(t, x) = \mathfrak{T}_x(x)u(t, \mathfrak{T}(x))$ (see subsection 2.1 below) such that all sufficiently regular solutions u to (1) become solutions of

$$\rho_t = \left(\rho[a(x)\rho^{m-1}]_x\right)_x, \tag{3}$$

and vice versa; here $a \in W^{2,\infty}(\mathbb{R})$ is a strictly positive function, determined from b via $\mathfrak{T}$. Well-known formal arguments, see e.g. [28], indicate that the time-dependent density functions $\rho(t)$ satisfying (3) are "curves of steepest descent" in the energy landscape of the entropy functional

$$\mathcal{F}(\rho) := \frac{1}{m} \int_{\mathbb{R}} a(x)\rho^m \, \mathrm{d}x, \tag{4}$$

with respect to the L^2-Wasserstein distance.

Indeed, if $\mathcal{F}$ would be a geodesically λ-convex functional, then the theory of λ-contractive gradient flows — see e.g. [1] — could be applied to conclude the existence of a unique gradient flow for $\mathcal{F}$ in the space of probability measures. This flow's curves would be weak solutions to (3), and by inversion of the coordinate transformation above, we could conclude well-posedness for (1). Unfortunately, $\mathcal{F}$ does apparently not have the required convexity property; see Section 3 for details.

Date: March 15, 2013.

Nevertheless, the variational structure behind (3) is on the basis for our proof of *existence* of solutions to (1). Specifically, we employ the JKO (or "minimizing movement") approximation scheme to obtain time-discrete curves ρ_τ of steepest descent for $\mathcal{F}$. We prove that any family of such discrete approximations ρ_τ possesses a weak time-continuous limit curve ρ_*. By inversion of the coordinate change, this provides a candidate u_* for a solution to (1). We emphasize that this construction does not require geodesic λ-convexity of $\mathcal{F}$; boundedness from below, coercivity and lower semi-continuity are sufficient.

At this point, the variational framework of minimizing movements provides strong tools which allow us to prove that u_* is actually an entropy solution for (1) in the sense of Carrillo [3]: we show that $u_*^m \in L^2_{\mathrm{loc}}(0, \infty; H^1(\mathbb{R}))$, and that for every non-negative test function $\varphi \in C_c^\infty(\mathbb{R}_+ \times \mathbb{R})$ and for every $k \in \mathbb{R}_+$:

$$\int_0^T \int_{\mathbb{R}} |u_* - k|\varphi_t \, \mathrm{d}y \, \mathrm{d}t$$
$$- \int_0^T \int_{\mathbb{R}} \mathrm{Sgn}(u_* - k)\Big(\big[(u_*^m)_y + b(u_*^m - k^m) \big] \varphi_y - b_y k^m \varphi \Big) \, \mathrm{d}y \, \mathrm{d}t \geq \mathfrak{D}(u_*)[\varphi] \geq 0. \tag{5}$$

We also obtain a non-trivial lower bound on the dissipation $\mathfrak{D}(u_*)$; see Proposition 4.11 for details.

The derivation of the entropy inequality (5) is the core element of our proof. First, a time-discrete version of this estimate is proven directly for the JKO scheme, and this is passed to the limit. In the derivation of the discrete estimates, the key idea is — as usual — to choose appropriate variations of the minimizers. Here we build on the ideas that have already been employed in [16] for the derivation of the weak formulation of the linear Fokker-Planck equation and have been generalized later in [22] to the "flow interchange lemma", see Lemma 4.2: variations are performed by an auxiliary gradient flow, which — in contrast to the gradient flow of (4) itself — is λ-convex and thus satisfies certain variational inequalities.

Once that (5) has been establish, we adapt the doubling of the variables method in [18] to our case and show that u_* is actually the *unique* entropy solution for the given initial condition u^0. A posteriori, we conclude uniqueness of the limit curve ρ_* of the approximation scheme as well.

The following theorem summarizes our main result in an informal way; the precise statement is given in Theorem 2.7 in Section 2.4. For a possible slight generalization avoiding the scaling, see Remark 2.9.

Theorem 1.1. *Let an initial condition $u^0 \in L^\infty(\mathbb{R})$ of finite second moment be given. Then every curve obtained from the JKO approximation for $\mathcal{F}$ in the limit of continuous time corresponds — by a chance of coordinates — to the unique entropy solution for* (1).

A similar observation about the connection between entropy solutions and gradient flows has been made recently by Gigli and Otto [15] in the context of the inviscid Burger's equation, see also [6, 24] for related results. The interpretation of the coincidence between entropy solutions and gradient flows is that both types of solutions can be characterized by diminishing an underlying entropy functional "as fast as possible". For previous results on the well-posedness of scalar conservation laws in Wasserstein spaces, we refer to [2, 4].

Apart from revealing an interesting connection between the two seemingly unrelated theories of entropy solutions and Wasserstein gradient flows, our simple example indicates a possible general strategy to prove existence *and* uniqueness of certain nonlinear evolution equations which can be cast in the form of a Wasserstein gradient flow of a functional which is not necessarily geodesically λ-convex. First, use methods from the calculus of variations to construct a candidate for a solution; a priori, there might be several. Second, show that this candidate is an entropy solution by deriving further a priori estimates in the variational framework. Third, conclude uniqueness using doubling of the variables. This strategy provides a new method to obtain entropy solutions, alternative e.g. to the classical vanishing viscosity approach [8], to the wave-front-tracking algorithm [7], or to semigroup theory [5]. In particular, the variational approach does not require the solution of auxiliary regularized problems.

To conclude this introduction, we stress again that the goal of this paper is to establish a link between gradient flows and entropy solutions but not to refine the results on existence and

uniqueness of weak solutions to (1). This said, we remark that — as far as our model equation (1) is concerned — uniqueness of weak L^1-solutions could be proven also by other means, e.g. with the methods developed in [23].

The paper is organized as follows. In Section 2, we formulate the problem and state our main results. Section 3 recalls some basic facts from the theory of gradient flows. In Lemma 3.4, we provide a result on the contractivity of quite general reaction-diffusion equations in the Wasserstein metric which might be of independent interest. In Section 4 we prove convergence of the scheme and show that the limit curves are entropy solutions. Section 5 contains (for the sake of completeness) the uniqueness proof of entropy solutions.

NOTATION

Measures, densities and Wasserstein distance. Here $\mathcal{P}_2(\mathbb{R})$ denotes the set of probability *densities* $\rho \in L^1(\mathbb{R})$ with finite second moment $\int x^2 \rho(x)\, dx$. Note that the symbol $\mathcal{P}_2(\mathbb{R})$ is frequently used in the literature for the (wider) space of probability *measures* on $\mathbb{R}$, which we shall denote by $\overline{\mathcal{P}_2(\mathbb{R})}$ instead. Given a measure $\mu \in \overline{\mathcal{P}_2(\mathbb{R})}$, its probability distribution function $U : \mathbb{R} \to [0,1]$ and pseudo-inverse $G : [0,1] \to \mathbb{R} \cup \{\pm\infty\}$, respectively, are given by

$$U(x) = \mu\big((-\infty, x)\big), \quad G(\omega) = \sup\big\{x \in \mathbb{R} \,\big|\, U(x) \leq \omega\big\}.$$

The space $\overline{\mathcal{P}_2(\mathbb{R})}$ is endowed with the L^2-Wasserstein distance $\mathbf{W}_2$, defined by

$$\mathbf{W}_2(\mu, \tilde{\mu}) = \left(\int_0^1 \big[G(\omega) - \tilde{G}(\omega)\big]^2 \, d\omega \right)^{1/2}, \tag{6}$$

where G and $\tilde{G}$ are the pseudo-inverse distribution functions of μ and $\tilde{\mu}$, respectively. The pair $(\overline{\mathcal{P}_2(\mathbb{R})}, \mathbf{W}_2)$ is a complete metric space. We refer to [1, Chapter 6.0] or to [28, Chapter 7.3] for a detailed explanation.

Mollifications. We will frequently use the following elementary functions: the absolute value $\mathrm{Abs}(x) := |x|$, the positive part $\mathrm{PosPar}(x) = (x)_+ = (x + |x|)/2$, the sign $\mathrm{Sgn}(x) = x/|x|$ (with $\mathrm{Sgn}(0) = 0$), and the unit step function $\mathrm{Stp}(x) = (x + |x|)/2|x|$ (with $\mathrm{Stp}(0) = 0$). In fact, we will mostly use their regularized versions obtained by *mollification*: denote by $\delta_1 : \mathbb{R} \to \mathbb{R}$ the standard mollifier

$$\delta_1(y) = \begin{cases} Z^{-1} \exp[-1/(1 - y^2)] & \text{for all } y \in\,]-1, 1[, \\ 0 & \text{otherwise,} \end{cases} \tag{7}$$

where $Z > 0$ is chosen s.t. δ_1 has unit integral. For $\epsilon > 0$, define the ϵ-mollifier $\delta_\epsilon : \mathbb{R} \to \mathbb{R}$ by $\delta_\epsilon(y) = \epsilon^{-1} \delta_1(\epsilon^{-1} y)$. Accordingly, we denote by $\mathrm{Abs}_\epsilon = \mathrm{Abs} \star \delta_\epsilon$ and $\mathrm{PosPar}_\epsilon = \mathrm{PosPar} \star \delta_\epsilon$ the ϵ-mollifications of Abs and PosPar, respectively. Notice that $\mathrm{Sgn}'_\epsilon = 2\delta_\epsilon$ and $\mathrm{Stp}'_\epsilon = \delta_\epsilon$, which means in particular that Sgn_ϵ and Stp_ϵ are non-decreasing functions.

2. STATEMENT OF THE PROBLEM AND RESULTS

2.1. **Coordinate transformation.** In this subsection we shall establish the correspondence between equations (1) and (3). We define a spatial coordinate change $y = \mathfrak{T}(x)$ as solution to the initial value problem

$$\mathfrak{T}_x(x) = \alpha(\mathfrak{T}(x)) \quad \text{for } x \in \mathbb{R}, \qquad \mathfrak{T}(0) = 0, \tag{8}$$

where $\alpha : \mathbb{R} \to \mathbb{R}_+$ is given by

$$\alpha(y) = \alpha_0 \exp\left(-\frac{m-1}{2m} \int_0^y b(\eta) d\eta \right), \tag{9}$$

α_0 being an arbitrary positive constant. The assumptions (2) on b ensure

$$\alpha \in W^{2,\infty}(\mathbb{R}) \quad \text{and} \quad \alpha \geq \underline{\alpha} \quad \text{for some } \underline{\alpha} > 0. \tag{10}$$

Therefore, (8) admits a unique global solution $\mathfrak{T} : \mathbb{R} \to \mathbb{R}$, with $\mathfrak{T}_x(x) \geq \underline{\alpha}$. Accordingly, introduce the corresponding transformation $\mathbf{S} : L^1(\mathbb{R}) \to L^1(\mathbb{R})$ of densities by

$$\mathbf{S}[u] := \mathfrak{T}_x \, u \circ \mathfrak{T}. \tag{11}$$

Finally, the transformed inhomogeneity $a : \mathbb{R} \to \mathbb{R}$ is given by

$$a = \frac{m}{m-1}(\mathfrak{T}_x)^{-(m+1)}. \tag{12}$$

Proposition 2.1. *The function a satisfies*

  (a1) $a(x) \geq \underline{a} > 0$ for all $x \in \mathbb{R}$,
  (a2) $a \in W^{2,\infty}(\mathbb{R})$,

and the transformation $\rho(t) := \mathbf{S}[u(t)]$ of an arbitrary weak solution u to (1) with initial datum $u_0 \in L^1(\mathbb{R})$ satisfies equation (3) with initial datum $\rho^0 = \mathbf{S}[u^0]$.

Proof. By definition of a in (12), the properties (a1) and (a2) are direct consequences of (10), which are the properties induced on α by the assumptions (2) on b.

Now, let u be a weak solution to (1), that is

$$-\int_0^\infty \int_{\mathbb{R}} u\varphi_t \, \mathrm{d}y \, \mathrm{d}t = \int_0^\infty \int_{\mathbb{R}} u^m \varphi_{yy} \, \mathrm{d}y \, \mathrm{d}t - \int_0^\infty \int_{\mathbb{R}} u^m b\varphi_y \, \mathrm{d}y \, \mathrm{d}t \tag{13}$$

for all test functions $\varphi \in C_c^\infty(\mathbb{R}_+ \times \mathbb{R})$. Define the transformed test function $\psi \in C_c^\infty(\mathbb{R}_+ \times \mathbb{R})$ by

$$\psi(t, x) = \varphi(t, \mathfrak{T}(x)).$$

Using (8) and its "derivative" $\mathfrak{T}_{xx} = \alpha_y \circ \mathfrak{T} \, \mathfrak{T}_x$, one easily verifies that

$$\varphi_y(\mathfrak{T}(x), t)\mathfrak{T}_x(x) = \psi_x(t, x), \quad \varphi_{yy}(\mathfrak{T}(x), t)\mathfrak{T}_x(x)^2 = \psi_{xx}(x, t) - \psi_x(x, t)\alpha_y(\mathfrak{T}(x)).$$

Substitute these representations and

$$u(\mathfrak{T}(x), t) = \frac{\rho(x, t)}{\mathfrak{T}_x(x)} \quad \text{with} \quad \rho(t) = \mathbf{S}[u(t)] \tag{14}$$

into (13). A change of variables $(t, y) = (t, \mathfrak{T}(x))$ under the integrals yields

$$\begin{aligned}
-\int_0^\infty \int_{\mathbb{R}} \rho\psi_t \, \mathrm{d}x \, \mathrm{d}t = &\int_0^\infty \int_{\mathbb{R}} \rho^m \psi_{xx} \mathfrak{T}_x^{-(m+1)} \, \mathrm{d}x \, \mathrm{d}t \\
&- \int_0^\infty \int_{\mathbb{R}} \rho^m \psi_x \big[(\alpha_y \circ \mathfrak{T})\mathfrak{T}_x^{-(m+1)} + (b \circ \mathfrak{T})\mathfrak{T}_x^{-m}\big] \, \mathrm{d}x \, \mathrm{d}t.
\end{aligned} \tag{15}$$

It remains to rewrite the integrands using a instead of b and $\mathfrak{T}$. Since

$$\alpha_y \circ \mathfrak{T} = -\frac{m-1}{2m}(b \circ \mathfrak{T})\mathfrak{T}_x \tag{16}$$

by (8) and (9), a direct calculation yields

$$\frac{1}{m}a_x \overset{(8)}{=} -\frac{m+1}{m-1}(\alpha_y \circ \mathfrak{T})\mathfrak{T}_x^{-(m+1)} \overset{(16)}{=} (\alpha_y \circ \mathfrak{T})\mathfrak{T}_x^{-(m+1)} + (b \circ \mathfrak{T})\mathfrak{T}_x^{-m}. \tag{17}$$

Substitute (12) and (17), respectively, in the first and the second integral on the right hand side of (15), to obtain

$$-\int_0^\infty \int_{\mathbb{R}} \rho\psi_t \, \mathrm{d}x \, \mathrm{d}t = \frac{m-1}{m} \int_0^\infty \int_{\mathbb{R}} a\rho^m \psi_{xx} \, \mathrm{d}x \, \mathrm{d}t - \frac{1}{m} \int_0^\infty \int_{\mathbb{R}} a_x \rho^m \psi_x \, \mathrm{d}x \, \mathrm{d}t,$$

which is a weak formulation of (3). □

Remark 2.2. *The function a in Proposition 2.1 is not uniquely determined: different choices of $\alpha_0 > 0$ in (9) change a by a positive factor. On the other hand, if the function a is given, then b and the corresponding change of variable $y = \mathfrak{T}(x)$ can be recovered from a in a unique way. Indeed, it follows from (8) and (12) that*

$$\mathfrak{T}(x) = \int_0^x \left[\frac{m-1}{m}a(\xi)\right]^{-\frac{1}{m+1}} \mathrm{d}\xi, \tag{18}$$

which, in combination with (16), *implies further that*

$$b \circ \mathfrak{T}\, \mathfrak{T}_x = \frac{2m}{m^2 - 1}(\log a)_x.$$

Remark 2.3. *The scaling $\mathfrak{T}$ is still well-defined if the assumption $b \in L^1(\mathbb{R})$ is relaxed to $b \propto |y|^{-1}$ as $|y| \to +\infty$. However, assumptions (a1) & (a2) may be not satisfied in those cases. Therefore we require $b \in L^1(\mathbb{R})$. Let us also emphasise that the case $b \equiv \mathrm{const}$ cannot be included, since the solution to* (8) *would blow up at some negative x.*

2.2. **Entropy solutions for** (1). The notion of entropy solution for (1) used here is a variant of the one originally introduced by Carrillo [3] and later adapted by Karlsen et al. [18]. To motivate this definition, consider the usual viscous approximation of (1),

$$u_t = (b(y)u^m)_y + (u^m)_{yy} + \nu u_{yy}, \tag{19}$$

which possesses smooth and classical solutions u_ν for every $\nu > 0$, provided that the initial condition u^0 is smooth enough. The following formal considerations are made under the hypothesis that u_ν converges — locally uniformly on $\mathbb{R}_+ \times \mathbb{R}$ and in $L^2_{\mathrm{loc}}([0,T[\times\mathbb{R})$ — to a limit function as $\nu \downarrow 0$.

Similarly as in the classical approach by Kružkov [20] for scalar conservation laws, we would like to derive from (19) an evolution *inequality* for all functions of the form $|u_\nu - k|$ with given $k \in \mathbb{R}_+$. As usual, the calculations are carried out with a suitable approximation of $|u_\nu - k|$; in our case, we multiply equation (19) by $\mathrm{Sgn}_\epsilon(u_\nu^m - k^m)$ with a mollification parameter $\epsilon > 0$ and rewrite it in the following way:

$$\mathrm{Sgn}_\epsilon(u_\nu^m - k^m)(u_\nu)_t = \mathrm{Sgn}_\epsilon(u_\nu^m - k^m)\big[(u_\nu^m)_y + b(u_\nu^m - k^m) + \nu(u_\nu)_y\big]_y + \mathrm{Sgn}_\epsilon(u_\nu^m - k^m)b_y k^m. \tag{20}$$

Now let $T > 0$, integrate (20) against a non-negative test function $\varphi \in C_c^\infty(]0,T[\times\mathbb{R})$, and integrate by parts in the first term on the right-hand side:

$$-\int_0^T \int_{\mathbb{R}} \varphi\, \mathrm{Sgn}_\epsilon(u_\nu^m - k^m)(u_\nu)_t \,\mathrm{d}y\,\mathrm{d}t = \tag{21}$$

$$\int_0^T \int_{\mathbb{R}} \mathrm{Sgn}_\epsilon(u_\nu^m - k^m)\Big(\big[(u_\nu^m)_y + b(u_\nu^m - k^m) + \nu(u_\nu)_y\big]\varphi_y - b_y k^m \varphi\Big)\,\mathrm{d}y\,\mathrm{d}t \tag{22}$$

$$+ \int_0^T \int_{\mathbb{R}} \varphi b\, \mathrm{Sgn}'_\epsilon(u_\nu^m - k^m)(u_\nu^m - k^m)(u_\nu^m)_y \,\mathrm{d}y\,\mathrm{d}t \tag{23}$$

$$+ \int_0^T \int_{\mathbb{R}} \mathrm{Sgn}'_\epsilon(u_\nu^m - k^m)\big[(u_\nu^m)_y^2 + m\nu u_\nu^{m-1}(u_\nu)_y^2\big]\varphi \,\mathrm{d}y\,\mathrm{d}t. \tag{24}$$

For further simplification, observe that inside the integrand in (21),

$$\lim_{\epsilon\downarrow 0} \mathrm{Sgn}_\epsilon(u_\nu^m - k^m)(u_\nu)_t = \mathrm{Sgn}(u_\nu - k)(u_\nu)_t = |u_\nu - k|_t$$

at every point (t,y) with $u(t,y) \neq k$. Likewise, in (22), the term with ν becomes $\mathrm{Sgn}(u_\nu - k)(u_\nu)_y = |u_\nu - k|_y$ in the limit $\epsilon \downarrow 0$. After integration by parts, its contribution is

$$-\nu \int_0^T \int_{\mathbb{R}} |u_\nu - k|\varphi_{yy} \,\mathrm{d}y\,\mathrm{d}t.$$

This term is negligible in the limit $\nu \downarrow 0$. Next, the integrand in (23) can be rewritten as

$$\varphi b[R_\epsilon(u_\nu^m - k^m)]_y \quad \text{with} \quad R_\epsilon(s) = \int_0^s r\, \mathrm{Sgn}'_\epsilon(r)\,\mathrm{d}r.$$

Since R_ϵ converges to zero uniformly for $\epsilon \downarrow 0$, it follows — after an integration by parts — that the integral (23) vanishes in that limit. Finally, the integral (24) obviously gives a non-negative

contribution. Thus in the limit $\epsilon \downarrow 0$, and after an integration by parts with respect to time in (21), the equality (21)–(24) implies that

$$
\int_0^T \int_{\mathbb{R}} |u_\nu - k| \varphi_t \, \mathrm{d}y \, \mathrm{d}t \geq \int_0^T \int_{\mathbb{R}} \mathrm{Sgn}(u_\nu - k) \Big(\big[(u_\nu^m)_y + b(u_\nu^m - k^m) \big] \varphi_y - b_y k^m \varphi \Big) \, \mathrm{d}y \, \mathrm{d}t
$$
$$
+ \limsup_{\epsilon \downarrow 0} \int_0^T \int_{\mathbb{R}} \mathrm{Sgn}'_\epsilon (u_\nu^m - k^m) \big[(u_\nu^m)_y \big]^2 \varphi \, \mathrm{d}y \, \mathrm{d}t. \tag{25}
$$

We take this as the defining inequality for entropy solutions.

Definition 2.4 (Definition of entropy solution). *Let $u^0 \in L^1 \cap L^\infty(\mathbb{R})$. A non-negative measurable function $u : \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}$ is an* entropy solution *to (1) with initial condition u^0 if $u \in L^1 \cap L^\infty(]0, T[\times \mathbb{R})$ for all $T > 0$, if $u^m \in L^2_{loc}(0, \infty; H^1(\mathbb{R}))$, if $u(t) \to u^0$ in $L^1(\mathbb{R})$ as $t \downarrow 0$, and if inequality (25) is satisfied for all nonnegative test functions $\varphi \in C_c^\infty(\mathbb{R}_+ \times \mathbb{R})$, and for all $k \in \mathbb{R}_+$.*

Remark 2.5. *The inclusion of the very particular dissipation term in inequality (25) seems a bit ad hoc. This term, however, plays a key role in the proof of uniqueness, see Section 5. Moreover, the choice of this term is less arbitrary than it appears: any smooth and uniformly convergent approximation of the sign function could be used in place of the mollification Sgn_ϵ there.*

Finally, we remark that a substantial part of the article [18] is devoted to proving that all functions u of a certain regularity which satisfy (25) without the dissipation term actually satisfy it also with the dissipation term. For us, the dissipation term results very naturally from our construction of solutions u.

Remark 2.6. *For a function u of the specified regularity, in order to be an entropy solution it is sufficient that inequality (25) is satisfied for all test function φ of the form $\varphi(t, y) = \theta(t)\phi(y)$ with arbitrary non-negative functions $\theta \in C_c^\infty(\mathbb{R}_+)$ and $\phi \in C_c^\infty(\mathbb{R})$. This follows immediately since the latter products lie dense in $C_c^\infty(\mathbb{R}_+ \times \mathbb{R})$, see e.g. [14, Theorem 4.3.1].*

In difference to [18], we do not require — and, in fact, cannot prove — that u is a continuous curve in $L^1(\mathbb{R})$ for $t > 0$.

2.3. **The JKO scheme.** The JKO scheme [16] is a variant of the time-discrete implicit Euler approximation for the solution of gradient flows in the non-smooth metric setting of the L^2-Wasserstein distance. In its core, it is a special case of De Giorgi's minimizing movement scheme [10]; see the book [1] for an extensive theory.

We apply the JKO scheme to the functional $\mathcal{F}$ defined in (4) with a defined implicitly in Proposition 2.1 and inital condition $\rho^0 := \mathbf{S}[u^0]$, see (11). To this end, let a time step $\tau > 0$ be given. For every $\sigma \in \mathcal{P}_2(\mathbb{R})$, introduce the associated Yoshida penalization $\mathcal{F}_\tau(\cdot; \sigma)$ of $\mathcal{F}$ by

$$
\mathcal{F}_\tau(\rho; \sigma) = \frac{1}{2\tau} \mathbf{W}_2(\rho, \sigma)^2 + \mathcal{F}(\rho).
$$

Let further an initial condition $\rho^0 \in \mathcal{P}_2(\mathbb{R})$ with $\mathcal{F}(\rho^0) < +\infty$ be given. We define a sequence of densities $\rho_\tau^n \in \mathcal{P}_2(\mathbb{R})$ inductively as follows:

(1) $\rho_\tau^0 := \rho^0$.
(2) For $n \geq 1$, let $\rho_\tau^n \in \mathcal{P}_2(\mathbb{R})$ be the (unique global) minimizer of $\mathcal{F}_\tau(\cdot; \rho_\tau^{n-1})$.

In Lemma 4.1 we prove well-definiteness of this scheme. In the following, we denote by $\bar{\rho}_\tau : [0, \infty[\to \mathcal{P}_2(\mathbb{R})$ the piecewise constant interpolation of the sequence $(\rho_\tau^n)_{n \in \mathbb{N}}$, with

$$
\bar{\rho}_\tau(t) = \rho_\tau^n \quad \text{for } (n-1)\tau < t \leq n\tau. \tag{26}
$$

2.4. **Main results.** We are now in the position to give the precise statement of the two main results of this paper.

Theorem 2.7. *Let $\rho^0 \in \mathcal{P}_2(\mathbb{R}) \cap L^\infty(\mathbb{R})$, and let a satisfy conditions (a1)&(a2) of Proposition 2.1. Define discrete curves $\bar{\rho}_\tau : [0, \infty[\to \mathcal{P}_2(\mathbb{R})$ by means of the JKO scheme from Section 2.3 above. Then every vanishing sequence $(\tau_k)_{k \in \mathbb{N}}$ of time steps contains a subsequence (not relabeled) such that the ρ_{τ_k} converge — in $L^m(]0, T[\times \mathbb{R})$, and also uniformly in $\mathbf{W}_2$ on each finite time interval $[0, T]$ — to a curve $\rho_* : [0, \infty[\to \mathcal{P}_2(\mathbb{R})$ that is continuous with respect to $\mathbf{W}_2$. The*

rescaled function $u_* = \mathbf{S}^{-1}[\rho_*]$ is an entropy solution to (1) in the sense of Definition 2.4, with initial condition $u^0 = \mathbf{S}^{-1}[\rho^0]$.

The proof of Theorem 2.7 is performed in Section 4.

Theorem 2.8. *For each initial condition $u^0 \in \mathcal{P}_2(\mathbb{R}) \cap L^\infty(\mathbb{R})$, there exists only one entropy solution to (1) in the sense of Definition 2.4. Consequently, every sequence of time-discrete approximations $\bar{\rho}_{\tau_k}$ converges to the same limit ρ_*.*

The proof of Theorem 2.8 is taken from [18], with minor modifications. In order to make the present paper self-contained, we outline the argument and review the relevant calculations in Section 5.

Remark 2.9. *A sensible notion of entropy solution (providing uniqueness under suitable hypotheses) can be introduced also directly for the equation (3), without passing through the scaling established in Proposition 2.1, see [17]. We prefer to deal with equation (1), which is a porous medium equation with a nonlinear convection perturbation, and is therefore of independent interest.*

3. Convex functionals and contractive gradient flows

Contractive gradient flows (or "κ-flows", see Definition 3.1 below) constitute our key tool for the derivation of a priori estimates on solutions ρ to (3), or, equivalently, for solutions u to (1). First, we recall the definition of κ-flows and their relation to geodesically λ-convex functionals. Then we prove κ-contractivity of a class of flows that is relevant for our needs. We refer to [1, 28, 29] as references on the general gradient flow theory on Wasserstein spaces.

3.1. **Fundamental definitions and relations.** A curve $\rho : I \to \mathcal{P}_2(\mathbb{R})$ is called *absolutely continuous* in $\mathbf{W}_2$ on the interval $I \subset \mathbb{R}$ if there exists a function $g \in L^1_{\mathrm{loc}}(I)$ such that

$$\mathbf{W}_2(\rho(t), \rho(s)) \leq \left| \int_s^t g(\tau) d\tau \right| \quad \text{for all } t, s \in I.$$

An absolutely continuous curve $\rho : [0,1] \to \mathcal{P}_2(\mathbb{R})$ is a *constant speed geodesic* if

$$\mathbf{W}_2(\rho(s), \rho(t)) = |t - s| \mathbf{W}_2(\rho(0), \rho(1)) \quad \text{for all } t, s \in [0,1].$$

Definition 3.1 (κ-flow). *A semigroup $\mathfrak{S}_\Psi : [0,\infty[\times \mathcal{P}_2(\mathbb{R}) \to \mathcal{P}_2(\mathbb{R})$ is a κ-flow for a functional $\Psi : \mathcal{P}_2(\mathbb{R}) \to \mathbb{R} \cup \{+\infty\}$ with respect to $\mathbf{W}_2$ if, for arbitrary $\rho \in \mathcal{P}_2(\mathbb{R})$, the curve $s \mapsto \mathfrak{S}_\Psi^s \rho$ is absolutely continuous on $[0,\infty[$ and satisfies the evolution variational inequality (EVI)*

$$\frac{1}{2} \frac{d^+}{d\sigma} \Big|_{\sigma=s} \mathbf{W}_2\big(\mathfrak{S}_\Psi^\sigma \rho, \tilde{\rho}\big)^2 + \frac{\kappa}{2} \mathbf{W}_2(\mathfrak{S}_\Psi^s \rho, \tilde{\rho})^2 \leq \Psi(\tilde{\rho}) - \Psi(\mathfrak{S}_s^\Psi \rho) \tag{27}$$

for all $s > 0$, with respect to every comparison measure $\tilde{\rho} \in \mathcal{P}_2(\mathbb{R})$ for which $\Psi(\tilde{\rho}) < \infty$.

Remark 3.2. *The symbol $d^+/d\sigma$ stands for the limit superior of the respective difference quotients, and equals to the derivative if the latter exists.*

The fact that a functional Ψ admits a κ-flow is essentially equivalent to the λ-convexity of Ψ along geodesics. The characterization of κ-flows by convexity will not play a role in our further considerations, but we cite the respective result for the sake of completeness, see [9, 25] for further details.

Theorem 3.3. *Assume that the functional $\Psi : \mathcal{P}_2(\mathbb{R}) \to \mathbb{R} \cup \{+\infty\}$ is λ-convex (along geodesics), with a modulus of convexity $\lambda \in \mathbb{R}$. That is, along every constant speed geodesic $\rho : [0,1] \to \mathcal{P}_2(\mathbb{R})$,*

$$\Psi[\rho(t)] \leq (1-t)\Psi[\rho(0)] + t\Psi[\rho(1)] - \frac{\lambda}{2} t(1-t) \mathbf{W}_2(\rho(0), \rho(1))^2 \tag{28}$$

holds for every $t \in [0,1]$. Then Ψ possesses a uniquely determined κ-flow, with some $\kappa \geq \lambda$. Conversely, if a functional Ψ possesses a κ-flow, and if Ψ is monotonically non-increasing along that flow, then Ψ is λ-convex, with some $\lambda \geq \kappa$.

3.2. **A special class of κ-flows.** In this section, we derive a sufficient condition under which the functional

$$\Psi(\eta) = \int_{\mathbb{R}} F(x, \eta(x)) \, \mathrm{d}x \tag{29}$$

for a sufficiently smooth function $F : \mathbb{R} \times \mathbb{R}_+ \to \mathbb{R}$ admits a κ-flow. For a concise formulation of that condition, we introduce the adjoint function $H : \mathbb{R} \times \mathbb{R}_+ \to \mathbb{R}$ by

$$H(x, \xi) = \xi F(x, 1/\xi), \tag{30}$$

which satisfies

$$H_x(x, \xi) = \xi F_x(x, 1/\xi), \quad H_\xi(x, \xi) = F(x, 1/\xi) - 1/\xi \, F_\eta(x, 1/\xi). \tag{31}$$

Lemma 3.4. *Let $F \in C^2(\mathbb{R} \times \mathbb{R}_+)$ be given, and assume that there exist constants $0 < c < C$ such that*

$$c \le \eta F_{\eta\eta}(x, \eta) \le C \quad and \quad |F_{x\eta}(x, \eta)| \le C \quad for \ all \ (x, \eta) \in \mathbb{R} \times \mathbb{R}_+. \tag{32}$$

Assume further that there is some $\kappa \in \mathbb{R}$ such that $H_\kappa : \mathbb{R} \times \mathbb{R}_+ \to \mathbb{R}$ with

$$H_\kappa(x, \xi) = H(x, \xi) - \frac{\kappa}{2} x^2 \tag{33}$$

is (jointly) convex on $\mathbb{R} \times \mathbb{R}_+$. Then, the solution operator $\mathfrak{S}_\Psi$ of the evolution equation

$$\eta_t = \mathrm{D}_x(\eta \mathrm{D}_x[F_\eta(x, \eta)]) \tag{34}$$

is a κ-flow for the functional Ψ from (29).

The convexity condition (33) can be rephrased as

$$\mathrm{D}^2 H_\kappa = \begin{pmatrix} H_{xx} - \kappa & H_{x\xi} \\ H_{x\xi} & H_{\xi\xi} \end{pmatrix} \quad \text{is positive semi-definite for all } (x, \xi), \tag{35}$$

and the condition (32) is equivalent to

$$c \le \xi^2 H_{\xi\xi}(x, \xi) \le C \quad and \quad |H_x(x, \xi) - \xi H_{x\xi}(x, \xi)| \le C \quad for \ all \ x \in \mathbb{R}, \ \xi \in \mathbb{R}_+. \tag{36}$$

Proof. We need to show that the EVI

$$\frac{1}{2} \frac{\mathrm{d}}{\mathrm{d}t} \mathbf{W}_2(\mathfrak{S}_\Psi^t \eta^0, \tilde{\eta})^2 + \frac{\kappa}{2} \mathbf{W}_2(\mathfrak{S}_\Psi^t \eta^0, \tilde{\eta})^2 \le \Psi(\tilde{\eta}) - \Psi(\mathfrak{S}_\Psi^t \eta^0) \tag{37}$$

holds for all smooth and strictly positive densities $\eta^0, \tilde{\eta} \in \mathcal{P}_2(\mathbb{R})$ at almost every time $t > 0$.

Instead of proving (37) directly for $\mathfrak{S}_\Psi$, we prove it for a family of regularizations and then pass to the limit. For every $N \in \mathbb{N}$, we set

$$\widetilde{F}^N(x, \eta) = \begin{cases} F(x, \eta) & \text{if } |x| \le N - 1 \\ F(N - 1, \eta) & \text{if } x > N - 1 \\ F(-N + 1, \eta) & \text{if } x < -N + 1 \end{cases},$$

and consider F^N to be a C^∞ mollification of $\widetilde{F}^N$ as $N \to +\infty$ (by a compactly supported smoothing kernel) such that $F_x \equiv 0$ in the set $\{|x| \ge N - 1/2\}$ for N large enough. Note that F^N can be chosen such that it satisfies the conditions (32) with constants c, C independent of N. Let also define the approximate adjoint function $H^N(x, \xi) = \xi(F^N(x, 1/\xi)$. Note that $F^N(x, \eta) \to F(x, \eta)$ for all $(x, \eta) \in \mathbb{R} \times [0, +\infty[$ as $N \to +\infty$. Let $\widetilde{\Psi}^N(\eta) = \int_{\mathbb{R}} F^N(x, \eta(x)) dx$. The regularized functional Ψ_N is defined by $\Psi_N(\eta) = \widetilde{\Psi}^N(\eta)$ if $\eta \in \mathcal{P}_2(\mathbb{R})$ is supported in $[-N, N]$, and $\Psi_N(\eta) = +\infty$ otherwise. We show that the solution operator $\mathfrak{S}_N$ to the boundary value problem

$$\eta_t = \mathrm{D}_x(\eta F_{\eta\eta}^N(x, \eta)\eta_x + \eta F_{\eta x}^N(x, \eta)), \quad \eta_x(t, N) = \eta_x(t, -N) = 0 \tag{38}$$

is a κ-flow for Ψ_N, i.e., the analogue of (37) holds. The densities $\tilde{\eta}$ and η^0 are approximated by

$$\tilde{\eta}_N = (\eta + \tilde{\mu}_N)\mathbf{1}_{[-N,N]}, \quad \eta_N^0 = (\eta^0 + \mu_N^0)\mathbf{1}_{[-N,N]},$$

with constants $\tilde{\mu}_N, \mu_N^0 > 0$ such that $\tilde{\eta}_N, \eta_N^0 \in \mathcal{P}_2(\mathbb{R})$. Notice that $\Psi_N(\tilde{\eta}_N), \Psi_N(\eta_N^0)$ converge to $\Psi(\tilde{\eta}), \Psi(\eta^0)$, respectively, as $N \to \infty$.

By the lower bound on $\eta F_{\eta\eta}$ required in (32) (and satisfied by F^N as noted before), the equation (38) is uniformly parabolic. Moreover, $F^N_{x\eta} \equiv 0$ on $x \in \{-N, N\}$, so that (38) preserves the total mass. Therefore, the boundary value problem with initial condition η^0_N possesses a solution η with $\eta(t) = \mathfrak{S}^t_N(\eta^0_N)$ for all $t \geq 0$ such that each $\eta(t) \in \mathcal{P}_2(\mathbb{R})$ restricts to a smooth and strictly positive function on $[-N, N]$ and vanishes outside of that interval. The associated distribution function $U : [0, \infty[\times[-N, N] \to [0, 1]$ with

$$U(t; x) = \int_{-\infty}^{x} \eta(t; y)\, \mathrm{d}y$$

is a smooth diffeomorphism from $[-N, N]$ to $[0, 1]$, for every $t \in \mathbb{R}_+$; this follows from smoothness, strict positivity and mass preservation of η. The inverse $G : [0, \infty[\times[0, 1] \to [-N, N]$ of U with respect to x satisfies by definition

$$G(t; U(t; x)) = x \quad \text{for all } x \in [-N, N]. \tag{39}$$

In particular, we have

$$G(t; 0) = -N \text{ and } G(t; 1) = N. \tag{40}$$

We differentiate equation (38) with respect to x and with respect to t, respectively, to obtain

$$1 = \eta(t; x)\, G_\omega(t; U(t; x)), \quad \text{and} \tag{41}$$
$$0 = G_t(t; U(t; x)) + U_t(t; x)\, G_\omega(t; U(t; x)). \tag{42}$$

On the other hand, integration of (34) with respect to x yields

$$U_t(t; x) = \eta(t; x)\mathrm{D}_x\big[F^N_\eta\big(x, \eta(t; x)\big)\big].$$

We multiply by $G_\omega(t; U(t; x))$, substitute (42) for U_t and (41) for η. This yields

$$G_t(t; U(t; x)) = -\mathrm{D}_x\big[F^N_\eta\big(x, \eta(t; x)\big)\big]. \tag{43}$$

We rewrite this (omitting the dependence on t) using again (41) and the properties of H from (31):

$$-\mathrm{D}_x[F^N_\eta(x, \eta(x))] = \frac{1}{\eta(x)}\big(\mathrm{D}_x[F^N(x, \eta(x)) - \eta(x)F^N_\eta(x, \eta(x))]$$
$$- \mathrm{D}_x F^N(x, \eta(x)) + \eta_x(x)F^N_\eta(x, \eta(x))\big)$$
$$= \frac{1}{\eta(x)}\mathrm{D}_x[H^N_\xi(x, 1/\eta(x))] - H^N_x(x, 1/\eta(x))$$
$$= G_\omega(U(x))\,\mathrm{D}_x\big[H^N_\xi\big(G(U(x)), G_\omega(U(x))\big)\big] - H^N_x\big(G(U(x)), G_\omega(U(x))\big).$$

In combination with (43), we have

$$G_t = \mathrm{D}_\omega[H^N_\xi(G, G_\omega)] - H^N_x(G, G_\omega). \tag{44}$$

Further, observe that a change of variables $\omega = U(x)$ leads to

$$\Psi_N(\eta) = \int_{-N}^{N} F^N(x, \eta(x))\, \mathrm{d}x = \int_{-N}^{N} F^N\Big(G(U(x)), \frac{1}{G_\omega(U(x))}\Big)\, \mathrm{d}x$$
$$= \int_0^1 G_\omega(\omega)F^N\Big(G(\omega), \frac{1}{G_\omega(\omega)}\Big)\, \mathrm{d}\omega = \int_0^1 H^N(G(\omega), G_\omega(\omega))\, \mathrm{d}\omega. \tag{45}$$

Now, let $\tilde{G}$ be the inverse distribution function of $\tilde{\eta}_N$. Using the representation (6) of the Wasserstein distance, we obtain

$$\mathbf{W}_2\big(\eta(t), \tilde{\eta}\big)^2 = \int_0^1 [G(t; \omega) - \tilde{G}(\omega)]^2\, \mathrm{d}\omega. \tag{46}$$

Now we can prove (37). Combining (46) with the evolution equation (44) and the representation (45) of Ψ, we find

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{W}_2(\eta(t),\tilde{\eta}_N)^2 + \frac{\kappa}{2}\mathbf{W}_2(\eta(t),\tilde{\eta}_N)^2 = \int_0^1 G_t[G-\tilde{G}]\,\mathrm{d}\omega + \frac{\kappa}{2}\int_0^1 [G-\tilde{G}]^2\,\mathrm{d}\omega$$

$$= \int_0^1 \mathrm{D}_\omega[H_\xi^N(G,G_\omega)][G-\tilde{G}]\,\mathrm{d}\omega - \int_0^1 H_x^N(G,G_\omega)[G-\tilde{G}]\,\mathrm{d}\omega + \frac{\kappa}{2}\int_0^1 [G-\tilde{G}]^2\,\mathrm{d}\omega$$

$$\stackrel{(\star)}{=} \int_0^1 H_\xi^N(G,G_\omega)[\tilde{G}_\omega - G_\omega]\,\mathrm{d}\omega + \int_0^1 H_x^N(G,G_\omega)[\tilde{G}-G]\,\mathrm{d}\omega + \frac{\kappa}{2}\int_0^1 [\tilde{G}-G]^2\,\mathrm{d}\omega$$

$$\stackrel{(\star\star)}{\leq} \int_0^1 \left[H^N(\tilde{G},\tilde{G}_\omega) - H^N(G,G_\omega) \right]\,\mathrm{d}\omega = \Psi_N(\tilde{\eta}_N) - \Psi_N(\eta(t)).$$

In step $(\star\star)$ we have used the joint convexity of (33). The equality $(\star)$ is established integrating by parts; the boundary terms vanish because of the boundary conditions (40) satisfied by G and $\tilde{G}$. This proves the analogue of (37) for each Ψ_N with its respective flow $\mathfrak{S}_N$.

Finally, we pass to the limit $N \to \infty$. For definiteness of notation, we denote the solutions to (38) by $\eta_N(t) = \mathfrak{S}_N^t(\eta_N^0)$ from now on.

To begin with, we show that a subsequence $(\eta_{N'}) \subseteq (\eta_N)$ converges to a limit function η_∞ satisfying the unregularized evolution equation (34). For the following calculations, the bounds (32) (satisfied by F^N as well, uniformly w.r.t. N) are essential. First, we derive an $H^1(\mathbb{R})$-estimate:

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\int_{-N}^N \eta_N^2\,\mathrm{d}x = -\int_{-N}^N \eta_N F_{\eta\eta}^N(\eta_N,x)\,(\eta_N)_x^2\,\mathrm{d}x - \int_{-N}^N (\eta_N)_x\,F_{x\eta}^N(\eta_N,x)\eta_N\,\mathrm{d}x$$

$$\leq -c\int_{-N}^N (\eta_N)_x^2\,\mathrm{d}x + \frac{c}{2}\int_{-N}^N (\eta_N)_x^2\,\mathrm{d}x + \frac{C^2}{2c}\int_{-N}^N \eta_N^2\,\mathrm{d}x.$$

Integration with respect to $t \in]0,T[$ yields an N-uniform bound on η_N in $L^2(0,T;H^1(\mathbb{R}))$ in terms of $\|\eta^0\|_{L^2}$. By Alaoglu's theorem, this provides weak convergence of a subsequence $\eta_{N'}$ to a limit η_∞ in $L^2(0,\infty;H^1(\mathbb{R}))$. Second, we prove an $H^{-1}(\mathbb{R})$-estimate on $(\eta_N)_t$. To this end, let $\theta_{xx} = (\eta_N)_t$ with $\theta_x(t,N) = \theta_x(t,-N) = 0$, and calculate, using again (32):

$$\int_{-N}^N \theta_x^2\,\mathrm{d}x = -\int_{-N}^N \theta\theta_{xx}\,\mathrm{d}x = -\int_{-N}^N \theta(\eta_N)_t\,\mathrm{d}x = \int_{-N}^N \theta_x\,(\eta_N)_x\,\eta_N F_{\eta\eta}^N\,\mathrm{d}x + \int_{-N}^N \theta_x\,\eta_N F_{\eta x}^N\,\mathrm{d}x$$

$$\leq \int_{-N}^N \theta_x^2\,\mathrm{d}x + \frac{C^2}{2}\int_{-N}^N (\eta_N)_x^2\,\mathrm{d}x + \frac{C^2}{2}\int_{-N}^N \eta_N^2\,\mathrm{d}x,$$

which, in combination with the $H^1(\mathbb{R})$-bound from above, shows that $(\eta_N)_t$ is N-uniformly bounded in $L^2(0,T;H^{-1}(\mathbb{R}))$. The Aubin-Lions lemma, see e.g. [27], shows that the convergence of $\eta_{N'}$ to η_∞ is actually strong in $L^2_{\mathrm{loc}}(\mathbb{R}_+ \times \mathbb{R})$.

Strong convergence in $L^2_{\mathrm{loc}}(\mathbb{R}_+ \times \mathbb{R})$ and weak convergence in $L^2(0,\infty;H^1(\mathbb{R}))$ allows us to pass to the limit in the weak formulation of (38),

$$-\int_0^\infty \int_{\mathbb{R}} \varphi_t \eta_{N'}\,\mathrm{d}x\,\mathrm{d}t = -\int_0^\infty \int_{\mathbb{R}} \varphi_x\left[\eta_{N'} F_{\eta\eta}^{N'}(x,\eta_{N'})(\eta_{N'})_x + \eta_{N'} F_{x\eta}^{N'}(x,\eta_{N'}) \right]\,\mathrm{d}x\,\mathrm{d}t$$

where $\varphi \in C_c^\infty(\mathbb{R}_+ \times \mathbb{R})$ is a test function. Here we use that the nonlinearities $\eta F_{\eta\eta}(x,\eta)$ and $\eta F_{x\eta}(x,\eta)$ grow at most linearly with respect to η by assumption (32), and so they converge in $L^2_{\mathrm{loc}}(\mathbb{R}_+ \times \mathbb{R})$, too, by dominated convergence. Since $F^{N'}$ converges to F on the compact support of φ, we conclude that the limit η_∞ is a weak (and thus — by parabolic regularity theory — also the unique classical) solution to the Cauchy problem (34) with initial condition η^0. In other words, we have $\mathfrak{S}_\Psi^t(\eta_0) = \eta_\infty(t)$ for almost every $t \geq 0$, where $\mathfrak{S}_\Psi$ is the solution operator to (34).

Passage to the limit in the EVI (37) is more subtle because the convergence of the time derivatives $(\eta_{N'})_t$ is a priori only weak in $L^2(0,\infty;H^{-1}(\mathbb{R}))$. In [9], an equivalent time-integrated version of (37) is given, and the latter is stable under strong L^2_{loc}-convergence. Notice that $\tilde{\eta}_N$ can be chosen such that $\Psi_N(\tilde{\eta}_N) \to \Psi(\tilde{\eta})$ (e. g. by suitably cutting off $\tilde{\eta}$ on $|x| \leq N-1$). Moreover, the lower semi-continuity of the term $\Psi_N(\eta_N(t))$ can be proven by first restricting to compact subsets

$K \subset \mathbb{R}$ so that $F^N \equiv F$ on K for N large enough. We refer the interested reader to [9, Section 3] for further details. $\square$

Lemma 3.4 gives an indication that the functional $\mathcal{F}$ from (4) is *not* geodesically λ-convex. Note that, by Theorem 3.3, convexity of $\mathcal{F}$ would imply the existence of a κ-flow for Ψ. However, with $F(x, \eta) = a(x)\eta^m$, we have $H(x, \xi) = a(x)\xi^{1-m}$, and thus, with H_κ defined in (33),

$$\mathrm{D}^2 H_\kappa(x, \xi) = \begin{pmatrix} a_{xx}(x)\xi^{1-m} - \kappa & -(m-1)a_x(x)\xi^{-m} \\ -(m-1)a_x(x)\xi^{-m} & m(m-1)a(x)\xi^{-(m+1)} \end{pmatrix}.$$

Unless a is a constant, there must exist an $\bar{x} \in \mathbb{R}$ with $a_{xx}(\bar{x}) < 0$. Thus, no matter how κ is chosen, there exists further a sufficiently small $\bar{\xi} \in \mathbb{R}_+$ such that the top left element of $\mathrm{D}^2 H_\kappa(\bar{x}, \bar{\xi})$ is negative, that is $a_{xx}(\bar{x})\bar{\xi}^{-(m-1)} - \kappa < 0$. Consequently, $\mathrm{D}^2 H_\kappa(\bar{x}, \bar{\xi})$ is not positive semi-definite. Note, however, that we have only shown sufficiency of the joint convexity of (33) for the existence of a κ-flow.

4. Construction of entropy solutions

This section is devoted to the proof of Theorem 2.7.

Throughout this section, ρ_τ^n denotes the nth iterate in the JKO scheme for given time step $\tau > 0$ as described in Section 2.3, see also (47) below. Further, $\bar{\rho}_\tau : [0, \infty[\to \mathcal{P}_2(\mathbb{R})$ denotes the piecewise constant interpolation of the sequence $(\rho_\tau^n)_{n \in \mathbb{N}}$, with (26).

4.1. **Well posedness of the JKO scheme.** Recall the inductive definition of the ρ_τ^n,

$$\rho_\tau^n \in \mathcal{P}_2(\mathbb{R}) \quad \text{is the unique minimizer of} \quad \mathcal{F}_\tau(\rho; \rho_\tau^{n-1}) = \frac{1}{2\tau} \mathbf{W}_2(\rho, \rho_\tau^{n-1})^2 + \mathcal{F}(\rho). \qquad (47)$$

Lemma 4.1. *The time-discrete scheme is well-defined in the sense that for any given $\rho_\tau^{n-1} \in \mathcal{P}_2(\mathbb{R})$, the functional $\mathcal{F}_\tau(\cdot; \rho_\tau^{n-1})$ indeed admits a unique minimizer $\rho_\tau^n \in \mathcal{P}_2(\mathbb{R})$.*

Proof. We argue that the direct methods from the calculus of variations apply to the minimization problem for $\mathcal{F}_\tau(\cdot; \rho_\tau^{n-1})$ on $\mathcal{P}_2(\mathbb{R})$.

First, observe that $\rho_\tau^{n-1} \in \mathcal{P}_2(\mathbb{R})$ implies that the $\mathcal{F}(\rho_\tau^{n-1})$-sublevel of this functional is non-empty and weakly-$\star$ relatively compact. Indeed, ρ_τ^{n-1} itself belongs to that sublevel, and if $\rho^* \in \overline{\mathcal{P}_2(\mathbb{R})}$ is any other element of this sublevel, then $\mathbf{W}_2(\rho^*, \rho_\tau^{n-1})^2 \leq \mathcal{F}(\rho_\tau^{n-1})$, since $\mathcal{F}$ is a non-negative functional. This bound on the Wasserstein distance provides a control on the second moment of ρ^* in terms of ρ_τ^{n-1} alone, and so the sublevel is a tight collection of measures, and thus relatively compact in the weak-$\star$ topology. Further, both $\mathcal{F}$ and the distance $\mathbf{W}_2(\cdot, \rho_\tau^{n-1})$ are lower semi-continuous with respect to narrow convergence. This, together with the fact that $\mathcal{F}_\tau(\cdot; \rho_\tau^{n-1})$ is bounded from below, guarantees the existence of at least one minimizer $\rho_\tau^n \in \overline{\mathcal{P}_2(\mathbb{R})}$. Finally, since necessarily $\mathcal{F}(\rho_\tau^n)$ is finite, ρ_τ^n belongs to the subspace of densities $\mathcal{P}_2(\mathbb{R})$.

To prove uniqueness of the minimizer, it suffices to observe that the squared distance $\mathbf{W}_2(\cdot, \rho_\tau^{n-1})^2$ is convex, and that the potential $\mathcal{F}$ is *strictly* convex (recall that $m > 1$) in the sense of linear interpolation of measures. Thus, there can be at most one minimizer. $\square$

4.2. **Basic energy estimates and flow interchange.** Immediately from the construction in (47), we obtain the canonical energy estimate

$$\frac{1}{2\tau} \sum_{n=1}^{N} \mathbf{W}_2(\rho_\tau^n, \rho_\tau^{n-1})^2 \leq \mathcal{F}(\rho_\tau^0) - \mathcal{F}(\rho_\tau^N) \quad \text{for all } N \in \mathbb{N}.$$

which induces the following uniform bound on ρ_τ^n in $L^m(\mathbb{R})$:

$$\frac{a}{m} \|\rho_\tau^n\|_{L^m}^m \leq \mathcal{F}(\rho_\tau^n) \leq \mathcal{F}(\rho^0). \qquad (48)$$

Since $\mathcal{F}$ is non-negative, it follows further that

$$\sum_{n=1}^{\infty} \mathbf{W}_2(\rho_\tau^n, \rho_\tau^{n-1})^2 \leq 2\tau \mathcal{F}(\rho_\tau^0). \qquad (49)$$

For our purposes, we derive stronger a priori estimates by variations of the minimizers ρ_τ^n along specific κ-flows, following the general strategy from [22, Section 3].

Lemma 4.2. *Let* $\Psi : \mathcal{P}_2(\mathbb{R}) \to]-\infty, +\infty]$ *be a lower semi-continuous functional on* $\mathcal{P}_2(\mathbb{R})$ *which possesses a* κ-*flow* $\mathfrak{S}_\Psi$. *Define further the dissipation of* $\mathcal{F}$ *along* $\mathfrak{S}_\Psi$ *by*

$$\mathfrak{D}_\Psi(\rho) := \limsup_{s\downarrow 0} \frac{1}{s} \big[\mathcal{F}(\rho) - \mathcal{F}(\mathfrak{S}_\Psi^s \rho) \big]$$

for every $\rho \in \mathcal{P}_2(\mathbb{R})$. *If* ρ_τ^{n-1} *and* ρ_τ^n *are two consecutive steps of the minimizing movement scheme* (47), *then*

$$\Psi(\rho_\tau^{n-1}) - \Psi(\rho_\tau^n) \geq \tau \mathfrak{D}_\Psi(\rho_\tau^n) + \frac{\kappa}{2} \mathbf{W}_2(\rho_\tau^n, \rho_\tau^{n-1})^2. \tag{50}$$

In particular, $\Psi(\rho_\tau^{n-1}) < \infty$ *implies* $\mathfrak{D}_\Psi(\rho_\tau^n) < \infty$.

Proof. Since (50) is trivial for $\Psi(\rho_\tau^{n-1}) = +\infty$, there is no loss in generality to assume $\Psi(\rho_\tau^{n-1}) < \infty$. We can thus use the EVI inequality (27) with $\rho := \rho_\tau^n$ for the comparison measure $\tilde{\rho} := \rho_\tau^{n-1}$. By lower semi-continuity of Ψ, we conclude

$$\Psi(\rho_\tau^{n-1}) - \Psi(\rho_\tau^n) \geq \limsup_{s\downarrow 0} \big(\Psi(\rho_\tau^{n-1}) - \Psi(\mathfrak{S}_\Psi^s \rho_\tau^n) \big) \tag{51}$$

$$\geq \frac{1}{2} \limsup_{s\downarrow 0} \Big(\frac{\mathrm{d}^+}{\mathrm{d}\sigma} \Big|_{\sigma=s} \mathbf{W}_2 \big(\mathfrak{S}_\Psi^\sigma \rho_\tau^n, \rho_\tau^{n-1} \big)^2 \Big) + \frac{\kappa}{2} \mathbf{W}_2(\rho_\tau^n, \rho_\tau^{n-1})^2. \tag{52}$$

Notice that we have used the (absolute) $\mathbf{W}_2$-continuity of the curve $s \mapsto \mathfrak{S}_\Psi^s \rho_\tau^n$ at $s = 0$ in the first step. The absolute continuity implies further that

$$\limsup_{s\downarrow 0} \Big(\frac{\mathrm{d}^+}{\mathrm{d}\sigma} \Big|_{\sigma=s} \mathbf{W}_2 \big(\mathfrak{S}_\Psi^\sigma \rho_\tau^n, \rho_\tau^{n-1} \big)^2 \Big) \geq \limsup_{s\downarrow 0} \frac{1}{s} \big(\mathbf{W}_2(\mathfrak{S}_\Psi^s \rho_\tau^n, \rho_\tau^{n-1})^2 - \mathbf{W}_2(\rho_\tau^n, \rho_\tau^{n-1})^2 \big). \tag{53}$$

By definition of ρ_τ^n as a minimizer of $\mathcal{F}_\tau$ from (47), we have

$$\mathcal{F}_\tau(\mathfrak{S}_\Psi^s \rho_\tau^n; \rho_\tau^{n-1}) \geq \mathcal{F}_\tau(\rho_\tau^n; \rho_\tau^{n-1})$$

for every $s \geq 0$, and thus

$$\mathbf{W}_2(\mathfrak{S}_\Psi^s \rho_\tau^n, \rho_\tau^{n-1})^2 - \mathbf{W}_2(\rho_\tau^n, \rho_\tau^{n-1})^2 \geq 2\tau \big[\mathcal{F}(\rho_\tau^n) - \mathcal{F}(\mathfrak{S}_\Psi^s \rho_\tau^n) \big]. \tag{54}$$

Inserting (54) into (53), and then (53) into (52) yields (50). □

Corollary 4.3. *Under the hypotheses of Lemma 4.2, let the* κ-*flow* $\mathfrak{S}_\Psi$ *be such that for every* $n \in \mathbb{N}$, *the curve* $s \mapsto \mathfrak{S}_\Psi^s \rho_\tau^n$ *lies in* $L^m(\mathbb{R})$, *where it is differentiable for* $s > 0$ *and continuous at* $s = 0$. *And moreover, let a functional* $\mathfrak{K} : \mathcal{P}_2(\mathbb{R}) \to]-\infty, \infty]$ *satisfy*

$$\liminf_{s\downarrow 0} \Big(-\frac{\mathrm{d}}{\mathrm{d}\sigma} \Big|_{\sigma=s} \mathcal{F}(\mathfrak{S}_\Psi^\sigma \rho_\tau^n) \Big) \geq \mathfrak{K}(\rho_\tau^n). \tag{55}$$

Then the following two estimates hold.

$$\text{For every } n \in \mathbb{N}: \quad \Psi(\rho_\tau^{n-1}) - \Psi(\rho_\tau^n) \geq \tau \mathfrak{K}(\rho_\tau^n) + \frac{\kappa}{2} \mathbf{W}_2(\rho_\tau^n, \rho_\tau^{n-1})^2; \tag{56}$$

$$\text{for every } N \in \mathbb{N}: \quad \Psi(\rho_\tau^N) \leq \Psi(\rho^0) - \tau \sum_{n=1}^{N} \mathfrak{K}(\rho_\tau^n) + \tau \max(0, -\kappa) \mathcal{F}(\rho^0). \tag{57}$$

Proof. First, we estimate $\mathfrak{D}_\Psi$ from below by $\mathfrak{K}$. The $L^m(\mathbb{R})$-regularity assumptions on $\mathfrak{S}^\Psi$ imply that $s \mapsto \mathcal{F}(\mathfrak{S}_s^\Psi \rho_\tau^n)$ is differentiable for $s > 0$ and continuous at $s = 0$. Thus, by the fundamental theorem of calculus,

$$\mathfrak{D}^\Psi(\rho_\tau^n) = \limsup_{\bar{s}\downarrow 0} \frac{1}{\bar{s}} \big[\mathcal{F}(\rho_\tau^n) - \mathcal{F}(\mathfrak{S}_{\bar{s}}^\Psi \rho_\tau^n) \big] = \limsup_{\bar{s}\downarrow 0} \int_0^1 \Big(-\frac{\mathrm{d}}{\mathrm{d}\sigma} \Big|_{\sigma=\bar{s}z} \mathcal{F}(\mathfrak{S}_\sigma^\Psi \rho_\tau^n) \Big) \, \mathrm{d}z$$

$$\geq \int_0^1 \liminf_{\bar{s}\downarrow 0} \Big(-\frac{\mathrm{d}}{\mathrm{d}\sigma} \Big|_{\sigma=\bar{s}z} \mathcal{F}(\mathfrak{S}_\sigma^\Psi \rho_\tau^n) \Big) \, \mathrm{d}z \geq \mathfrak{K}(\rho_\tau^n),$$

by Fatou's lemma and (55). Now estimate (56) follows directly from (50). Estimate (57) is obtained by arranging (56) in a telescopic sum and combining it with the energy estimate (49). □

4.3. **Refined a priori estimates.** The main ingredient in order to get suitable compactness is the following a priori estimate.

Lemma 4.4. *There is a constant A depending only on ρ^0 (and in particular not on τ) such that the piecewise constant interpolants $\bar{\rho}_\tau$ satisfy*

$$\left\|\bar{\rho}_\tau^{m/2}\right\|_{L^2(0,T;H^1(\mathbb{R}))} \le A(1+T) \quad \text{for all } T > 0. \tag{58}$$

In particular, $\bar{\rho}_\tau(t)^{m/2} \in H^1(\mathbb{R})$ for every $t > 0$.

For the proof of Lemma 4.4, we need a τ-uniform control on the second moment of $\bar{\rho}_\tau$, which is established in a further lemma.

Lemma 4.5. *There is a constant B depending only on ρ^0 such that*

$$\int_\mathbb{R} |x|^2 \bar{\rho}_\tau(T,x)\,dx \le B(1+T) \quad \text{for all } T > 0. \tag{59}$$

The proof of Lemma 4.5 is obtained by a straight-forward application of the "flow interchange" Lemma 4.2. Specifically, we choose as auxiliary functional $\Psi := \mathbf{m}$ the second moment,

$$\mathbf{m}(\rho) = \int_\mathbb{R} x^2 \rho(x)\,dx.$$

This functional possesses a κ-flow (with $\kappa = 2$), which is explicitly given by

$$\left(\mathfrak{S}_\mathbf{m}^s \eta\right)(x) = e^{2s}\eta(e^{2s}x).$$

Clearly, $s \mapsto \mathfrak{S}_\mathbf{m}^s \eta$ defines a curve in $L^m(\mathbb{R})$ for every $\eta \in L^m(\mathbb{R})$; this curve is differentiable for every $s \ge 0$.

Proof of Lemma 4.5. In order to obtain a sensible bound $\mathfrak{K}_\mathbf{m}$ in (55), we compute the s-derivative of $\mathcal{F}$ along $\mathfrak{S}_\mathbf{m}^s$ for a given $\eta_0 \in \mathcal{P}_2(\mathbb{R}) \cap L^m(\mathbb{R})$:

$$\frac{d}{ds}\Big|_{s=0} \mathcal{F}\left(\mathfrak{S}_\Psi^s \eta_0\right) = \frac{d}{ds}\Big|_{s=0} \left(e^{2ms}\int_\mathbb{R} \frac{a(x)}{m}\eta_0(e^{2s}x)^m\,dx\right) = \frac{d}{ds}\Big|_{s=0}\left(e^{2(m-1)s}\int_\mathbb{R} \frac{a(e^{-2s}z)}{m}\eta_0(z)^m\,dz\right)$$

$$= 2(m-1)\int_\mathbb{R} \frac{a(z)}{m}\eta_0(z)^m\,dz - 2\int_\mathbb{R} \frac{za_x(z)}{m}\eta_0(z)^m\,dz \le M\mathcal{F}(\eta_0)$$

with the constant

$$M := 2(m-1) - 2\inf_{z\in\mathbb{R}} \frac{za_x(z)}{a(z)} \ge 0.$$

We apply Corollary 4.3 with $\mathfrak{K}_\mathbf{m} := -M\mathcal{F}$; estimate (57) becomes

$$\int_\mathbb{R} x^2 \rho_\tau^n(x)\,dx \le \int_\mathbb{R} x^2 \rho^0(x)\,dx + \tau\sum_{n=1}^N \mathfrak{K}_\mathbf{m}(\rho_\tau^n) \le \int_\mathbb{R} x^2 \rho^0(x)\,dx + M\mathcal{F}(\rho^0)\,N\tau,$$

using the monotonicity (48) of the potential. From here, (59) follows immediately. □

We now turn to the proof of Lemma 4.4 above. Also here, the key element is an application of Lemma 4.2. This time, the auxiliary functional $\Psi := \mathbf{H}$ is the *entropy*,

$$\mathbf{H}(\rho) = \int_\mathbb{R} \rho(x)\log\rho(x)\,dx.$$

One of the celebrated results from the theory of optimal transportation is that this functional possesses a κ-flow $\mathfrak{S}_\mathbf{H}$, with $\kappa = 0$, which is given by the heat semi-group, see e.g. [9, 16, 28]. More precisely: for given $\eta_0 \in \mathcal{P}_2(\mathbb{R})$, the curve $s \mapsto \eta(s) := \mathfrak{S}_\mathbf{H}^s \eta_0$ solves the initial value problem

$$\eta_s = \eta_{xx}, \quad \eta(0) = \eta_0, \tag{60}$$

in the classical sense: $\eta(s)$ is a positive density function for every $s > 0$, it is continuously differentiable as a map from $\mathbb{R}_+$ to $C^\infty(\mathbb{R}) \cap L^1(\mathbb{R})$, and if $\eta_0 \in L^m(\mathbb{R})$, then $\eta(s)$ converges to η_0 in $L^m(\mathbb{R})$ as $s \downarrow 0$.

In order to extract information from the flow interchange estimate for $\mathbf{H}$, we need another technical ingredient.

Lemma 4.6. *There is a constant C depending only on ρ^0 such that*

$$-C(1+T) \le \mathbf{H}(\bar\rho_\tau(t)) \le C \quad \text{for all } T > 0. \tag{61}$$

Proof of Lemma 4.6. One verifies by elementary calculations that

$$-\frac{2}{e}s^{1/2} < s\log s < \frac{1}{(m-1)e}s^m \quad \text{for all } s > 0.$$

And so, for every $\eta \in \mathcal{P}_2(\mathbb{R})$, we have on one hand that

$$\mathbf{H}(\eta) = \int_{\mathbb{R}} \eta(x)\log\eta(x)\,\mathrm{d}x \le \frac{1}{(m-1)e}\int_{\mathbb{R}} \eta(x)^m\,\mathrm{d}x \le \frac{m}{\underline{a}(m-1)e}\mathcal{F}(\eta),$$

while on the other hand,

$$\mathbf{H}(\eta) \ge -\frac{2}{e}\int_{\mathbb{R}}(1+x^2)^{-1/2}(1+x^2)^{1/2}\eta(x)^{-1/2}\,\mathrm{d}x$$

$$\ge -\frac{2}{e}\left(\int_{\mathbb{R}}\frac{\mathrm{d}x}{1+x^2}\right)^{1/2}\left(\int_{\mathbb{R}}(1+x^2)\eta(x)\,\mathrm{d}x\right)^{1/2} \ge -\frac{2\sqrt{\pi}}{e}\left(1+\int_{\mathbb{R}}x^2\eta(x)\,\mathrm{d}x\right)^{1/2}.$$

Now, with $\bar\rho_\tau(t)$ in place of η, it is straightforward to conclude both estimates in (61) by using the energy bound (48) and the moment control (59), respectively. $\qquad\square$

Proof of Lemma 4.4. By the preceding discussion, $\mathfrak{S}_{\mathbf{H}}$ satisfies the hypotheses of Corollary 4.3; we shall define a suitable lower bound $\mathfrak{K}_{\mathbf{H}}$ for the use in (55). By the spatial regularity of $\eta(s)$ for every $s > 0$, the following calculations are justified:

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}s}\mathcal{F}(\mathfrak{S}_{\mathbf{H}}^s\eta_0) &= \frac{1}{m}\int_{\mathbb{R}}a(x)\big(\eta(s,x)^m\big)_s\,\mathrm{d}x = \int_{\mathbb{R}}a(x)\eta(s,x)^{m-1}\eta_{xx}(s,x)\,\mathrm{d}x \\
&= -(m-1)\int_{\mathbb{R}}a(x)\eta(s,x)^{m-2}\eta_x(s,x)^2\,\mathrm{d}x - \int_{\mathbb{R}}a_x(x)\eta(s,x)^{m-1}\eta_x(s,x)\,\mathrm{d}x \\
&= -\frac{4(m-1)}{m^2}\int_{\mathbb{R}}a(x)\big[\big(\eta(s,x)^{m/2}\big)_x\big]^2\,\mathrm{d}x + \frac{1}{m}\int_{\mathbb{R}}a_{xx}(x)\eta(s,x)^m\,\mathrm{d}x \\
&\le -\frac{4(m-1)\underline{a}}{m^2}\int_{\mathbb{R}}\big[\big(\eta(s,x)^{m/2}\big)_x\big]^2\,\mathrm{d}x + \frac{\overline{a_{xx}}}{m}\int_{\mathbb{R}}\eta(s,x)^m\,\mathrm{d}x,
\end{aligned} \tag{62}$$

where we have used the notation $\overline{a_{xx}} := \sup a_{xx} < +\infty$ in view of the property (a2) in Proposition 2.1. Consequently, we define

$$\mathfrak{K}_{\mathbf{H}}(\rho) := \frac{4(m-1)\underline{a}}{m^2}\int_{\mathbb{R}}\big[\big(\rho^{m/2}\big)_x\big]^2\,\mathrm{d}x - \frac{\overline{a_{xx}}}{m}\int_{\mathbb{R}}\rho^m\,\mathrm{d}x. \tag{63}$$

It is easily seen that (55) is satisfied with this choice of $\mathfrak{K}$: indeed, since $\mathfrak{S}_{\mathbf{H}}^s$ is continuous in $L^m(\mathbb{R})$ at $s = 0$, it suffices to observe that $\mathfrak{K}_{\mathbf{H}}$ is lower semi-continuous in $L^m(\mathbb{R})$ by Lemma A.1. So Corollary 4.3 is applicable, and from (57), we conclude that

$$\mathbf{H}(\rho_\tau^N) + \frac{4(m-1)\underline{a}}{m^2}\tau\sum_{n=1}^N\big\|\big((\rho_\tau^n)^{m/2}\big)_x\big\|_{L^2}^2 \le \mathbf{H}(\rho^0) + \frac{\overline{a_{xx}}}{m}\tau\sum_{n=1}^N\big\|\rho_\tau^n\big\|_{L^m}^m \tag{64}$$

for every $N \in \mathbb{N}$. In combination with the bound (61) on $\mathbf{H}$, it follows that

$$\frac{4(m-1)\underline{a}}{m^2}\tau\sum_{n=1}^N\big\|(\rho_\tau^n)^{m/2}\big\|_{H^1}^2 \le \mathbf{H}(\rho^0) - \mathbf{H}(\rho_\tau^N) + \frac{\overline{a_{xx}}}{m}\tau\sum_{n=1}^N\big\|\rho_\tau^n\big\|_{L^m}^m + \frac{4(m-1)\underline{a}}{m^2}\tau\sum_{n=1}^N\big\|\rho_\tau^n\big\|_{L^m}^m$$

$$\le C(2 + N\tau) + \left[\frac{\overline{a_{xx}}}{\underline{a}} + \frac{4(m-1)}{m}\right]\mathcal{F}(\rho^0)N\tau.$$

The energy estimate (48) has been used to derive the last line. From here, it is immediate to conclude (58). □

4.4. **Weak and strong convergence of the discrete curves.** Our first convergence result is a standard consequence of the energy estimate.

Lemma 4.7. *Every vanishing sequence $(\tau_k)_{k\in\mathbb{N}}$ of time steps $\tau_k > 0$ contains a (non-relabelled) subsequence such that $\bar{\rho}_{\tau_k}$ converges — uniformly on compact time intervals — in $\mathbf{W}_2$ to a Hölder continuous limit curve $\rho_* : [0, \infty[\to \mathcal{P}_2(\mathbb{R})$.*

Proof. From (49), we conclude the Hölder-type estimate

$$\mathbf{W}_2\big(\bar{\rho}_\tau(t), \bar{\rho}_\tau(s)\big) \leq \sqrt{2\mathcal{F}(\rho_\tau^0)} \max(\tau, |t-s|)^{1/2} \tag{65}$$

for arbitrary $s, t \geq 0$. The claim is now obtained as a consequence of (a refined version of) the Ascoli-Arzelá theorem [1, Proposition 3.3.1]. □

In the following, let a sequence (τ_k) be fixed along which $\bar{\rho}_{\tau_k}$ converges to some limit ρ_* in the sense described in Lemma 4.7. For the derivation of the entropy formulation (25) for $u_* = \mathbf{S}^{-1}[\rho_*]$ in the next section, the following stronger convergence result is needed.

Proposition 4.8 (Strong L^m-compactness). *The curves $\bar{\rho}_{\tau_k} : [0, \infty[\to \mathcal{P}_2(\mathbb{R})$ converge to ρ_* in $L^m(]0, T[\times\mathbb{R})$ for every $T > 0$.*

The proof of Proposition 4.8 is obtained via an extension of the Aubin-Lions lemma as given in [26]. The precise statement is recalled here for convenience:

Theorem 4.9 (Theorem 2 in [26]). *On a Banach space X, let be given*

- *a normal coercive integrand $\mathfrak{F} : X \to [0, +\infty]$, i.e., $\mathfrak{F}$ is lower semi-continuous and its sublevels are relatively compact in X;*
- *a pseudo-distance $g : X \times X \to [0, +\infty]$, i.e., g is lower semi-continuous, and $g(\rho, \eta) = 0$ for any $\rho, \eta \in X$ with $\mathfrak{F}(\rho) < \infty$, $\mathfrak{F}(\eta) < \infty$ implies $\rho = \eta$.*

Let further U be a set of measurable functions $u :]0, T[\to X$, with a fixed $T > 0$. Under the hypotheses that

$$\sup_{u\in U} \int_0^T \mathfrak{F}(u(t))\, \mathrm{d}t < \infty \quad \text{and} \quad \limsup_{h\downarrow 0} \sup_{u\in U} \int_0^{T-h} g(u(t+h), u(t))\, \mathrm{d}t = 0, \tag{66}$$

U contains an infinite sequence $(u_n)_{n\in\mathbb{N}}$ that converges in measure (with respect to $t \in]0, T[$) to a limit $u_* :]0, T[\to X$.

Proof of Proposition 4.8. Fix some $T > 0$. We verify the hypotheses of Theorem 4.9 for a specific choice of X, $\mathfrak{F}$, g and U. First, let $X := L^m(\mathbb{R})$. Next, define

$$g(\rho, \eta) := \begin{cases} \mathbf{W}_2(\rho, \eta) & \text{if } \rho, \eta \in \mathcal{P}_2(\mathbb{R}), \\ +\infty & \text{otherwise.} \end{cases}$$

Finally, let $\mathfrak{F}$ be given by

$$\mathfrak{F}(\rho) = \begin{cases} \int_\mathbb{R} \big[(\rho(x)^{m/2})_x\big]^2 \mathrm{d}x + \int_\mathbb{R} x^2 \rho(x)\, \mathrm{d}x & \text{if } \rho \in \mathcal{P}_2(\mathbb{R}) \text{ and } (\rho^{m/2})_x \in L^2(\mathbb{R}), \\ +\infty & \text{otherwise.} \end{cases}$$

Since any elements ρ and η in the proper domain of $\mathfrak{F}$ belong to $\mathcal{P}_2(\mathbb{R})$, it is clear that $0 = g(\rho, \eta) = \mathbf{W}_2(\rho, \eta)$ implies $\rho = \eta$. Further, the lower semi-continuity of $\mathfrak{F}$ on $L^m(\mathbb{R})$ follows from Lemma A.1 in the Appendix A. Next we show that — for any given $c > 0$ — the sublevel $A_c := \{\rho \in L^m(\mathbb{R}) | \mathfrak{F}(\rho) \leq c\}$ is relatively compact in $L^m(\mathbb{R})$. To this end, we shall prove below that $B_c := \{\eta = \rho^{m/2} | \rho \in A_c\}$ is relatively compact in $L^2(\mathbb{R})$; since the map $\iota : L^2(\mathbb{R}) \to L^m(\mathbb{R})$ with $\iota(\eta) = \eta^{2/m}$ is continuous, it then follows that $A_c = \iota(B_c)$ is a relatively compact set in $L^m(\mathbb{R})$. To show relative compactness of B_c in $L^2(\mathbb{R})$, we verify the hypotheses of the Fréchet-Kolmogorov theorem, see e.g. [12, Theorem IV.8.20].

B_c *is bounded in* $L^2(\mathbb{R})$. For any given $q \geq 1$, the Gagliardo-Nirenberg interpolation inequality (in combination with Hölder's inequality if $m > 2$) provides

$$\int_{\mathbb{R}} \eta^{2q} \, dx \leq K \left(\int_{\mathbb{R}} (\eta_x)^2 \, dx \right)^{\frac{mq-1}{m+1}} \left(\int_{\mathbb{R}} \eta^{2/m} \, dx \right)^{\frac{m(q+1)}{m+1}}. \tag{67}$$

Boundedness of B_c in $L^2(\mathbb{R})$ follows directly with the choice $q = 1$; recall that $\eta^{2/m} = \rho$ lies in the proper domain of $\mathfrak{F}$ and thus has integral one.

B_c *is tight under translations.* For every $\eta \in B_c$ and any $h > 0$, we have that

$$\int_{\mathbb{R}} |\eta(x+h) - \eta(x)|^2 \, dx = \int_{\mathbb{R}} \left| \int_0^h \eta_x(x+z) \, dz \right|^2 dx \leq h \int_0^h \int_{\mathbb{R}} \left(\eta_x(x) \right)^2 dx \, dz \leq ch^2.$$

Thus the integral converges to zero uniformly on B_c as $h \downarrow 0$.

Elements of B_c *are uniformly decaying at infinity.* For every $\eta \in B_c$ and any $R > 0$, we have

$$\int_{|x|>R} \eta(x)^2 \, dx \leq \frac{1}{R} \int_{\mathbb{R}} |x|\eta(x)^2 \, dx \leq \frac{1}{R} \left(\int_{\mathbb{R}} x^2 \eta(x)^{2/m} \, dx \right)^{1/2} \left(\int_{\mathbb{R}} \eta(x)^{4-2/m} \, dx \right)^{1/2}.$$

Inside the last expression, the first integral is less than c, and also the second one is controlled in terms of c, using inequality (67) with $q = 2 - 1/m > 1$.

In conclusion, B_c satisfies the hypotheses of the Fréchet-Kolmogorov compactness theorem. It follows that $\mathfrak{F}$ has compact sublevels.

We turn to verify the two hypotheses in (66) for the set $U := \{\bar{\rho}_{\tau_k} | k \in \mathbb{N}\}$. The first hypothesis is satisfied because of (58) and (59). We establish the second hypotheses as a consequence of the τ-uniform approximate Hölder continuity (65). For this, pick $T > 0$ and $h \in]0, 1[$ arbitrary. Given $k \in \mathbb{N}$, define $N_k \in \mathbb{N}$ such that $(N_k - 1)\tau_k < T \leq N_k \tau_k$. We distinguish two cases. If $0 < h < \tau_k$, then (writing $\tau = \tau_k$ and $N = N_k$ for ease of notation)

$$\int_0^{T-h} \mathbf{W}_2\big(\bar{\rho}_\tau(t+h), \bar{\rho}_\tau(t)\big) \, dt \leq \tau \sum_{n=0}^{N-1} \frac{h}{\tau} \mathbf{W}_2(\rho_\tau^{n+1}, \rho_\tau^n) \leq hN^{1/2} \left(\sum_{\ell=0}^{\infty} \mathbf{W}_2(\rho_\tau^{\ell+1}, \rho_\tau^\ell)^2 \right)^{1/2}$$

$$\leq h\big(2\tau N \mathcal{F}(\rho^0)\big)^{1/2} \leq \big(2(T+1)\mathcal{F}(\rho^0)\big)^{1/2} h$$

by (49). If, on the other hand, $h \geq \tau_k$, then there is a $J \in \mathbb{N}$ with $h \leq J\tau_k \leq 2h$, and so

$$\int_0^{T-h} \mathbf{W}_2\big(\bar{\rho}_\tau(t+h), \bar{\rho}_\tau(t)\big) \, dt \leq \tau \sum_{n=0}^{N-1} \sum_{j=0}^{J-1} \mathbf{W}_2(\rho_\tau^{n+j+1}, \rho_\tau^{n+j})$$

$$\leq \tau \sum_{n=0}^{N-1} \left(J^{1/2} \left[\sum_{\ell=0}^{\infty} \mathbf{W}_2(\rho_\tau^{\ell+1}, \rho_\tau^\ell)^2 \right]^{1/2} \right) \leq N\tau \big(2J\tau\mathcal{F}(\rho^0)\big)^{1/2} \leq 2(T+1)\mathcal{F}(\rho^0)^{1/2} h^{1/2}.$$

Theorem 4.9 now provides for every subsequence $(\tau_{k'}) \subseteq (\tau_k)$ the existence of a subsubsequence $(\tau_{k''}) \subseteq (\tau_{k''})$ such that $\bar{\rho}_{\tau_{k''}}$ converges in measure with respect to $t \in]0, T[$ in $L^m(\mathbb{R})$ to some limit ρ_+. By convergence of $\bar{\rho}_{\tau_k}(t)$ to $\rho_*(t)$ in $\mathbf{W}_2$ for every $t \in [0, T]$, it follows that $\rho_+ = \rho_*$. By the usual arguments, we conclude that the entire sequence $(\bar{\rho}_{\tau_k})$ converges to ρ_* in measure. In combination with the τ-uniform bound (48) of $\bar{\rho}_\tau$ in $L^m(\mathbb{R})$, we can invoke Lebesgue's dominated convergence theorem to conclude strong convergence of $\bar{\rho}_{\tau_k}$ to ρ_* in $L^m(0, T; L^m(\mathbb{R}))$. $\qquad\square$

Corollary 4.10. *For every* $T > 0$*, we have* $\rho_*^{m/2} \in L^2(0, T; H^1(\mathbb{R}))$ *and* $(\rho_*^m)_x \in L^1(0, T; L^1(\mathbb{R}))$.

Proof. Fix $T > 0$. By estimate (58), $\bar{\rho}_{\tau_k}^{m/2}$ is uniformly bounded in the reflexive Banach space $L^2(0, T; H^1(\mathbb{R}))$. By Alaoglu's theorem, there is a subsequence $(\tau_{k'}) \subseteq (\tau_k)$ such that $\bar{\rho}_{\tau_{k'}}^{m/2}$ converges weakly to some limit ζ in that space. Since $\bar{\rho}_{\tau_{k'}}^{m/2}$ converges strongly to $\rho_*^{m/2}$ in $L^2(0, T; L^2(\mathbb{R}))$ by Proposition (4.8), it follows that $\rho_*^{m/2} = \zeta \in L^2(0, T; H^1(\mathbb{R}))$.

The second claim is a trivial consequence of the representation $(\rho_*^m)_x = 2\rho_*^{m/2}(\rho_*^{m/2})_x$. $\qquad\square$

4.5. **Derivation of the entropy formulation.** In this subsection we show that $u_* = \mathbf{S}^{-1}[\rho_*]$ satisfies the entropy formulation (25). The following proposition plays the role of Lemma 2.4 in [18]. We emphasize that although the entropy inequality contains the same dissipation term as in [18], we derive it from a completely different source. Here, it results naturally from the variational construction by minimizing movements.

Proposition 4.11. *Define $u_* = \mathbf{S}^{-1}[\rho_*]$ from ρ_* via scaling, see (11). Then u_* satisfies the entropy inequality (25) for any $k > 0$ and for any non-negative test function $\varphi \in C_c^\infty(\mathbb{R}_+ \times \mathbb{R})$.*

Proof. In view of Remark 2.6, it suffices to prove the estimate (25) for all test functions φ of the form $\varphi(t, y) = \theta(t)\phi(y)$, with arbitrary non-negative $\theta \in C_c^\infty(\mathbb{R}_+)$ and $\phi \in C_c^\infty(\mathbb{R})$. Let ϕ and θ as well as $k > 0$ be fixed in the following.

The proof of the entropy inequality results from another application of Lemma 4.2. For given parameters $\nu > 0$ and $0 < \epsilon < k$, define

$$\Psi_{\epsilon,\nu}(\eta) = \int_{\mathbb{R}} S_\epsilon\Big(\frac{\eta(x)}{\mathfrak{T}_x(x)}\Big)\phi \circ \mathfrak{T}(x)\mathfrak{T}_x(x)\,\mathrm{d}x + \nu\mathbf{H}(\eta), \tag{68}$$

where the function $S_\epsilon : \mathbb{R} \to \mathbb{R}$ is given by

$$S_\epsilon(s) := \int_0^s \mathrm{Sgn}_\epsilon(r^m - k^m)\,\mathrm{d}r,$$

and has first and second derivatives:

$$S_\epsilon'(s) = \mathrm{Sgn}_\epsilon(s^m - k^m), \quad S_\epsilon''(s) = 2ms^{m-1}\delta_\epsilon(s^m - k^m).$$

Observe that there is a constant K_ϵ such that

$$|S_\epsilon(s)| \le s, \quad |S_\epsilon'(s)| \le 1, \quad 0 \le S_\epsilon''(s) \le K_\epsilon s^{-1} \quad \text{for all } s \in \mathbb{R}_+, \tag{69}$$

which can be verified by elementary calculations, using the definitions of δ_ϵ and Sgn_ϵ.

As a preliminary step, we show that there is a κ-flow associated to $\Psi_{\epsilon,\nu}$, which is given as the solution operator $\mathfrak{S}_{\epsilon,\nu}$ to

$$\eta_s = \Big(\eta\Big[\mathrm{Sgn}_\epsilon\Big(\Big(\frac{\eta}{\mathfrak{T}_x}\Big)^m - k^m\Big)\phi \circ \mathfrak{T}\Big]_x\Big)_x + \nu\eta_{xx}. \tag{70}$$

This will be achieved by application of Lemma 3.4. In the situation at hand, we have

$$F(x, \eta) = S_\epsilon\Big(\frac{\eta}{\mathfrak{T}_x(x)}\Big)\phi \circ \mathfrak{T}(x)\mathfrak{T}_x(x) + \nu\eta\log\eta,$$

and it is easily seen that (32) is satisfied (with $c = \nu > 0$). The associated function H reads

$$H(x, \xi) = \xi S_\epsilon\Big(\frac{1}{\xi\mathfrak{T}_x(x)}\Big)\phi \circ \mathfrak{T}(x)\mathfrak{T}_x(x) - \nu\log\xi.$$

We calculate the entries of the matrix in (35):

$$H_{\xi\xi}(x, \xi) = \xi^{-3}S_\epsilon''\Big(\frac{1}{\xi\mathfrak{T}_x(x)}\Big)\frac{\phi \circ \mathfrak{T}(x)}{\mathfrak{T}_x(x)} + \nu\xi^{-2},$$

$$H_{xx}(x, \xi) = \xi S_\epsilon\Big(\frac{1}{\xi\mathfrak{T}_x(x)}\Big)f_1(x) + S_\epsilon'\Big(\frac{1}{\xi\mathfrak{T}_x(x)}\Big)f_2(x) + \xi^{-1}S_\epsilon''\Big(\frac{1}{\xi\mathfrak{T}_x(x)}\Big)f_3(x),$$

$$H_{x\xi}(x, \xi) = S_\epsilon\Big(\frac{1}{\xi\mathfrak{T}_x(x)}\Big)f_4(x) + \xi^{-1}S_\epsilon'\Big(\frac{1}{\xi\mathfrak{T}_x(x)}\Big)f_5(x) + \xi^{-2}S_\epsilon''\Big(\frac{1}{\xi\mathfrak{T}_x(x)}\Big)f_6(x).$$

Here f_1 to f_6 are continuous functions of compact support in $\mathbb{R}$, explicitly expressible in terms of $\mathfrak{T}$ and its derivatives. Using the properties (69), it is easily seen that there is a constant M_ϵ such that

$$H_{\xi\xi}(x, \xi) \ge \nu\xi^{-2}, \quad H_{xx}(x, \xi) \ge -M_\epsilon, \quad |H_{x\xi}(x, \xi)| \le M_\epsilon\xi^{-1}$$

holds for all $\xi \in \mathbb{R}_+$ and uniformly in $x \in \mathbb{R}$. Thus, for a suitable choice of $\kappa = \kappa_{\epsilon,\nu} < 0$, the matrix in (35) is positive for every $\xi \in \mathbb{R}_+$ and $x \in \mathbb{R}$, and so H is jointly convex. By Lemma 3.4, the solution operator $\mathfrak{S}_{\epsilon,\nu}$ for (70) is a $\kappa_{\epsilon,\nu}$-flow for $\Psi_{\epsilon,\nu}$.

In order to apply Corollary 4.3, we need to calculate the derivative of $\mathcal{F}$ along solutions η to (70). For simplification, introduce the rescaling $v(t) = \mathbf{S}^{-1}[\eta(t)]$, see (11). Further, recall the definitions of $\mathfrak{T}$ and of $\alpha = \mathfrak{T}_x \circ \mathfrak{T}^{-1}$. Using the functional $\mathfrak{K}_\mathbf{H}$ defined in (63), we have

$$\frac{\mathrm{d}}{\mathrm{d}s}\mathcal{F}(\mathfrak{S}^s_{\epsilon,\nu}\eta_0) = -\int_\mathbb{R} \eta\big[a\eta^{m-1}\big]_x\big[\operatorname{Sgn}_\epsilon\big((\eta/\mathfrak{T}_x)^m - k^m\big)\phi\circ\mathfrak{T}\big]_x \mathrm{d}x + \nu\int_\mathbb{R} a\eta^{m-1}\eta_{xx}$$

$$\leq -\int_\mathbb{R}\{\alpha v\}\circ\mathfrak{T}\big[\{(a\circ\mathfrak{T}^{-1})\cdot(\alpha v)^{m-1}\}\circ\mathfrak{T}\big]_x\big[\{\operatorname{Sgn}_\epsilon(v^m - k^m)\phi\}\circ\mathfrak{T}\big]_x \mathrm{d}x - \nu\mathfrak{K}_\mathbf{H}(\eta)$$

$$\overset{(12)}{=} -\int_\mathbb{R}\alpha^2 v\Big[\frac{m\alpha^{-2}}{m-1}v^{m-1}\Big]_y\big[\operatorname{Sgn}_\epsilon(v^m - k^m)\phi\big]_y\mathrm{d}y - \nu\mathfrak{K}_\mathbf{H}(\eta)$$

$$\overset{(16)}{=} -\int_\mathbb{R}\big[(v^m)_y + bv^m\big][\operatorname{Sgn}_\epsilon(v^m - k^m)\phi]_y\mathrm{d}y - \nu\mathfrak{K}_\mathbf{H}(\eta)$$

$$= -\int_\mathbb{R}\big[(v^m)_y + b(v^m - k^m)\big]\operatorname{Sgn}_\epsilon(v^m - k^m)\phi_y\mathrm{d}y - \int_\mathbb{R}k^m b\big[\operatorname{Sgn}_\epsilon(v^m - k^m)\phi\big]_y\mathrm{d}y$$

$$\quad - \int_\mathbb{R}\big[(v^m)_y + b(v^m - k^m)\big](v^m)_y\operatorname{Sgn}'_\epsilon(v^m - k^m)\phi\mathrm{d}y - \nu\mathfrak{K}_\mathbf{H}(\eta)$$

$$= \int_\mathbb{R}R_\epsilon(v^m)\phi_{yy}\mathrm{d}y - \int_\mathbb{R}\big(b(v^m - k^m)\phi_y - k^m b_y\phi\big)\operatorname{Sgn}_\epsilon(v^m - k^m)\mathrm{d}y$$

$$\quad - \int_\mathbb{R}\big[P_\epsilon(v^m)_y\big]^2\phi\mathrm{d}y - \int_\mathbb{R}bQ_\epsilon(v^m)_y\phi\mathrm{d}y - \nu\mathfrak{K}_\mathbf{H}(\eta).$$

In the last step, we have implicitly defined the smooth functions $P_\epsilon, Q_\epsilon, R_\epsilon : \mathbb{R}_+ \to \mathbb{R}$ such that

$$R'_\epsilon(s) = \operatorname{Sgn}_\epsilon(s - k^m), \quad P'_\epsilon(s)^2 = \operatorname{Sgn}'_\epsilon(s - k^m), \quad Q'_\epsilon(s) = (s - k^m)\operatorname{Sgn}'_\epsilon(s - k^m).$$

Accordingly, we define, still with $v = \mathbf{S}^{-1}[\eta]$,

$$\mathfrak{K}_{\epsilon,\nu}(\eta) = -\int_\mathbb{R}S_\epsilon(v^m)\phi_{yy}\mathrm{d}y + \int_\mathbb{R}\big(b(v^m - k^m)\phi_y - k^m b_y\phi\big)\operatorname{Sgn}_\epsilon(v^m - k^m)\mathrm{d}y$$

$$\quad + \int_\mathbb{R}\big[P_\epsilon(v^m)_y\big]^2\phi\mathrm{d}y - \int_\mathbb{R}(b\phi)_y Q_\epsilon(v^m)\mathrm{d}y + \nu\mathfrak{K}_\mathbf{H}(\eta).$$

Considered as a functional of v, the right-hand side is lower semi-continuous with respect to strong convergence of v in $L^m(\mathbb{R})$. Indeed, all of the integral expressions are even continuous in $L^m(\mathbb{R})$, except for the one involving P_ϵ, for which lower semi-continuity can be concluded by means of Lemma A.1. Since convergence of η in $L^m(\mathbb{R})$ is equivalent to convergence of $v = \mathbf{S}^{-1}[\eta]$ in $L^m(\mathbb{R})$, the functional $\mathfrak{K}_{\epsilon,\nu}$ is lower-semicontinuous with respect to η. Now since $\mathfrak{S}^s_{\epsilon,\nu}\eta_0$ converges to η_0 in $L^m(\mathbb{R})$ for every $\eta_0 \in L^m(\mathbb{R})$, we conclude that $\mathfrak{K}_{\epsilon,\nu}$ satisfies the condition (55) for the application of Corollary 4.3.

We shall now derive a refined version of estimate (57). To this end, recall that $\theta \in C^\infty_c(\mathbb{R}_+)$ is a temporal test function. Multiply (56) by $\theta(n\tau)$ and sum over n to find (since $\kappa_{\epsilon,\nu} < 0$)

$$\tau\sum_{n=1}^\infty\Psi_{\epsilon,\nu}(\rho^n_\tau)\frac{\theta(n\tau) - \theta((n+1)\tau)}{\tau} \geq \tau\sum_{n=1}^\infty\mathfrak{K}_{\epsilon,\nu}(\rho^n_\tau)\theta(n\tau) + \kappa_{\epsilon,\nu}\tau\mathcal{F}(\rho^0)\sup|\theta|.$$

Thanks to the strong convergence of $\bar\rho_\tau$ to ρ_* in $L^m(]0, T[\times\mathbb{R})$ and the lower semi-continuity of $\mathfrak{K}_{\epsilon,\nu}$, we can pass to the time-continuous limit $\tau \downarrow 0$ and obtain

$$\int_0^T\Psi_{\epsilon,\nu}(\rho_*(t))\theta_t(t)\mathrm{d}t \geq \int_0^T\theta(t)\mathfrak{K}_{\epsilon,\nu}(\rho_*(t))\mathrm{d}t.$$

Taking into account estimate (61) on $\mathbf{H}(\rho_*(t))$, we can now pass to the inviscid limit $\nu \downarrow 0$ and find, recalling $\varphi(t, y) = \theta(t)\phi(y)$, that

$$\int_0^T\int_\mathbb{R}S_\epsilon(u_*)\varphi_t\mathrm{d}y\mathrm{d}t \geq \int_0^T\int_\mathbb{R}\Big(\big[(u^m_*)_y + b(u^m_* - k^m)\big]\varphi_y - k^m b_y\varphi\Big)\operatorname{Sgn}_\epsilon(u^m_* - k^m)\mathrm{d}y\mathrm{d}t$$

$$\quad + \int_0^T\int_\mathbb{R}\big[P_\epsilon(u^m_*)_y\big]^2\varphi\mathrm{d}y\mathrm{d}t - \int_0^T\int_\mathbb{R}(b\varphi)_y Q_\epsilon(u^m_*)\mathrm{d}y\mathrm{d}t.$$

Notice that we have integrated by parts in the first term,

$$-\int_0^T \int_{\mathbb{R}} R_\epsilon(u_*^m)\varphi_{yy}\,\mathrm{d}y\,\mathrm{d}t = \int_0^T \int_{\mathbb{R}} (u_*^m)_y \operatorname{Sgn}_\epsilon(u_*^m - k^m)\varphi_y\,\mathrm{d}y\,\mathrm{d}t,$$

which is admissible since $(\rho_*^m)_x \in L^1(0,T;L^1(\mathbb{R}))$ by Corollary 4.10. In the final step, we pass to the limit $\epsilon \downarrow 0$. By the uniform convergence of $S_\epsilon(s)$ to $|s-k|-k$, by the uniform convergence of Q_ϵ to zero, and since

$$\big[P_\epsilon(u_*)_y\big]^2 = \operatorname{Sgn}_\epsilon'(u_*^m - k^m)\big[(u_*^m)_y\big]^2$$

by definition of P_ϵ, we finally obtain (25). $\qquad\square$

4.6. L^∞ bounds.
Definition 2.4 of entropy solutions requires $u_* \in L^1 \cap L^\infty(]0,T[\times\mathbb{R})$ for every $T > 0$. The L^1-bound is obvious from our construction. Below, we prove the L^∞-bound.

Proposition 4.12. *Assume that $\rho^0 \in \mathcal{P}_2(\mathbb{R}) \cap L^\infty(\mathbb{R})$, and let $k > 0$ be such that $\rho^0(x) \leq ka(x)^{-1/(m-1)}$ for almost all $x \in \mathbb{R}$. Then $\rho_*(T,x) \leq ka(x)^{-1/(m-1)}$ for all $T > 0$ and almost every $x \in \mathbb{R}$.*

Proof. Once again, the proof is obtained in application of Lemma 4.2. As auxiliary functional, we choose

$$\Psi_{\epsilon,\nu}(\eta) = \int_{\mathbb{R}} \operatorname{PosPar}_\epsilon\big(a(x)^{\frac{1}{m-1}}\eta(x) - k\big)a(x)^{\frac{-1}{m-1}}\,\mathrm{d}x + \nu\mathbf{H}(\eta)$$

with $\nu > 0$ and $\epsilon \in]0,k[$. We verify that $\Psi_{\epsilon,\nu}$ satisfies the assumptions required in Lemma 3.4. With the short hand notation $A = a^{1/(m-1)}$, we have

$$H(x,\xi) = \xi \operatorname{PosPar}_\epsilon\Big(\frac{A(x)}{\xi} - k\Big)A(x)^{-1} - \nu\log\xi.$$

Recall that $\operatorname{PosPar}_\epsilon' = \operatorname{Stp}_\epsilon$ and $\operatorname{PosPar}_\epsilon'' = \delta_\epsilon$, and observe that there is some $K_{k,\epsilon}$ such that

$$0 \leq \operatorname{PosPar}_\epsilon(s-k) \leq s, \quad 0 \leq \operatorname{Stp}_\epsilon(s-k) \leq 1, \quad 0 \leq \delta_\epsilon(s-k) \leq K_{k,\epsilon}s^{-1} \qquad (71)$$

for all $s \in \mathbb{R}_+$. We obtain for the second derivatives of H:

$$H_{\xi\xi}(x,\xi) = \frac{A(x)}{\xi^3}\delta_\epsilon\Big(\frac{A(x)}{\xi} - k\Big) + \nu\xi^{-2},$$

$$H_{xx}(x,\xi) = f_1(x)\xi \operatorname{PosPar}_\epsilon\Big(\frac{A(x)}{\xi} - k\Big) + f_2(x)\operatorname{Stp}_\epsilon\Big(\frac{A(x)}{\xi} - k\Big) + \frac{f_3(x)}{\xi}\delta_\epsilon\Big(\frac{A(x)}{\xi} - k\Big),$$

$$H_{x\xi}(x,\xi) = f_4(x)\operatorname{PosPar}_\epsilon\Big(\frac{A(x)}{\xi} - k\Big) + \frac{f_5(x)}{\xi}\operatorname{Stp}_\epsilon\Big(\frac{A(x)}{\xi} - k\Big) + \frac{f_6(x)}{\xi^2}\delta_\epsilon\Big(\frac{A(x)}{\xi} - k\Big),$$

where the functions f_1 to f_6 are explicitly expressible in terms of A, A_x and A_{xx}, and are uniformly bounded in $x \in \mathbb{R}$. Taking into account (71) and that $A \geq \underline{a}^{1/(m-1)}$, we arrive at the uniform bounds

$$H_{\xi\xi}(x,\xi) \geq \nu\xi^{-2}, \quad H_{xx}(x,\xi) \geq -M_\epsilon, \quad |H_{x\xi}(x,\xi)| \leq M_\epsilon\xi^{-1}$$

with a suitable constant M_ϵ. Consequently, by application of Lemma 3.4, the solution operator $\mathfrak{S}_\Psi$ to

$$\eta_s = \mathrm{D}_x\big(\eta\mathrm{D}_x\big[\operatorname{Stp}_\epsilon(a^{\frac{1}{m-1}}\eta - k)\big]\big) + \nu\eta_{xx}$$

defines a κ-flow for $\Psi_{\epsilon,\nu}$, with some $\kappa = \kappa_{\epsilon,\nu} < 0$. The regularizing effect of the viscous term is strong enough to justify the following calculations:

$$\frac{\mathrm{d}}{\mathrm{d}s}\mathcal{F}(\mathfrak{S}_\Psi^s\eta_0) = \int_{\mathbb{R}} a\eta^{m-1}\big(\eta\big[\operatorname{Stp}_\epsilon(a^{\frac{1}{m-1}}\eta - k)\big]_x\big)_x\,\mathrm{d}x + \nu\int_{\mathbb{R}} a\eta^{m-1}\eta_{xx}\,\mathrm{d}x$$

$$\leq -\int_{\mathbb{R}} \eta\big[(a^{\frac{1}{m-1}}\eta)^{m-1}\big]_x\big[\operatorname{Stp}_\epsilon(a^{\frac{1}{m-1}}\eta - k)\big]_x\,\mathrm{d}x + \frac{\nu\overline{a_{xx}}}{m}\int_{\mathbb{R}}\eta^m\,\mathrm{d}x.$$

Here we have used that the second integral with ν can be estimated as in (62), neglecting a positive term. The product under the first integral is always non-negative since both the functions

$s \mapsto s^{m-1}$ and $s \mapsto \mathrm{Stp}_\epsilon(s-k)$ are differentiable and increasing. Thus, condition (55) is satisfied with

$$\mathfrak{K}(\eta) := -\frac{\nu \overline{a_{xx}}}{m} \int_{\mathbb{R}} \eta^m \, \mathrm{d}x,$$

which is obviously continuous with respect to strong convergence in $L^m(\mathbb{R})$. Estimate (57) in combination with (48) yields

$$\Psi_{\epsilon,\nu}(\rho_\tau^N) \leq \Psi_{\epsilon,\nu}(\rho^0) + \frac{\nu \overline{a_{xx}}}{m} \tau \sum_{n=1}^{N} \int_{\mathbb{R}} (\rho_\tau^n)^m \, \mathrm{d}x + \tau(-\kappa_{\epsilon,\nu})\mathcal{F}(\rho^0)$$

$$\leq \Psi_{\epsilon,\nu}(\rho^0) + \Big(\frac{\nu \overline{a_{xx}}}{m\underline{a}} N\tau + \tau(-\kappa_{\epsilon,\nu})\Big)\mathcal{F}(\rho^0).$$

For fixed positive parameters ν and ϵ, the modulus $\kappa_{\epsilon,\nu} < 0$ of convexity is a τ-independent constant. We can thus pass to the limit $\tau_k \downarrow 0$ in (57) and obtain, using the lower semi-continuity of $\Psi_{\epsilon,\nu}$ in $\mathbf{W}_2$, that

$$\Psi_{\epsilon,\nu}(\rho_*(T)) \leq \Psi_{\epsilon,\nu}(\rho^0) + \frac{\nu \overline{a_{xx}}}{m\underline{a}} T \mathcal{F}(\rho^0)$$

for every $T \geq 0$. Using further that $\mathbf{H}(\rho_*(T))$ is a finite quantity, see (61), we can pass to the limit $\nu \downarrow 0$ and obtain

$$\int_{\mathbb{R}} \mathrm{PosPar}_\epsilon \big(a(x)^{\frac{1}{m-1}} \rho_*(T,x) - k\big) a(x)^{\frac{-1}{m-1}} \, \mathrm{d}x \leq \int_{\mathbb{R}} \mathrm{PosPar}_\epsilon \big(a(x)^{\frac{1}{m-1}} \rho^0(x) - k\big) a(x)^{\frac{-1}{m-1}} \, \mathrm{d}x.$$

By the properties of a, and since $\rho^0, \rho_*(t) \in L^m(\mathbb{R})$, we can further pass to the limit $\epsilon \downarrow 0$, which yields

$$\int_{\mathbb{R}} \big[\rho_*(T,x) - ka(x)^{\frac{-1}{m-1}}\big]_+ \, \mathrm{d}x \leq \int_{\mathbb{R}} \big[\rho^0(x) - ka(x)^{\frac{-1}{m-1}}\big]_+ \, \mathrm{d}x.$$

Since the integral on the right-hand side is zero by hypothesis, so is the integral on the left-hand side. This proves the claim. $\qquad\square$

Corollary 4.13. *Provided that $\rho^0 \in L^\infty(\mathbb{R})$, it follows that $\rho_*^m \in L^2(0,T;H^1(\mathbb{R}))$.*

Proof. By Proposition 4.12,

$$K := \sup_{t\in\mathbb{R}_+} \operatorname*{ess\,sup}_{x\in\mathbb{R}} \rho_*(t,x) < \infty,$$

and so

$$\int_{\mathbb{R}} \big[(\rho_*(t)^m)_x\big]^2 \, \mathrm{d}x = \int_{\mathbb{R}} \big[2\rho_*(t)^{m/2}(\rho_*(t)^{m/2})_x\big]^2 \, \mathrm{d}x \leq 4K^m \int_{\mathbb{R}} \big[(\rho_*(t)^{m/2})_x\big]^2 \, \mathrm{d}x$$

for all $t \geq 0$. The claim now follows from Corollary 4.10. $\qquad\square$

4.7. **Continuity at $t = 0$.** It remains to verify that u_* attains the initial condition u^0.

Proposition 4.14. $u_*(t) \to u^0$ in $L^1(\mathbb{R})$ as $t \downarrow 0$.

Proof. Since the rescaling $\mathbf{S}$ from (11) is a homeomorphism on $L^1(\mathbb{R})$, it suffices to show that $\rho_*(t) \to \rho^0$ in $L^1(\mathbb{R})$ as $t \downarrow 0$. By the lower semi-continuity of $\mathcal{F}$, we can pass to the time-continuous limit $\tau \downarrow 0$ in (48) and obtain

$$\mathcal{F}(\rho_*(t)) \leq \liminf_{\tau\downarrow 0} \mathcal{F}(\bar\rho_\tau(t)) \leq \mathcal{F}(\rho^0)$$

for every $t > 0$, which implies that

$$\limsup_{t\downarrow 0} \mathcal{F}(\rho_*(t)) \leq \mathcal{F}(\rho^0). \tag{72}$$

On the other hand, since the limiting curve ρ_* is continuous in $\mathbf{W}_2$, once again the lower semi-continuity of $\mathcal{F}$ in $\mathbf{W}_2$ yields

$$\mathcal{F}(\rho^0) \leq \liminf_{t\downarrow 0} \mathcal{F}(\rho_*(t)).$$

In combination with (72), we have that

$$\lim_{t\downarrow 0} \mathcal{F}(\rho_*(t)) = \mathcal{F}(\rho^0). \tag{73}$$

By definition, $\mathcal{F}(\rho)$ is the mth power of the L^m-norm of ρ with respect to the non-uniform background measure $m^{-1}a(x)\,\mathrm{d}x$. Since $m > 1$, and since the weight function a satisfies the bounds (a1)&(a2), we can conclude by standard arguments, see e.g. [21, Theorem 2.11], that weak convergence of $\rho_*(t)$ and convergence (73) together imply strong convergence of $\rho_*(t)$ to ρ^0 in $L^m(\mathbb{R})$. To obtain convergence in $L^1(\mathbb{R})$, we apply the generalized Hölder inequality with exponents $2m/(m-1)$, 2 and $2m$:

$$\int_{\mathbb{R}} |\rho_*(t) - \rho^0|\,\mathrm{d}x \le \int_{\mathbb{R}} (1+|x|^2)^{-\frac{1}{2}} \left[(1+|x|^2)|(\rho_*(t)+\rho^0)\right]^{\frac{1}{2}} \left[|\rho_*(t)-\rho^0|\right]^{\frac{1}{2}}\,\mathrm{d}x$$

$$\le \left[\int_{\mathbb{R}} (1+|x|^2)^{-\frac{m}{m-1}}\,\mathrm{d}x\right]^{\frac{m-1}{2m}} \left[\int_{\mathbb{R}} (1+|x|^2)(\rho_*(t)+\rho^0)\,\mathrm{d}x\right]^{\frac{1}{2}} \|\rho_*(t)-\rho^0\|_{L^m}^{\frac{1}{2m}}.$$

The first integral on the right hand side is clearly finite, and the second integral remains uniformly bounded as $t \downarrow 0$, since continuity in $\mathbf{W}_2$ implies continuity of the second moment. The last term vanishes for $t \downarrow 0$ because of the strong convergence of $\rho_*(t)$ in $L^m(\mathbb{R})$. $\qquad\square$

4.8. **Proof of Theorem 2.7.** At this point, we have proven that u_* meets all the requirements for being an entropy solution as stated in Definition 2.4: we have $u_* \in L^1 \cap L^\infty(\mathbb{R})$ by Proposition 4.12 and $u_*^m \in L^2(0,T;H^1(\mathbb{R}))$ by Corollary 4.13; we further have continuity of u_* at $t = 0$ by Lemma 4.14, and the validity of the entropy inequality has been verified in Proposition 4.11. Finally, convergence of the time-discrete approximation scheme in Wasserstein and in L^m have been shown in Lemma 4.7 and Proposition 4.8, respectively.

5. Uniqueness of entropy solutions

In this section we prove Theorem 2.8, using the *doubling of the variables* device. Since we follow almost literally the proof of [18, Theorem 1.1], we restrict ourselves to the key steps and refer the interested reader to the original article [18] for more details.

For a fixed $T > 0$ we shall use the notation $\Pi_T = \mathbb{R}\times]0,T[$. Let $\varphi \in C_c^\infty(\Pi_T \times \Pi_T)$ be a non-negative test function, and assume that u and v are entropy solutions in the sense of Definition 2.4. For brevity, we write $v = v(x,t)$, $u = u(y,s)$ and $\varphi = \varphi(x,t,y,s)$; note that $y \in \mathbb{R}$ is a "doubling variable" that is not related to y in (1). From Proposition 4.11 we obtain

$$-\iint_{\Pi_T}\iint_{\Pi_T} \big(|v-u|\varphi_t + \mathrm{Sgn}(v-u)[(v^m-u^m)b(x) - (v^m)_x]\varphi_x$$
$$- \mathrm{Sgn}(v-u)b_x(x)u^m\varphi\big)\,\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}y\,\mathrm{d}s \tag{74}$$
$$\le -\limsup_{\epsilon\downarrow 0}\iint_{\Pi_T}\iint_{\Pi_T} (v^m)_x^2\,\mathrm{Sgn}_\epsilon'(v^m-u^m)\varphi\,\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}y\,\mathrm{d}s.$$

Recall that Sgn_ϵ is a smooth uniformly convergent approximation of the sign function, which is obtained by mollification with δ_ϵ. Since $(v^m)_x \in L^2(\Pi_T)$, the following integration by parts is justified:

$$-\iint_{\Pi_T} \mathrm{Sgn}_\epsilon(v^m-u^m)(v^m)_x\varphi_y\,\mathrm{d}y\,\mathrm{d}s = \iint_{\Pi_T} (\mathrm{Sgn}_\epsilon(v^m-u^m))_y(v^m)_x\varphi\,\mathrm{d}y\,\mathrm{d}s. \tag{75}$$

We integrate (75) w.r.t. $(x,t) \in \Pi_T$ and send $\epsilon \downarrow 0$. By the dominated convergence theorem, that yields

$$-\iint_{\Pi_T}\iint_{\Pi_T} \mathrm{Sgn}(v-u)(v^m)_x\varphi_y\,\mathrm{d}y\,\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}t$$
$$= -\lim_{\epsilon\downarrow 0}\iint_{\Pi_T}\iint_{\Pi_T} (u^m)_y(v^m)_x\,\mathrm{Sgn}_\epsilon'(v^m-u^m)\varphi\,\mathrm{d}y\,\mathrm{d}s\,\mathrm{d}x\,\mathrm{d}t. \tag{76}$$

Adding (74) and (76) we get

$$
-\iint_{\Pi_T} \iint_{\Pi_T} \left(|v-u|\varphi_t + \mathrm{Sgn}(v-u)\left[(v^m - u^m)b(x)\varphi_x - (v^m)_x(\varphi_x + \varphi_y) \right] \right.
$$
$$
\left. - \mathrm{Sgn}(v-u)b_x(x)u^m\varphi \right) \mathrm{d}x\,\mathrm{d}t\,\mathrm{d}y\,\mathrm{d}s \tag{77}
$$
$$
\leq -\limsup_{\epsilon \to 0} \iint_{\Pi_T} \iint_{\Pi_T} \left[(v^m)_x^2 - (u^m)_y(v^m)_x \right] \mathrm{Sgn}'_\epsilon(v^m - u^m)\varphi\,\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}y\,\mathrm{d}s.
$$

The terms in (77) containing b can be rewritten as follows:

$$
\mathrm{Sgn}(v-u)(v^m - u^m)b(x)\varphi_x - \mathrm{Sgn}(v-u)b_x(x)u^m\varphi
$$
$$
= \mathrm{Sgn}(v-u)(v^m b(x) - u^m b(y))\varphi_x + \mathrm{Sgn}(v-u)(u^m(b(y) - b(x))\varphi)_x. \tag{78}
$$

Now, we repeat the previous steps with a simultaneous interchange of the roles of u and v and the roles of (y,s) and (x,t). Summation of (77) with its respective counter part yields

$$
-\iint_{\Pi_T} \iint_{\Pi_T} \left(|v-u|(\varphi_t + \varphi_s) + \mathrm{Sgn}(v-u)\left[v^m b(x) - u^m b(y) \right](\varphi_x + \varphi_y) \right.
$$
$$
+ |v^m - u^m|(\varphi_{xx} + 2\varphi_{xy} + \varphi_{yy})
$$
$$
\left. + \mathrm{Sgn}(v-u)\left[(u^m(b(y) - b(x))\varphi)_x - (v^m(b(x) - b(y))\varphi)_y \right] \right) \mathrm{d}x\,\mathrm{d}t\,\mathrm{d}y\,\mathrm{d}s \tag{79}
$$
$$
\leq -\limsup_{\epsilon \to 0} \iint_{\Pi_T} \iint_{\Pi_T} |(v^m)_x - (u^m)_y|^2 \mathrm{Sgn}'_\epsilon(v^m - u^m)\varphi\,\mathrm{d}x\,\mathrm{d}t\,\mathrm{d}y\,\mathrm{d}s \leq 0,
$$

where we have used the identity

$$
\mathrm{Sgn}(v-u)((v^m)_x - (u^m)_y) = \left[|v^m - u^m|_x + |v^m - u^m|_y \right]
$$

and integrated by parts. We emphasize that for the estimation on the right-hand side of (79), the dissipation term from the definition (25) has been essential.

As usual, $\varphi = \varphi(x,t,y,s)$ is chosen in product form,

$$
\varphi(x,t,y,s) = \psi\left(\frac{x+y}{2}, \frac{t+s}{2} \right)\omega_\sigma\left(\frac{x-y}{2} \right)\delta_\sigma\left(\frac{t-s}{2} \right),
$$

where $0 \leq \psi \in C_c^\infty(\Pi_T)$ is a test function, $\delta_\sigma = \sigma^{-1}\delta_1(t/\sigma)$ for $\sigma > 0$ and δ_1 is the mollifier defined in (7), and

$$
\omega_\sigma(x) = \frac{1}{2\sigma}\delta_1\left(\frac{|x|^2}{\sigma^2} \right).
$$

Accordingly, we introduce a new set of variables $(\bar{x}, \bar{t}, z, \tau)$ with

$$
\bar{x} = \frac{x+y}{2}, \quad \bar{t} = \frac{t+s}{2}, \quad z = \frac{x-y}{2}, \quad \tau = \frac{t-s}{2},
$$

for which we have (by the usual abuse of notation)

$$
\partial_{\bar{t}} = \partial_t + \partial_s, \quad \partial_{\bar{x}} = \partial_x + \partial_y, \quad \partial_{\bar{x}\bar{x}} = \partial_{xx} + \partial_{yy} + 2\partial_{xy}.
$$

With the understanding that $u = u(y,s)$ and $v = v(x,t)$ while $\psi = \psi(\bar{x}, \bar{t})$, $\omega_\sigma = \omega_\sigma(z)$ and $\delta_\sigma = \delta_\sigma(\tau)$, the inequality (79) can be written as

$$
0 \geq J_\sigma := -\iint_{\Pi_T} \iint_{\Pi_T} \left[\left\{ |v-u|\psi_{\bar{t}} + \mathrm{Sgn}(v-u)(v^m b(x) - u^m b(y))\psi_{\bar{x}} + |v^m - u^m|\psi_{\bar{x}\bar{x}} \right. \right.
$$
$$
+ \mathrm{Sgn}(v-u)\left[((u^m(b(y) - b(x)))_x - (v^m(b(x) - b(y)))_y)\psi \right.
$$
$$
\left. \left. + u^m(b(y) - b(x))\psi_x - v^m(b(x) - b(y))\psi_y \right] \right\}\omega_\sigma\delta_\sigma \tag{80}
$$
$$
\left. + \frac{1}{2}(b(x) - b(y))|v^m - u^m|\psi\delta_\sigma(\omega_\sigma)_z \right] \mathrm{d}x\,\mathrm{d}t\,\mathrm{d}y\,\mathrm{d}s.
$$

By assumption, $b \in W^{1,\infty}(\mathbb{R})$. Moreover, since u and v belong to $L^\infty(\Pi_T)$, and since $(\delta_1)_x(x) \leq 0$ for $x > 0$, we can find a constant K (depending on b) such that

$$
\begin{aligned}
(b(x) - b(y))|v^m - u^m|(\omega_\sigma)_z &= (b(x) - b(y))|v^m - u^m|\frac{z}{\sigma^3}(\delta_1)_x\left(\frac{z^2}{\sigma^2}\right) \\
&\leq K|v - u|\frac{z^2}{\sigma^2}\sigma^{-1}\chi_{|z|\leq\sigma}.
\end{aligned}
\tag{81}
$$

Now we perform the limit $\sigma \downarrow 0$, which concentrates the support of φ on the diagonals $x = y$ and $t = s$. By Lebesgue's differentiation theorem we then obtain

$$
\begin{aligned}
0 \geq \lim_{\sigma\downarrow 0} J_\sigma \geq -\iint_{\Pi_T} &\left[\left\{|v - u|\psi_t + b(x)|v(x,t)^m - u(x,t)^m|\psi_x + |v(x,t)^m - u(x,t)^m|\psi_{xx}\right.\right. \\
&\left.\left. + b_x(x)|v(x,t)^m - u(x,t)^m|\psi\right\} + K|v(x,t) - u(x,t)|\psi\right] \mathrm{d}x\,\mathrm{d}t.
\end{aligned}
\tag{82}
$$

Regrouping terms and using again that $b \in W^{1,\infty}(\mathbb{R})$, we arrive at the key estimate

$$
-\iint_{\Pi_T} \left(|u - v|\psi_t + b|v^m - u^m|\psi_x + |v^m - u^m|\psi_{xx}\right) \mathrm{d}x\,\mathrm{d}t \leq C\iint_{\Pi_T} |v - u|\psi\,\mathrm{d}x\,\mathrm{d}t,
\tag{83}
$$

for some $C > 0$ depending on K and on b.

At this point, we make the classical special choice for the test function ψ. For any given $0 < t_1 < t_2 < T$ and $r > 0$, let $\theta_\sigma \in C_c^\infty(]0,T[)$ be the σ-mollification of the characteristic function of the interval $[t_1, t_2]$, and let $\phi_r \in C_c^\infty(\mathbb{R})$ be such that $\phi_r(x) = 1$ for $|x| \leq r$ and $\phi_r(x) = 0$ for $|x| \geq r + 1$. Setting $\psi(x,t) = \phi_r(x)\theta_\sigma(t)$ yields

$$
\begin{aligned}
&\lim_{r\to\infty}\iint_{\Pi_T}\left(b|v^m - u^m|\psi_x + |v^m - u^m|\psi_{xx}\right)\mathrm{d}x\,\mathrm{d}t \\
&\leq \bar{C}\lim_{r\to\infty}\iint_{\Pi_T\cap||x|-r|\leq 1}\left(v + u\right)\mathrm{d}x\,\mathrm{d}t = 0
\end{aligned}
$$

by the dominated convergence theorem (with respect to $t \in [0,T]$), because u and v are t-uniformly bounded in $L^1(\mathbb{R})$. Therefore, by sending $r \to \infty$ in (83), we get

$$
-\int_0^T\int_{\mathbb{R}} |v - u|(\theta_\sigma)_t(t)\,\mathrm{d}x\,\mathrm{d}t \leq C\int_0^T\int_{\mathbb{R}} |v - u|\theta_\sigma(t)\,\mathrm{d}x\,\mathrm{d}t.
\tag{84}
$$

Finally, passing to $\sigma \downarrow 0$, we obtain

$$
\|u(t_2) - v(t_2)\|_{L^1} \leq \|u(t_1) - v(t_1)\|_{L^1} + C\int_{t_1}^{t_2} \|u(\tau) - v(\tau)\|_{L^1}\,\mathrm{d}\tau,
\tag{85}
$$

hence we can use the Gronwall inequality in (85) (integral form for measurable functions, cf. [13]) to obtain

$$
\|u(t_2) - v(t_2)\|_{L^1} \leq \|u(t_1) - v(t_1)\|_{L^1}(1 + C(t_2 - t_1)e^{C(t_2-t_1)}).
\tag{86}
$$

Since u and v are right continuous at $t = 0$, we can perform the limit $t_1 \downarrow 0$ in (86) and obtain stability of the entropy solutions u, v in the $L^1(\mathbb{R})$ norm. In particular, if $v(0) = u(0)$, then $v(t) = u(t)$ for all $t > 0$, which shows uniqueness.

Appendix A. A lemma on lower semi-continuity

Lemma A.1. *Let $H \in C^1(\mathbb{R}_+)$, and let $\phi \in C^0(\mathbb{R})$ be a bounded non-negative function. Define the functional $\Psi : \mathcal{P}_2(\mathbb{R}) \to [0,\infty]$ by*

$$
\Psi(\eta) = \int_{\mathbb{R}} \phi(x)\left[\mathrm{D}_x H(\eta(x))\right]^2 \mathrm{d}x
$$

whenever the integral is well-defined, and $+\infty$ otherwise. Then Ψ is sequentially lower semi-continuous in $L^p(\mathbb{R})$, for arbitrary $p \geq 1$.

Proof. Let $(\eta_n)_{n \in \mathbb{N}}$ be a sequence that converges to η_0 in $L^p(\mathbb{R})$, with $\bar{\Psi} := \sup_n \Psi(\eta_n) < \infty$. Without loss of generality, we may even assume that η_n converges to η_0 almost everywhere on $\mathbb{R}$. For every $\epsilon > 0$, introduce the functional $\Psi_\epsilon : \mathcal{P}_2(\mathbb{R}) \to [0, \infty]$ with

$$\Psi_\epsilon(\eta) = \int_{\Omega_\epsilon} \phi[\mathrm{D}_x H(\eta)]^2 \, dx \quad \text{where} \quad \Omega_\epsilon := \left\{ x \in [-\epsilon^{-1}, \epsilon^{-1}] \big| \phi(x) \geq \epsilon \right\},$$

with the understanding that $\Psi_\epsilon(\eta) = +\infty$ unless $H(\eta) \in H^1(\Omega_\epsilon)$. Since ϕ is non-negative,

$$\epsilon \int_{\Omega_\epsilon} [\mathrm{D}_x H(\eta_n)]^2 \, dx \leq \Psi_\epsilon(\eta_n) \leq \Psi(\eta_n) \leq \bar{\Psi}.$$

Thus, the functions $H(\eta_n)$ are n-uniformly bounded in $H^1(\Omega_\epsilon)$. By Alaoglu's theorem, every subsequence $(H(\eta_{n'}))_{n' \in \mathbb{N}}$ contains a subsubsequence $(H(\eta_{n''}))_{n'' \in \mathbb{N}}$ that converges weakly in $H^1(\Omega_\epsilon)$ to some limit h. Further, since Ω_ϵ is compact by definition, Rellich's Lemma applies, and $H(\eta_{n''})$ converges to h strongly in $L^2(\Omega_\epsilon)$. We have assumed that η_n converges pointwise almost everywhere to η_0, hence $H(\eta_{n''})$ converges to $H(\eta_0)$ almost everywhere, and we conclude $h = H(\eta_0) \in H^1(\Omega_\epsilon)$ — independently of the chosen subsequence. From here, it follows that

$$\mathrm{D}_x H(\eta_n) \rightharpoonup \mathrm{D}_x H(\eta_0) \quad \text{in } L^2(\Omega_\epsilon). \tag{87}$$

By elementary calculations, one verifies that

$$\Psi_\epsilon(\eta_n) \geq \Psi_\epsilon(\eta_0) + 2 \int_{\Omega_\epsilon} \phi\big(\mathrm{D}_x H(\eta_n) - \mathrm{D}_x H(\eta_0)\big) \mathrm{D}_x H(\eta_0) \, dx.$$

Using (87), the limit $n \to \infty$ provides

$$\Psi_\epsilon(\eta_0) \leq \liminf_{n \to \infty} \Psi_\epsilon(\eta_n) \leq \bar{\Psi}.$$

To conclude the proof, observe that $\Psi_\epsilon(\eta_0) \to \Psi(\eta_0)$ in the limit $\epsilon \downarrow 0$ by the monotone convergence theorem. $\square$

ACKNOWLEDGMENTS

REFERENCES

[1] L. Ambrosio, N. Gigli, and G. Savaré. *Gradient flows in metric spaces and in the space of probability measures. 2nd ed.* Lectures in Mathematics, ETH Zürich. Basel: Birkhäuser., 2008.

[2] Y. Bolley, F. ad Brenier and G. Loeper. Contractive metrics for scalar conservation laws. *Journal of Hyperbolic Differential Equations*, 2:91–107, 2005.

[3] J. Carrillo. Entropy solutions for nonlinear degenerate problems. *Arch. Ration. Mech. Anal.*, 147(4):269–361, 1999.

[4] J. A. Carrillo, M. Di Francesco, and C. Lattanzio. Contractivity of wasserstein metrics and asymptotic profiles for scalar conservation laws. *Journal of Differential Equations*, 231:425–458, 2006.

[5] G. Crandall. The semigroup approach to first–order quasilinear equations in several space variables. *Israel J. Math.*, 12:108–132, 1972.

[6] C. Dafermos. The entropy rate admissibility criterion for solutions of hyperbolic conservation laws. *J. Diff. Equations*, 14:202–212, 1973.

[7] C. M. Dafermos. Polygonal approximations of solutions of the initial value problem for a conservation law. *J. Math. Anal. Appl.*, 38:33–41, 1972.

[8] C. M. Dafermos. *Hyperbolic conservation laws in continuum physics*, volume 325 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 2000.

[9] S. Daneri and G. Savaré. Eulerian calculus for the displacement convexity in the Wasserstein distance. *SIAM J. Math. Anal.*, 40(3):1104–1122, 2008.

[10] E. De Giorgi. New problems on minimizing movements. In *Boundary value problems for partial differential equations and applications*, volume 29 of *RMA Res. Notes Appl. Math.*, pages 81–98. Masson, Paris, 1993.

[11] Y. Du. *Order structure and topological methods in nonlinear partial differential equations. Vol. 1. Maximum principles and applications.* Series in Partial Differential Equations and Applications 2. Hackensack, NJ: World Scientific., 2006.

[12] N. Dunford and J. T. Schwartz. *Linear operators. Part I.* Wiley Classics Library. John Wiley & Sons Inc., New York, 1988. General theory, With the assistance of William G. Bade and Robert G. Bartle, Reprint of the 1958 original, A Wiley-Interscience Publication.

[13] S. N. Ethier and T. G. Kurtz. *Markov processes.* Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons Inc., New York, 1986. Characterization and convergence.

[14] F. G. Friedlander. *Introduction to the theory of distributions.* Cambridge University Press, Cambridge, 1982.

[15] N. Gigli and F. Otto. Entropic burgers' equation via a minimizing movement scheme based on the wasserstein metric. *Calculus of Variations and Partial Differential Equations*, Online First:1–26, 2012.

[16] R. Jordan, D. Kinderlehrer, and F. Otto. The variational formulation of the Fokker-Planck equation. *SIAM J. Math. Anal.*, 29(1):1–17, 1998.

[17] K. H. Karlsen and M. Ohlberger. A note on the uniqueness of entropy solutions of nonlinear degenerate parabolic equations. *J. Math. Anal. Appl.*, 275:439–458, 2002.

[18] K. H. Karlsen and N. H. Risebro. On the uniqueness and stability of entropy solutions of nonlinear degenerate parabolic equations with rough coefficients. *Discrete Contin. Dyn. Syst.*, 9(5):1081–1104, 2003.

[19] S. N. Kružkov. First order quasilinear equations in serveral independent variables. *Math. USSR Sb*, 10:217–243, 1970.

[20] S. N. Kružkov. Generalized solutions of the hamilton-jacobi equations of eikonal type. i. formulation of the problems; existence, uniqueness and stability theorems; some properties of the solutions. *Math. USSR Sb*, 27, 1975.

[21] E. H. Lieb and M. Loss. *Analysis.* Graduate Studies in Mathematics, 14 American Mathematical Society, Providence, RI, 1997.

[22] D. Matthes, R. J. McCann, and G. Savaré. A family of nonlinear fourth order equations of gradient flow type. *Comm. Partial Differential Equations*, 34(10-12):1352–1397, 2009.

[23] F. Otto. l^1-contraction and uniqueness for unstationary saturated-unsaturated porous media flow. *Adv. Math. Sci. Appl.*, 7:537–553, 1997.

[24] F. Otto. Evolution of microstructure in unstable porous media flow: a relaxational approach. *Comm. Pure Appl. Math.*, 52(7):873–915, 1999.

[25] F. Otto and M. Westdickenberg. Eulerian calculus for the contraction in the Wasserstein distance. *SIAM J. Math. Anal.*, 37(4):1227–1255 (electronic), 2005.

[26] R. Rossi and G. Savaré. Tightness, integral equicontinuity and compactness for evolution problems in Banach spaces. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)*, 2(2):395–431, 2003.

[27] R. E. Showalter. *Monotone operators in Banach space and nonlinear partial differential equations*, volume 49 of *Mathematical Surveys and Monographs.* American Mathematical Society, Providence, RI, 1997.

[28] C. Villani. *Topics in optimal transportation*, volume 58 of *Graduate Studies in Mathematics.* American Mathematical Society, Providence, RI, 2003.

[29] C. Villani. *Optimal transport. Old and new.* Grundlehren der Mathematischen Wissenschaften 338. Berlin: Springer. xxii, 973 p., 2009.

Marco Di Francesco, Department of Mathematical Sciences, University of Bath, Claverton Down, BA2 7AY, Bath (UK)
    E-mail address: m.difrancesco@bath.ac.uk

Daniel Matthes, Zentrum Mathematik, Technische Universität München, D-85747 Garching bei München (Germany)
    E-mail address: matthes@ma.tum.de