LONG-TIME BEHAVIOR OF A FULLY DISCRETE
LAGRANGIAN SCHEME FOR A FAMILY OF FOURTH
ORDER EQUATIONS

HORST OSBERGER
Zentrum Mathematik
TU München Boltzmannstr. 3
D-85748 Garching, Germany

(Communicated by José A. Carrillo)

Abstract. A fully discrete Lagrangian scheme for solving a family of fourth
order equations numerically is presented. The discretization is based on the
equations’ underlying gradient flow structure with respect to the Wasserstein
metric, and preserves numerous of their most important structural properties
by construction, like conservation of mass and entropy-dissipation.

In this paper, the long-time behavior of our discretization is analysed: We
show that discrete solutions decay exponentially to equilibrium at the same
rate as smooth solutions of the original problem. Moreover, we give a proof
of convergence of discrete entropy minimizers towards Barenblatt-profiles or
Gaussians, respectively, using Γ-convergence.

1. Introduction. In this paper, we propose and study a fully discrete numerical
scheme for a family of nonlinear fourth order equations of the type
\[ \partial_t u = -\left(u^{\alpha-1} u_x^\alpha x + \lambda xu_x\right)_x \quad \text{for } x \in \mathbb{R}, t > 0 \]
and \( u(0,.) = u^0 \) on \( \mathbb{R} \) at initial time \( t = 0 \). The initial density \( u^0 \geq 0 \) is assumed
to be compactly supported and integrable with total mass \( M \geq 0 \), and we further
require strict positivity of \( u^0 \) on \( \text{supp}(u^0) = [a,b] \). For the sake of simplicity, let us
further assume that \( M = 1 \). We are especially interested in the long-time behavior
of discrete solutions and their rate of decay towards equilibrium. For the exponent
in \( \alpha \), we consider values \( \alpha \in \left[\frac{1}{2},1\right] \), and assume \( \lambda \geq 0 \). The most famous examples
for parabolic equations described by (1) are the so-called DLSS equation for \( \alpha = \frac{1}{2} \),
(first analysed by Derrida, Lebowitz, Speer and Spohn in [23, 24] with application in
semi-conductor physics) and the thin-film equation for \( \alpha = 1 \) — indeed, references
are very rare in the literature for other values of \( \alpha \), except [44] of Matthes, McCann
and Savar.

Due to the physically motivated origin of (1) (especially for \( \alpha = \frac{1}{2} \) and \( \alpha = 1 \)),
it is not surprising that solutions to (1) carry many structural properties as for
instance nonnegativity, the conservation of mass and the dissipation of (several)
entropy functionals. In Section 2 we are going to list more properties of solutions
to (1). For the numerical approximation of solutions to (1), it is hence natural to ask for structure-preserving discretizations that inherit at least some of those properties. A minimum criteria for such a scheme should be the preservation of nonnegativity, which can already be a difficult task, if standard discretizations are used. So far, many (semi-)discretizations have been proposed in the literature, and most of them keep some basic structural properties of the equation’s underlying nature. Take for example [10, 16, 39, 41], where positivity appears as a conclusion of Lyaponov functionals—a logarithmic/power entropy [10, 16, 39] or some variant of a (perturbed) information functional. But there is only a small number of examples, where structural properties of (1) are adopted from the discretization by construction. A first approach in this direction was a fully Lagrangian discretization for the DLSS equation by Dring, Matthes and Milišić [25], which is based on its $L^2$-Wasserstein gradient flow representation and thus preserves nonnegativity and dissipation of the Fisher-information. A similar approach was then applied in [46], again for the special case $\alpha = \frac{1}{2}$, where we even showed convergence of our numerical scheme, which was—as far as we know—the first convergence proof of a fully discrete numerical scheme for the DLSS equation, which additionally dissipates two Lyapunov functionals.

1.1. Description of the numerical scheme. We are now going to present a scheme, which is practical, stable and easy to implement. In fact our discretization seems to be so mundane that one would not assume any special properties therein, at first glance. But we are going to show later in Section 2 that our numerical approximation can be derived as a natural restriction of a $L^2$-Wasserstein gradient flow in the potential landscape of the so-called perturbed information functional

$$\mathcal{F}_{\alpha,\lambda}(u) = \frac{1}{2\alpha} \int_{\mathbb{R}} \left( \partial_x u^\alpha \right)^2 \, dx + \frac{\lambda}{2} \int_{\mathbb{R}} |x|^2 u(x) \, dx \quad \text{(2)}$$

into a discrete Lagrangian setting, thus preserves a deep structure. The starting point for our discretization is the Lagrangian representation of (1). Since each $u(t, \cdot)$ is of mass $M$, there is a Lagrangian map $X(t, \cdot) : [0, M] \rightarrow \mathbb{R}$ — the so-called pseudo-inverse distribution function of $u(t, \cdot)$ — such that

$$\xi = \int_{-\infty}^{X(t,\xi)} u(t, x) \, dx \quad \text{for each } \xi \in [0, M]. \quad \text{(3)}$$

Written in terms of $X$, the Wasserstein gradient flow for $\mathcal{F}_{\alpha,\lambda}$ turns into an $L^2$-gradient flow for

$$\mathcal{F}_{\alpha,\lambda}(u \circ X) = \frac{1}{2\alpha} \int_0^M \left[ \frac{1}{X_\xi^\alpha} \right]^2 \frac{1}{X_\xi} \, d\xi + \frac{\lambda}{2} \int_0^M X^2 \, d\xi,$$

that is

$$\partial_t X = \frac{2\alpha}{(2\alpha + 1)^2} \partial_t \left(Z^{\alpha + \frac{1}{2}} \partial_\xi Z^{\alpha + \frac{1}{2}} \right) + \lambda X,$$

where

$$Z(t, \xi) := \frac{1}{\partial_\xi X(t,\xi)} = u(t, X(t,\xi)). \quad \text{(4)}$$

To build a bridge from (4) to the original (1), remember that (1) can be written as a transport equation,
\[
\partial_t u + (u v_\alpha)_x = 0 \quad \text{with velocity field } \quad v_\alpha = -\left( \frac{\delta \mathcal{F}_{\alpha,\lambda}(u)}{\delta u} \right)_x, \quad (5)
\]

where \(\delta \mathcal{F}_{\alpha,\lambda}(u)/\delta u\) denotes the Eulerian first variation. So take the time derivative in (3) and use (5), then a formal calculation yields

\[
0 = \partial_t X(t, \xi) u(t, X(t, \xi)) + \int_{-\infty}^{X(t, \xi)} \partial_t u(t, X(t, \xi)) \, dx \\
= \partial_t X(t, \xi) u(t, X(t, \xi)) - \int_{-\infty}^{X(t, \xi)} (u v_\alpha)_x(t, x) \, dx \\
= \partial_t X(t, \xi) u(t, X(t, \xi)) - (u v_\alpha) \circ X(t, \xi).
\]

This is equivalent to

\[
\partial_t X(t, \xi) = v_\alpha \circ X(t, \xi) \quad \text{for } (t, \xi) \in (0, +\infty) \times [0, M],
\]

which is further equivalent to (4).

Before we come to the proper definition of the numerical scheme, we fix a spatio-temporal discretization parameter \(\Delta = (\tau, \delta)\): Given \(\tau > 0\), introduce varying time step sizes \(\tau = (\tau_1, \tau_2, \ldots)\) with \(\tau_n \in (0, \tau]\), then a time decomposition of \([0, +\infty)\) is defined by \(\{t_n\}_{n \in \mathbb{N}}\) with \(t_n := \sum_{j=1}^{n} \tau_n\). As spatial discretization, fix \(K \in \mathbb{N}\) and \(\delta = M/K\), and declare an equidistant decomposition of the mass space \([0, M]\) through the set \(\{\xi_k\}_{k=0}^{K}\) with \(\xi_k := k\delta, k = 0, \ldots, K\).

Our numerical scheme is now defined as a standard discretization of (4):

**Numerical scheme.** Fix a discretization parameter \(\Delta = (\tau, \delta)\). Then for any \((\alpha, \lambda) \in [\frac{1}{2}, 1] \times [0, +\infty)\) and any initial density function \(u^0 \in L^1(\mathbb{R})\) satisfying the above requirements, a numerical scheme for (1) is recursively given as follows:

1. For \(n = 0\), define an initial sequence of monotone values \(x^0_{\Delta} := (x^0_0, \ldots, x^0_K) \in \mathbb{R}^{K+1}\) uniquely by \(x^0_0 = a, x^0_K = b\), and

\[
\xi_k = \int_{x^0_{k-1}}^{x^0_k} u^0(x) \, dx \quad \text{for any } k = 1, \ldots, K - 1.
\]

The vector \(x^0_{\Delta}\) describes a non-equidistant decomposition of the support \([a, b]\) of the initial density function \(u^0\). In any interval \([x^0_{k-1}, x^0_k]\), \(k = 1, \ldots, K\), the density \(u^0\) has mass \(\delta\).

2. For \(n \geq 1\), define recursively a monotone vector \(x^n_{\Delta} := (x^n_0, \ldots, x^n_K) \in \mathbb{R}^{K+1}\) as a solution of the system, consisting of \((K + 1)\)-many equations

\[
\begin{align*}
\frac{x^n_k - x^n_{k-1}}{\tau_n} &= \frac{2\alpha}{(2\alpha + 1)^2 \delta} \left[ (z^n_{k+\frac{1}{2}})^{\alpha + \frac{1}{2}} [D^2 \tilde{z}^{\alpha + \frac{1}{2}}]_{k+\frac{1}{2}} - (z^n_{k-\frac{1}{2}})^{\alpha + \frac{1}{2}} [D^2 \tilde{z}^{\alpha + \frac{1}{2}}]_{k-\frac{1}{2}} \right] \\
&\quad + \lambda x_k,
\end{align*}
\]

with \(k = 0, \ldots, K\), where the values \(z^n_{\ell-\frac{1}{2}} \geq 0\) are defined by

\[
z^n_{\ell-\frac{1}{2}} = \begin{cases} \frac{\delta}{x^n_{\ell-1} - x^n_{\ell-2}}, & \ell = 1, \ldots, K, \\
0, & \text{else}, \end{cases}
\]

and \([D^2 \tilde{z}^{\alpha + \frac{1}{2}}]_{k-\frac{1}{2}} := \delta^{-2} (z^n_{k+\frac{1}{2}} - 2z^n_{k-\frac{1}{2}} + z^n_{k-\frac{1}{2}})\).
We later show in Proposition 1 that the solvability of the system (6) is guaranteed.

The above procedure (1) – (2) yields a sequence of monotone vectors
\[
\tilde{x}_\Delta := (\tilde{x}_{\Delta}^0, \tilde{x}_{\Delta}^1, \ldots),
\]
and any entry $\tilde{x}_\Delta^n$ defines a spatial decomposition of the compact interval $[x_0^n, x_k^n] \subset \mathbb{R}, n \in \mathbb{N}$. Fixing $k = 0, \ldots, K$, the sequence $n \mapsto x_k^n$ defines a discrete temporal evolution of spatial grid points in $\mathbb{R}$, and if one assigns each interval $[x_{k-1}, x_k^n]$ a constant mass package $\delta$, the map $n \mapsto [x_{k-1}, x_k^n]$ characterizes the temporal movement of mass. Hence $\tilde{x}_\Delta$ is uniquely related to a sequence of local constant density functions $u_\Delta := (u_\Delta^0, u_\Delta^1, \ldots, u_\Delta^K)$, where each function $u_\Delta^k : \mathbb{R} \rightarrow \mathbb{R}_+$ satisfies
\[
u_\Delta^k(x) = u_\delta[\tilde{x}_\Delta^n] := \sum_{k=1}^K \frac{\delta}{x_k^n - x_{k-1}^n} \mathbb{I}_{[x_{k-1}, x_k]}(x). \tag{8}
\]

We will see later in Section 2.1 that the information functional $F_{\alpha, \lambda}$ can be derived using the dissipation of the entropy
\[
\mathcal{H}_{\alpha, \lambda}(\nu) = \int_{\mathbb{R}} \varphi_{\alpha}(\nu) \, dx + \frac{\Lambda_{\alpha, \lambda}}{2} \int_{\mathbb{R}} |\nu|^2 \varphi_{\alpha}(\nu) \, dx
\]
with
\[
\varphi_{\alpha}(s) := \begin{cases}
\Theta_{\alpha}^\alpha s^{\alpha+1/2} & , \alpha \in (1/2, 1], \\
\Theta_{1/2} s \ln(s) & , \alpha = 1/2,
\end{cases}
\]
where
\[
\Theta_{\alpha} := \sqrt{2\alpha/(2\alpha + 1)}, \quad \text{and} \quad \Lambda_{\alpha, \lambda} := \lambda/(2\alpha + 1). \tag{9}
\]
To discretize the entropy $\mathcal{H}_{\alpha, \lambda}$ and the perturbed information functional $F_{\alpha, \lambda}$, we introduce
\[
\mathbf{H}_{\alpha, \lambda}(\tilde{x}) := \delta \sum_{k=1}^K f_\alpha(z_{k-1/2}) + \frac{\Lambda_{\alpha, \lambda}}{2} \sum_{k=0}^K |x_k|^2
\]
with
\[
f_\alpha(s) := \begin{cases}
\Theta_{\alpha}^\alpha s^{\alpha-1/2} & , \alpha \in (0, 1/2], \\
\Theta_{1/2} s \ln(s) & , \alpha = 1/2,
\end{cases}
\]
and
\[
\mathbf{F}_{\alpha, \lambda}(\tilde{x}) := \Theta_{\alpha}^2 \delta \sum_{k=0}^K \left(\frac{z_{k+1/2}^{\alpha+\frac{1}{2}} - z_{k-1/2}^{\alpha+\frac{1}{2}}}{\delta}\right)^2 + \frac{\lambda}{2} \sum_{k=0}^K |x_k|^2. \tag{12}
\]

1.2. Related schemes. The construction of numerical schemes as a solution of discrete Wasserstein gradient flows with Lagrangian representation is not new in the literature. Many approaches in this spirit have been realised for second-order diffusion equation [9, 11, 13, 48], but also for chemotaxis systems [6], for non-local aggregation equations [17, 19], and for variants of the Boltzmann equation [32]. We further refer to [42] to the reader interested in a very general numerical treatment of Wasserstein gradient flows. In case of fourth order equations, there are some results for the thin-film equation and its more general version, the Hele-Shaw flow, see [20, 32], but convergence results are missing. Rigorous stability and convergence results for fully discrete schemes are rare and can just be found in [41, 45] for second order equations, and in [60] for the DLSS equation. However, there are results available for semi-discrete Lagrangian approximations, see e.g. [2, 20].
1.3. Main results. In this section, fix a discretization $\Delta = (\tau; \delta)$ with $\tau, \delta > 0$. For any solution $\bar{x}_\Delta$ of (11), we will further denote by $u_\Delta = (u_\Delta^0, u_\Delta^1, \ldots)$ the corresponding sequence of local constant density functions, as defined in (8).

All analytical results that will follow arise from the very fundamental observation that solutions to the scheme defined in Section 1.1 can be successively derived as solutions to the discrete minimizing movement scheme: For fixed $\bar{x}_\Delta^0$ and $n \geq 1$, define $\bar{x}_\Delta^n$ recursively as the minimizer of the functional

$$\bar{x} \mapsto \frac{\delta}{2\tau_n} \sum_{k=0}^{K} (x_k - x_{k-1})^2 + F_{\alpha,\lambda}(x).$$

An immediate consequence of the minimization procedure is that solutions $\bar{x}_\Delta^n$ dissipate the functional $F_{\alpha,\lambda}$.

Concerning the long-time behavior of solutions $\bar{x}_\Delta$, remarkable similarities to the continuous case appear. Assuming first the case $\lambda > 0$, it turns out that the unique minimizer $x_\delta$ of $H_{\alpha,\lambda}$ is even a minimizer of the discrete information functional $F_{\alpha,\lambda}$, and the corresponding set of density functions $u_\delta^{\min} = u_\delta(x_\delta)$ converges for $\delta \to 0$ towards a Barenblatt-profile $b_{\alpha,\lambda}$ or Gaussian $b_{1/2,\lambda}$, respectively, that is defined by

$$b_{\alpha,\lambda} = (a - b|x|^2)^{1/(\alpha - 1/2)} \quad b = \frac{\alpha - 1/2}{\sqrt{2\alpha}} \Lambda_{\alpha,\lambda} \quad \text{if } \alpha > 1/2 \quad \text{and}$$

$$b_{1/2,\lambda} = ae^{-\Lambda_{1/2,\lambda}|x|^2} \quad \text{if } \alpha = 1/2,$$

where $a \in \mathbb{R}$ is chosen to conserve unit mass and $\Lambda_{\alpha,\lambda}$ is defined as in (10). Beyond this, solutions $\bar{x}_\Delta^n$ satisfying (11) converge as $n \to \infty$ towards a minimizer $\hat{x}_\delta$ of $F_{\alpha,\lambda}$ with an exponential decay rate, which is “asymptotically equal” to the one obtained in the continuous case. The above claims are summarized in the following theorems:

**Theorem 1.1.** Assume $\lambda > 0$. Then the sequence of minimizers $u_\delta^{\min}$ satisfies

$$u_\delta^{\min} \xrightarrow{\delta \to 0} b_{\alpha,\lambda} \text{ strongly in } L^p(\mathbb{R}) \text{ for any } p \geq 1,$$

$$\hat{u}_\delta^{\min} \xrightarrow{\delta \to 0} b_{\alpha,\lambda} \text{ uniformly on } \mathbb{R},$$

where $\hat{u}_\delta^{\min}$ is a locally affine interpolation of $u_\delta^{\min}$ defined in Lemma 3.2.

**Theorem 1.2.** For $\lambda > 0$, any sequence of monotone vectors $\bar{x}_\Delta$ satisfying (13) dissipates the entropies $H_{\alpha,\lambda}$ and $F_{\alpha,\lambda}$ at least exponentially, i.e.

$$H_{\alpha,\lambda}(\bar{x}_\Delta^n) - H_{\alpha,\lambda}(\bar{x}_\Delta^0) \leq (H_{\alpha,\lambda}(\bar{x}_\Delta^0) - H_{\alpha,\lambda}(\bar{x}_\Delta^0)) e^{-\frac{2\lambda}{1+\tau}} t_n, \quad \text{and}$$

$$F_{\alpha,\lambda}(\bar{x}_\Delta^n) - F_{\alpha,\lambda}(\bar{x}_\Delta^0) \leq (F_{\alpha,\lambda}(\bar{x}_\Delta^0) - F_{\alpha,\lambda}(\bar{x}_\Delta^0)) e^{-\frac{2\lambda}{1+\tau}} t_n,$$

where $H_{\alpha,\lambda}^{\min} = H_{\alpha,\lambda}(\bar{x}_\Delta^{\min})$ and $F_{\alpha,\lambda}^{\min} = F_{\alpha,\lambda}(\bar{x}_\Delta^{\min})$. The associated sequence of densities $u_\Delta$ further fulfills

$$\|u_\Delta^n - u_\Delta^{\min}\|_{L^1(\mathbb{R})} \leq c_{\alpha,\lambda} (H_{\alpha,\lambda}(\bar{x}_\Delta^0) - H_{\alpha,\lambda}^{\min}) e^{-\frac{2\lambda}{1+\tau}} t_n,$$

for any time step $n \in \mathbb{N}$, where $c_{\alpha,\lambda} > 0$ depends only on $\alpha, \lambda$.

Let us now consider the zero-confinement case $\lambda = 0$. In the continuous setting, the long-time behavior of solutions to (12) with $\lambda = 0$ can be studied by a rescaling of solutions to (11) with $\lambda > 0$. We are able to translate this method into the discrete case and derive a discrete counterpart of [44 Corollary 5.5], which describes the
intermediate asymptotics of solutions that approach self-similar Barenblatt profiles as $t \to \infty$.

**Theorem 1.3.** Assume $\lambda = 0$ and take a sequence of monotone $\Delta^n$ satisfying 13. Then there exists a constant $c_\alpha > 0$ depending only on $\alpha$, such that

$$
\|u_n^\alpha - b_n^\alpha,0\|_{L^1(\mathbb{R})} \leq c_\alpha \sqrt{H_{\alpha,0}(\Delta^n)} - H_{\alpha,0}(R_n^\alpha)^{-1}
$$

with $R_n^\alpha := (1 + a_\tau(2\alpha + 3)t_n)^{-\frac{1}{b_\tau(2\alpha + 3)}}$, where $b_n^\alpha,0$ is a rescaled discrete Barenblatt profile and $a_\tau, b_\tau > 0$, such that $a_\tau, b_\tau \to 1$ for $\tau \to 0$, see Section 3.2 for more details.

Before we come to the analytical part of this paper, we want to point out the following: The ideas for the proofs of Theorem 1.2 and 1.3 are mainly guided by the techniques developed in [44]. The remarkable observation of this work is the fascinating structure preservation of our discretization, which allows us to adapt almost all calculations from the continuous theory for the discrete setting.

1.4. **Structure of paper.** In the following Section 2, we point out some of the main structural features of (1) and the functionals $H_{\alpha,\lambda}$ and $F_{\alpha,\lambda}$, and show that our scheme arises from a discrete $L^2$-Wasserstein gradient flow, so that many properties of the continuous flow are inherited. Section 3 treats the analysis of discrete equilibria in case of positive confinement $\lambda > 0$: We prove convergence of discrete stationary states to Barenblatt-profiles or Gaussians, respectively, and analyse the asymptotics of discrete solutions for $\lambda = 0$. Finally, some numerical experiments are presented in Section 4.

2. **Structural properties — continuous vs. discrete case.**

2.1. **Structural properties of (1).** The family of fourth order equations (1) carries a bunch of remarkable structural properties. The most fundamental one is the conservation of mass, which is a naturally given property, if one interprets solutions to (1) as a $L^2$-Wasserstein gradient flow of the perturbed information functional $F_{\alpha,\lambda}$ in (2). As an immediate consequence, $F_{\alpha,\lambda}$ is a Lyapunov functional, and one can find infinitely many other (formal) Lyapunov functionals at least for special choices of $\alpha$ — see [2] [12] [36] for $\alpha = \frac{1}{2}$ or [3] [18] [28] for $\alpha = 1$. Apart from $F_{\alpha,\lambda}$, one of the most important such Lyapunov functionals is given by the $\Lambda_{\alpha,\lambda}$-convex entropy $H_{\alpha,\lambda}$. It turns out that the functionals $F_{\alpha,\lambda}$ and $H_{\alpha,\lambda}$ are not just Lyapunov functionals, but share numerous remarkable similarities. One can indeed see (1) as a higher order extension of the second order porous media/heat equation

$$
\partial_t v = -\nabla W_2 H_{\alpha,\lambda}(v) = -\Theta_\alpha \partial_{xx}(v^\alpha) + \Lambda_{\alpha,\lambda}(xu)_x,
$$

which is nothing less than the $L^2$-Wasserstein gradient flow of $H_{\alpha,\lambda}$. Furthermore, the unperturbed functional $F_{\alpha,0}$, i.e. $\lambda = \Lambda_{\alpha,\lambda} = 0$, equals the dissipation of $H_{\alpha,0}$ along its own gradient flow,

$$
F_{\alpha,0}(v(s)) = -\frac{d}{ds}H_{\alpha,0}(v(s)).
$$

In view of the gradient flow structure, this relation makes (1) the “big brother” of the porous media/heat equation (21), see [22] [44] for structural consequences. Another astonishing common feature is the correlation of $F_{\alpha,\lambda}$ and $H_{\alpha,\lambda}$ by the so-called
fundamental entropy-information relation: For any $u \in \mathcal{P}(\mathbb{R})$ with $\mathcal{H}_{\alpha,\lambda}(u) < \infty$, one has that
\[
\mathcal{F}_{\alpha,\lambda}(u) = |\text{grad}_{W_2} \mathcal{H}_{\alpha,\lambda}|^2 + (2\alpha - 1)\Lambda_{\alpha,\lambda}\mathcal{H}_{\alpha,\lambda}(u)
\] for any $\lambda \geq 0$, (23) see [44, Corollary 2.3]. This equation is a crucial tool for the analysis of equilibria of both functionals and the corresponding long-time behavior of solutions to (1) and (21).

In addition to the above listing, a typical property of diffusion processes like (1) or (21) with positive confinement $\lambda, \Lambda_{\alpha,\lambda} > 0$ is the convergence towards unique stationary solutions $u^\infty$ and $v^\infty$, respectively, independent of the choice of initial data. It is maybe one of the most surprising facts that both equations, (1) and (21), share the same steady state, i.e. the stationary solutions $u^\infty$ and $v^\infty$ are identical. Those stationary states are solutions of the elliptic equations
\[
-(P_{\alpha}(u))_{xx} + \Lambda_{\alpha,\lambda}(xu)_x = 0,
\] with $P_{\alpha}(s) := \Theta_{\alpha}s^{\alpha+1/2}$, and have the form of Barenblatt profiles or Gaussians, respectively, see definition (14) and (15). This was first observed by Denzler and McCann in [22], and further studied in [44] using the Wasserstein gradient flow structure of both equations and their remarkable relation via (22).

In case of $\alpha \in \{\frac{1}{2}, 1\}$, the mathematical literature is full of numerous results, which is because of the physical importance of (1) in those limiting cases.

2.1.1. DLSS equation. As already mentioned at the very beginning, the DLSS equation — (1) with $\alpha = \frac{1}{2}$ — arises from the Toom model [23, 24] in one spatial dimension on the half-line $[0, +\infty)$, and was used to describe interface fluctuations, therein. Moreover, the DLSS equation also finds application in semi-conductor physics, namely as a simplified model (low-temperature, field-free) for a quantum drift diffusion system for electron densities, see [38].

From the analytical point of view, a big variety of results in different settings has been developed in the last view decades. For results on existence and uniqueness, we refer e.g. to [27, 28, 34, 29, 37, 38], and [12, 18, 14, 29, 37, 40, 44] for qualitative and quantitative descriptions of the long-time behavior. The main reason, which makes the research on this topic so non-trivial, is a lack of comparison/maximum principles as in the theory of second order equations (21). And, unfortunately, the absence of such analytical tools is not neglectable, as the work [7] of Bleher et.al shows: As soon as a solution $u$ of (1) with $\alpha = \frac{1}{2}$ is strictly positive, one can show that it is even $C^\infty$-smooth, but there are no regularity results available from the moment when $u$ touches zero. The problem of strictly positivity of such solutions seems to be a difficult task, since it is still open. This is why alternative theories for nonnegative weak solutions have more and more become matters of great interest, as e.g. an approach based on entropy methods developed in [29, 37].

2.1.2. Thin-film equation. The thin-film equation — (1) with $\alpha = 1$ — is of similar physical importance as the DLSS equation, since it gives a dimension-reduced description of the free-surface problem with the Navier-Stokes equation in the case of laminar flow, [47]. In case of linear mobility — which is exactly the case in our situation — the thin-film equation can also be used to describe the pinching of thin necks in a Hele-Shaw cell in one spatial dimension, and thus plays an extraordinary role in physical applications. To this topic, the literature provides some interesting
results in the framework of entropy methods, see [13, 18, 28]. In the (more general) case of nonnegative mobility functions \( m \), i.e.

\[
\partial_t u = - \div (m(u) \nabla u),
\]

one of the first achievements to this topic available in the mathematical literature was done by Bernis and Friedman [4]. The same equation is observed in [5], treating a vast number of results to numerous mobility functions of physical meaning. There are several other references in this direction, e.g. Grn et. al [3, 21, 33], concerning long-time behavior of solutions and the non-trivial question of spreading behavior of the support.

2.2. Structure-preservation of the numerical scheme. In this section, we try to get a better intuition of the scheme in Section 1.1. Foremost we will derive (6) as a discrete system of Euler-Lagrange equations of a variational problem that arises to get a better intuition of the scheme in Section 1.1. Foremost we will derive (6) as a discrete system of Euler-Lagrange equations of a variational problem that arises from a \( L^2 \)-Wasserstein gradient flow restricted on a discrete submanifold \( P_\delta(\mathbb{R}) \) of the space of probability measures \( P(\mathbb{R}) \) on \( \mathbb{R} \). This is why the numerical scheme satisfies several discrete analogues of the results discussed in the previous section. As the following section shows, some of the inherited properties are obtained by construction (e.g. preservation of mass and dissipation of the entropy), where others are caused by the underlying discrete gradient flow structure and a smart choice of a discrete \( L^2 \)-Wasserstein metric. Moreover, it is possible to prove that the entropy and the information functional share the same minimizer \( x^\text{min} \) even in the discrete case, and solutions of the discrete gradient flow converges with an exponential rate to this stationary state. The prove of this observation is more sophisticated, that is why we dedicate an own section (Section 3) to the treatment of this special property.

2.2.1. Ansatz space and discrete entropy/information functionals. The entropies \( H_{\alpha,\lambda} \) and \( F_{\alpha,\lambda} \) as defined in (4) and (2) are nonnegative functionals on \( P(\mathbb{R}) \). If we first consider the zero-confinement case \( \lambda = 0 \), one can derive in analogy to [15] the discretization in (11) of \( H_{\alpha,0} \) just by restriction to a finite-dimensional submanifold \( P_\delta(\mathbb{R}) \) of \( P(\mathbb{R}) \): For fixed \( K \in \mathbb{N} \), the set \( P_\delta(\mathbb{R}) \) consists of all local constant density functions \( u = u_\delta(x) \) (remember definition (5)), such that \( \vec{x} \in \mathbb{R}^{K+1} \) is a monotone vector, i.e.

\[
\vec{x} \in \mathbb{R}^\delta := \{(x_0, \ldots, x_K) \mid x_0 < x_1 < \ldots < x_{K-1} < x_K\} \subseteq \mathbb{R}^{K+1}.
\]

Such density functions \( u = u_\delta(x) \in P_\delta(\mathbb{R}) \) bear a one-to-one relation to their Lagrangians or Lagrangian maps, which are defined on the mass grid \([0, M]\) with uniform decomposition \((0 = \xi_0, \ldots, \xi_k, \ldots, \xi_K = M)\). More precisely, we define for \( \vec{x} \in \mathbb{R}^\delta \) the local affine and monotonically increasing function \( X = X_\delta(\vec{x}) : [0, M] \to \mathbb{R} \), such that \( X(\xi_k) = x_k \) for any \( k = 0, \ldots, K \). One then attains \( u \circ X = \frac{1}{\vec{x}_\delta} \) for \( u \in P_\delta(\mathbb{R}) \) and its corresponding Lagrangian map. For later analysis, we introduce in addition to the decomposition \( \{\xi_k\}_{k=0}^K \) the intermediate values \( (\xi_{k-\frac{1}{2}}, \xi_{\frac{1}{2}}, \ldots, \xi_{K-\frac{1}{2}}) \) by \( \xi_{k-\frac{1}{2}} = \frac{1}{2}(\xi_k + \xi_{k-1}) \) for \( k = 1, \ldots, K \).

In view of the entropy’s discretization, this implies

\[
H_{\alpha,0}(\vec{x}) = H_{\alpha,0}(u_\delta(\vec{x})) = \int_{\mathbb{R}} \varphi_\alpha(u_\delta(\vec{x})) \, dx = \delta \sum_{k=1}^{K} f_\alpha(z_{k-\frac{1}{2}}),
\]

using (11) and (9), a change of variables \( x = X_\delta(\vec{x}) \), and the definition (7) of the \( \vec{x} \)-dependent vectors \( \vec{z} \). This is perfectly compatible with (11). Obviously, one cannot derive the discrete information functional \( F_{\alpha,0} \) in the same way, since \( F_{\alpha,0} \) is not
defined on \( P_\delta(\mathbb{R}) \). So instead of restriction, we mimic property (22) that is for any \( \vec{x} \in \mathfrak{r}_\delta \)
\[
F_{\alpha,0}(\vec{x}) = \delta^{-1} \partial_{\vec{x}} H_{\alpha,0}(\vec{x})^T \partial_{\vec{x}} H_{\alpha,0}(\vec{x}) = \langle \nabla_\delta H_{\alpha,0}(\vec{x}), \nabla_\delta H_{\alpha,0}(\vec{x}) \rangle_\delta.
\]

Here, the \( k \)th component of \( \partial_{\vec{x}} f(\vec{x}) \) satisfies \( \partial_{\vec{x}} f(\vec{x})_k = \partial_{x_k} f(\vec{x}) \) for any \( k = 0, \ldots, K \) and arbitrary function \( f : \mathfrak{r}_\delta \rightarrow \mathbb{R} \). Moreover, we set \( \nabla_\delta f(\vec{x}) = \delta^{-1} \partial_{\vec{x}} f(\vec{x}) \), and introduce for \( \vec{v}, \vec{w} \in \mathbb{R}^{K+1} \) the scalar product \( \langle \cdot, \cdot \rangle_\delta \) by
\[
\langle \vec{v}, \vec{w} \rangle_\delta = \delta \sum_{k=0}^{K} v_k w_k \quad \text{with induced norm} \quad \| \vec{v} \|_\delta = \sqrt{\langle \vec{v}, \vec{v} \rangle_\delta}.
\]

**Example 2.1.** Each component \( z_\kappa \) of \( \vec{z} = \mathfrak{z}_\delta[\vec{x}] \) is a function on \( \mathfrak{r}_\delta \), and
\[
\partial_{\vec{x}} z_\kappa = -z_\kappa^2 \frac{e^{\kappa+\frac{1}{2}} - e^{\kappa-\frac{1}{2}}}{\delta}, \tag{26}
\]
where we denote for \( k = 0, \ldots, K \) by \( e_k \in \mathbb{R}^{K+1} \) the \((k+1)\)th canonical unit vector.

**Remark 1.** One of the most fundamental properties of the \( L^2 \)-Wasserstein metric \( \mathcal{W}_2 \) on \( \mathcal{P}(\mathbb{R}) \) in one space dimension is its explicit representation in terms of Lagrangian coordinates. We refer to [1, 49] for a comprehensive introduction to the topic. This enables us to prove the existence of \( K \)-independent constants \( c_1, c_2 > 0 \), such that
\[
c_1 \| \vec{x} - \vec{y} \|_\delta \leq \mathcal{W}_2(\mathfrak{z}_\delta[\vec{x}], \mathfrak{z}_\delta[\vec{y}]) \leq c_2 \| \vec{x} - \vec{y} \|_\delta \quad \text{for all} \quad \vec{x}, \vec{y} \in \mathfrak{r}_\delta. \tag{27}
\]
A proof of the analogue statement formulated on a domain \([a, b] \subset (-\infty, +\infty)\) is given in [45, Lemma 7], and can be easily recomposed for the whole set of real numbers.

Let us further introduce the sets of (semi)-indizes
\[
\mathfrak{i}_K^1 = \{0, 1, \ldots, K\}, \quad \text{and} \quad \mathfrak{i}_K^{1/2} = \left\{ \frac{1}{2}, \frac{3}{2}, \ldots, K - \frac{1}{2} \right\}.
\]
The calculation (26) in the above example yields the explicit representation of the gradient of \( H_{\alpha,\lambda}(\vec{x}) \),
\[
\partial_{\vec{x}} H_{\alpha,0}(\vec{x}) = \Theta_\alpha \delta \sum_{\kappa \in \mathfrak{i}_K^{1/2}} z_\kappa^{\alpha+\frac{1}{2}} \frac{e^{\kappa+\frac{1}{2}} - e^{\kappa-\frac{1}{2}}}{\delta}, \tag{28}
\]
and further of the discretized information functional
\[
F_{\alpha,0}(\vec{x}) = \| \nabla_\delta H_{\alpha,0}(\vec{x}) \|_\delta^2 = \Theta_\alpha^2 \delta \sum_{\kappa \in \mathfrak{i}_K^{1/2}} \left( \frac{z_\kappa^{\alpha+\frac{1}{2}} - z_\kappa^{\alpha-\frac{1}{2}}}{\delta} \right)^2.
\]

In the case of positive confinement \( \lambda > 0 \), note that the drift potential \( u \mapsto \int_\mathbb{R} |x|^2 u(x) \, dx \) attains an equivalent representation in terms of Lagrangian coordinates that is \( X \mapsto \int_0^{|X|} |X(\xi)|^2 \, d\xi \). In our setting, the simplest discretization of this functional is hence by summing-up over all values \( x_k \) weighted with \( \delta \). This yields
\[
H_{\alpha,\lambda}(\vec{x}) = H_{\alpha,0}(\vec{x}) + \frac{\lambda}{2} \delta \sum_{k \in \mathfrak{i}_K^1} |x_k|^2, \quad \text{and} \quad F_{\alpha,\lambda}(\vec{x}) = F_{\alpha,0}(\vec{x}) + \frac{\lambda}{2} \delta \sum_{k \in \mathfrak{i}_K^1} |x_k|^2
\]
as an extension to the case of positive \( \lambda \), which is nothing else than (11) and (12). Note in addition that \( \delta \sum_{k \in \mathfrak{i}_K^1} |x_k|^2 = \| \vec{x} \|_\delta^2 \).
A first structural property of the above simple discretization are convexity retention from the continuous to the discrete setting. For reasons of readability, the proof of the following lemma is located in the appendix.

**Lemma 2.2.** The functional $\bar{x} \mapsto H_{\alpha,\lambda}$ is $\Lambda_{\alpha,\lambda}$-convex, i.e.

$$H_{\alpha,\lambda}((1-s)\bar{x} + s\bar{y}) \leq (1-s)H_{\alpha,\lambda}(\bar{x}) + sH_{\alpha,\lambda}(\bar{y}) - \frac{\Lambda_{\alpha,\lambda}}{2} (1-s)s\|\bar{x} - \bar{y}\|_d^2$$  \hspace{1cm} (29)

for any $\bar{x}, \bar{y} \in \mathfrak{T}_d$ and $s \in (0,1)$. It therefore admits a unique minimizer $\bar{x}_d^{\min} \in \mathfrak{T}_d$. If we further assume $\Lambda_{\alpha,\lambda} > 0$, then one attains for any $\bar{x} \in \mathfrak{T}_d$

$$\frac{\Lambda_{\alpha,\lambda}}{2} \|\bar{x} - \bar{x}_d^{\min}\|_d^2 \leq H_{\alpha,\lambda}(\bar{x}) - H_{\alpha,\lambda}(\bar{x}_d^{\min}) \leq \frac{1}{2\Lambda_{\alpha,\lambda}} \|\nabla_{\delta} H_{\alpha,\lambda}(\bar{x})\|_d^2.$$  \hspace{1cm} (30)

As a further conclusion of our natural discretization, we get a discrete fundamental entropy-information relation analogously to the continuous case (23).

**Corollary 1.** For any $\lambda \geq 0$, every $\bar{x} \in \mathfrak{T}_d$ with $H_{\alpha,0}(\bar{x}) < \infty$ we have

$$F_{\alpha,\lambda}(\bar{x}) = \|\nabla_{\delta} H_{\alpha,\lambda}(\bar{x})\|_d^2 + (2\alpha - 1)\Lambda_{\alpha,\lambda} H_{\alpha,\lambda}(\bar{x})$$ for $\alpha \in \left(\frac{1}{2},1\right]$ and (31)

$$F_{1/2,\lambda}(\bar{x}) = \|\nabla_{\delta} H_{1/2,\lambda}(\bar{x})\|_d^2 + \Lambda_{1/2,\lambda}$$ for $\alpha = \frac{1}{2}$. \hspace{1cm} (32)

The proof of this corollary can again be found in the appendix.

**Remark 2.** Note that the above seemingly appearing discontinuity at $\alpha = \frac{1}{2}$ is not real. For $\alpha > \frac{1}{2}$, the second term in the right hand side of (31) is explicitly given by

$$(2\alpha - 1)\Lambda_{\alpha,\lambda} H_{\alpha,\lambda}(\bar{x}) = (2\alpha - 1)\Lambda_{\alpha,\lambda} \left(\Theta_{\alpha} \delta \sum_{\kappa \in \Pi_{\delta}^{1/2}} z_{\kappa}^{\alpha-1/2} \frac{z_{\kappa}^{\alpha-1/2}}{\alpha - 1/2} + \frac{\Lambda_{\alpha,\lambda}}{2} ||\bar{x}||_d^2\right)$$

$$= 2\Lambda_{\alpha,\lambda} \Theta_{\alpha} \delta \sum_{\kappa \in \Pi_{\delta}^{1/2}} z_{\kappa}^{\alpha-1/2} + (2\alpha - 1) \frac{\Lambda_{\alpha,\lambda}}{2} ||\bar{x}||_d^2.$$  \hspace{1cm} (33)

For $\alpha \downarrow \frac{1}{2}$, one has $\Lambda_{\alpha,\lambda} \to \Lambda_{1/2,\lambda}$, $\Theta_{\alpha} \to \frac{1}{2}$, and especially $\delta \sum_{\kappa \in \Pi_{\delta}^{1/2}} z_{\kappa}^{\alpha-1/2} \to M = 1$. The drift-term vanishes since $(2\alpha - 1) \to 0$.

For the following reason, the above representation of $F_{\alpha,\lambda}$ is indeed a little miracle: From a naive point of view, one would ideally hope to gain a discrete counterpart of the fundamental entropy-information relation (23), if one takes the one-to-one discretization of the $L^2$-Wasserstein metric, which is (in the language of Lagrangian vectors) realized by the norm $\bar{x} \mapsto W_2(\mathfrak{u}_d[\bar{x}], \mathfrak{u}_d[\bar{x}])$ instead of our simpler choice $\bar{x} \mapsto ||\bar{x}||_d^2$. Indeed, with this ansatz, the proof of the above statement would fail in the moment in which one tries to calculate the scalar product of $\partial_{\bar{x}} F_{\alpha,0}$ and $\partial_{\bar{x}} W_2(\mathfrak{u}_d[\bar{x}], \mathfrak{u}_d[\bar{x}])$. This is why our discretization of the $L^2$-Wasserstein metric by the norm $||\cdot||_d$ seems to be the right choice, if one is interested in a structure-preserving discretization.

**Corollary 2.** The unique minimizer $\bar{x}_d^{\min} \in \mathfrak{T}_d$ of $H_{\alpha,\lambda}$ is a minimizer of $F_{\alpha,\lambda}$ and it satisfies for any $\bar{x} \in \mathfrak{T}_d$

$$F_{\alpha,\lambda}(\bar{x}) - F_{\alpha,\lambda}(\bar{x}_d^{\min}) \leq \frac{2\alpha + 1}{2} \|\nabla_{\delta} H_{\alpha,\lambda}(\bar{x})\|_d^2.$$  \hspace{1cm} (33)
Proof. Equality (51) and $2\alpha - 1 \geq 0$ show that $\bar{x} \mapsto F_{\alpha,\lambda}(\bar{x})$ is minimal, iff $\|\nabla_x H_{\alpha,\lambda}(\bar{x})\|_\delta = 0$ and $H_{\alpha,\lambda}(\bar{x})$ is minimal, which is the case for $\bar{x} = \bar{x}_\delta^{\min}$. The representation in (51) further implies

$$F_{\alpha,\lambda}(\bar{x}) - F_{\alpha,\lambda}(\bar{x}_\delta^{\min}) = \|\nabla_x H_{\alpha,\lambda}(\bar{x})\|_\delta^2 + (2\alpha - 1)\Lambda_{\alpha,\lambda}(H_{\alpha,\lambda}(\bar{x}) - H_{\alpha,\lambda}(\bar{x}_\delta^{\min}))$$

$$\leq \left(1 + \frac{2\alpha - 1}{2}\right)\|\nabla_x H_{\alpha,\lambda}(\bar{x})\|_\delta^2,$$

where we used (30) in the last step. $\blacksquare$

2.2.2. Interpretation of the scheme as discrete Wasserstein gradient flow. Starting from the discretized perturbed information functional $F_{\alpha,\lambda}$ we approximate the spatially discrete gradient flow equation

$$\dot{x} = -\nabla_x F_{\alpha,\lambda}(x) \quad (34)$$

also in time, using minimizing movements. To this end, remember the temporal decomposition of $[0, +\infty)$ by

$$\{0 = t_0 < t_1 < \ldots < t_n < \ldots\}, \quad \text{where} \quad t_n = t_{n-1} + \tau_n,$$

using time step sizes $\tau := \{\tau_1, \tau_2, \ldots, \tau_n, \ldots\}$ with $\tau_n \leq \tau$ and $\tau > 0$. As before in the introduction, we combine the spatial and temporal mesh widths in a single discretization parameter $\Delta = (\tau; \delta)$. For each $\bar{y} \in \bar{\gamma}_\delta$, introduce the Yosida-regularized information functional $F_{\alpha}(\cdot; \cdot, \cdot, \cdot, \bar{y}) : [0, +\infty) \times [0, \tau] \times \bar{\gamma}_\delta$ by

$$F_{\alpha}(\lambda, \sigma, \bar{x}, \bar{y}) = \frac{1}{2\sigma} \|\bar{x} - \bar{y}\|_\delta^2 + F_{\alpha,\lambda}(\bar{x}). \quad (35)$$

A fully discrete approximation $(\bar{x}_\Delta^n)_{n=0}^\infty$ of (34) is defined inductively from a given initial datum $\bar{x}_0$ by choosing each $\bar{x}_\Delta^n$ as a global minimizer of $F_{\alpha}(\lambda, \tau_n, \cdot, \bar{x}_\Delta^{n-1})$. Below, we prove that such a minimizer always exists (see Lemma 2.3).

In practice, one wishes to define $\bar{x}_\Delta^n$ as — preferably unique — solution of the Euler-Lagrange equations associated to $F_{\alpha}(\lambda, \tau_n, \cdot, \bar{x}_\Delta^{n-1})$, which leads to the implicit Euler time stepping:

$$\bar{x}_\Delta^{n+1} - \bar{x}_\Delta^n = -\nabla_x F_{\alpha,\lambda}(\bar{x}). \quad (36)$$

Using the explicit representation of $\partial_x F_{\alpha,\lambda}$, it is immediately seen that (36) is indeed the same as (6). Equivalence of (36) and the minimization problem is guaranteed at least for sufficiently small $\tau > 0$, as the following proposition shows.

Proposition 1. For each discretization $\Delta$ and every initial condition $\bar{x}_0 \in \bar{\gamma}_\delta$, the sequence of equations (36) can be solved inductively. Moreover, if $\tau > 0$ is sufficiently small with respect to $\delta$ and $F_{\alpha,\lambda}(\bar{x})$, then each (36) possesses a unique solution with $F_{\alpha,\lambda}(\bar{x}) \leq F_{\alpha,\lambda}(\bar{x})$, and that solution is the unique global minimizer of $F_{\alpha}(\lambda, \tau_n, \cdot, \bar{x}_\Delta^{n-1})$.

The proof of this proposition is a consequence of the following rather technical lemma.

Lemma 2.3. Fix a spatial discretization parameter $\delta$ and a bound $C > 0$. Then for every $\bar{y} \in \bar{\gamma}_\delta$ with $F_{\alpha,\lambda}(\bar{y}) \leq C$, the following are true:
for each $\sigma > 0$, the function $F_\alpha(\lambda, \sigma, \cdot, \vec{y})$ possesses at least one global minimizer $\vec{x}^* \in I_\delta$;

• there exists a $\tau_C > 0$ independent of $\vec{y}$ such that for each $\sigma \in (0, \tau_C)$, the global minimizer $\vec{x}^* \in I_\delta$ is strict and unique, and it is the only critical point of $F_\alpha(\lambda, \sigma, \cdot, \vec{y})$ with $F_{\alpha, \lambda}(\vec{x}) \leq C$.

Proof. Fix $\vec{y} \in I_\delta$ with $F_{\alpha, \lambda}(\vec{y}) \leq C$, and define the nonempty (since it contains $\vec{y}$) sublevel $A_C := (F_\alpha(\lambda, \sigma, \cdot, \vec{y}))^{-1}((0, C]) \subset I_\delta$. First observe that any $\vec{x} \in A_C$ satisfies

$$\sqrt{2\sigma C} \geq ||\vec{y} - \vec{x}||_\delta \geq \delta^{\frac{1}{2}} ||\vec{y}||_\infty - \delta \frac{1}{2} ||\vec{x}||_\infty,$$

hence $||\vec{x}||_\infty$ is bounded from above by $\sqrt{2\sigma C} + ||\vec{y}||_\infty$. Especially,

$$\max_{\kappa \in I_K} |x_\kappa| \leq \sqrt{2\delta^{-1}\sigma C} + ||\vec{y}||_\infty =: L(\delta, \sigma, \vec{y}),$$

which means, in the sense of density functions, that any $u = u_\delta[\vec{x}]$ with $\vec{x} \in A_C$ is compactly supported in $[-L(\delta, \sigma, \vec{y}), L(\delta, \sigma, \vec{y})]$. Consequently, take again $\vec{x} \in I_\delta$ arbitrarily and declare $z_\kappa = \min_{\kappa \in I_K^2} z_\kappa$ and $z^* = \max_{\kappa \in I_K^2} z_\kappa$, then on the one hand the conservation of mass yields the boundedness of $z_\kappa$ from above,

$$1 = \int_R u_\delta[\vec{x}] \, dx = \sum_{\kappa \in I_K^2} z_\kappa^{-1} (x_{\kappa + \frac{1}{2}} - x_{\kappa - \frac{1}{2}}) \leq 2L(\delta, \sigma, \vec{y})(z_*)^{-1},$$

and on the other hand $F_{\alpha, \lambda}(\vec{x}) \leq F_{\alpha, \lambda}(\vec{y}) \leq C$ yields an upper bound for $z^*$, since

$$(z^*)^{\alpha + \frac{1}{2}} - (z_*)^{\alpha + \frac{1}{2}} \leq \sum_{\kappa \in I_K^2} |z_{\kappa + \frac{1}{2}}^{\alpha + \frac{1}{2}} - z_{\kappa - \frac{1}{2}}^{\alpha + \frac{1}{2}}|$$

$$\leq \left( \sum_{\kappa \in I_K^2} \delta \right)^{\frac{1}{2}} \left( \delta \sum_{\kappa \in I_K^2} \left( \frac{z_{\kappa + \frac{1}{2}}^{\alpha + \frac{1}{2}} - z_{\kappa - \frac{1}{2}}^{\alpha + \frac{1}{2}}}{\delta} \right)^2 \right)^{\frac{1}{2}},$$

and hence

$$z^* \leq \left( M\Theta^{-1} C + (2L(\delta, \sigma, \vec{y}))^{\alpha + \frac{1}{2}} \right)^{1/(\alpha + 1/2)}.$$

Collecting the above observations, we first conclude that $A_C \subset I_\delta$ is a compact subset of $R^{K+1}$, due to $|x_0|, |x_K| \leq L(\delta, \sigma, \vec{y})$ and the continuity of $F_{\alpha, \lambda}$. Moreover, every vector $\vec{x} \in A_C$ satisfies $x_{\kappa + \frac{1}{2}} - x_{\kappa - \frac{1}{2}} \geq \delta(z^*)^{-1} \geq \delta$ for all $\kappa \in I_K^2$ with a positive constant $\delta$ that depends on $C$ and $L(\delta, \sigma, \vec{y})$. Thus $A_C$ does not touch the boundary (in the ambient $R^{K+1}$) of $I_\delta$. Consequently, $A_C$ is closed and bounded in $I_\delta$, endowed with the trace topology.

The restriction of the continuous function $F_\alpha(\lambda, \sigma, \cdot, \vec{y})$ to the compact set $A_C$ possesses a minimizer $\vec{x}^* \in A_C$. We clearly have $F_{\alpha, \lambda}(\vec{x}^*) \leq F_{\alpha, \lambda}(\vec{y}) \leq C$, and so $\vec{x}^*$ lies in the interior of $A_C$ and therefore is a global minimizer of $F_{\alpha, \lambda}(\lambda, \sigma, \cdot, \vec{y})$. This proves the first claim.

Since $F_{\alpha, \lambda} : I_\delta \to R$ is smooth, its restriction to $A_C$ is $\lambda_C$-convex with some $\lambda_C \leq 0$, i.e., $\partial^2 F_{\alpha, \lambda}(\vec{x}) \geq \lambda_C 1_{K+1}$ for all $\vec{x} \in A_C$. Independently of $\vec{y}$, we have that

$$\partial^2 F_{\alpha, \lambda}(\lambda, \sigma, \vec{x}, \vec{y}) = \delta^2 F_{\alpha, \lambda}(\vec{x}) + \delta \frac{\vec{y}}{\tau} 1_{K+1},$$

for all $\vec{x} \in A_C$.
which means that \( \bar{x} \mapsto \mathbf{F}_\alpha(\lambda, \sigma, \bar{x}, \bar{y}) \) is strictly convex on \( A_C \) if
\[
0 < \sigma < \tau_C := \frac{\delta}{(-\lambda_C)}.
\]

Consequently, each such \( \mathbf{F}_\alpha(\lambda, \sigma, \cdot, \bar{y}) \) has at most one critical point \( \bar{x}^* \) in the interior of \( A_C \), and this \( \bar{x}^* \) is necessarily a strict global minimizer. \( \square \)

**Remark 3** (propagation of the support). Take a solution \( \bar{x}_\Delta \) of (35) with density functions \( u_\Delta \). As we already noted in the above proof, any density \( u^*_\Delta \) has compact support in \([-L^n, L^n] \) with \( L^n = L(\delta, \tau_n, \bar{x}_\Delta^{n-1}) \) as in (38). Hence
\[
\|\bar{x}_\Delta^n\|_\infty \leq \sqrt{2\delta^{-1} f_{\alpha,\lambda}(\bar{x}_\Delta^{n-1}) + \|\bar{x}_\Delta^{n-1}\|_\infty}
\]
\[
\Rightarrow \|\bar{x}_\Delta^n\|_\infty \leq \sqrt{2\delta^{-1} \sum_{j=1}^n \|\delta_j \mathbf{F}_{\alpha,\lambda}(\bar{x}_\Delta^{j-1}) + \|x_\Delta^0\|_\infty},
\]
which is the best we can assume in case of \( \lambda = 0 \). If \( \lambda > 0 \), one can find a much better bound on the support of \( u_\Delta \), namely by replacing (37) by,
\[
\|\bar{x}_\Delta^n\|_\infty \leq \delta^{-1} \|\bar{x}_\Delta^n\|_\delta \leq \frac{2\delta^{-1}}{\lambda} \mathbf{F}_{\alpha,\lambda}(\bar{x}_\Delta^{n-1}) \leq \frac{2\delta^{-1}}{\lambda} \mathbf{F}_{\alpha,\lambda}(\bar{x}_\Delta^0).
\]

3. **Analysis of equilibrium.** In that which follows, we will analyse the long-time behavior in the discrete setting and will especially prove Theorem 1.2. As we have already seen in [44], the scheme’s underlying variational structure is essential to get optimal decay rates. Due to our structure-preserving discretization, it is even possible to derive analogue, asymptotically equal decay rates for solutions to (35).

3.1. **Entropy dissipation – the case of positive confinement** \( \lambda > 0 \). In this section, we pursue the discrete rate of decay towards discrete equilibria and try to verify the statements in Theorem 1.2 to that effect. That is why we assume henceforth \( \lambda > 0 \).

**Lemma 3.1.** A solution \( \bar{x}_\Delta \) to the discrete minimizing movement scheme (13) dissipates the entropies \( \mathbf{H}_{\alpha,\lambda} \) and \( \mathbf{F}_{\alpha,\lambda} \) at least exponential, i.e.
\[
(1 + 2\tau_n \lambda)(\mathbf{H}_{\alpha,\lambda}(\bar{x}_\Delta^n) - \mathbf{H}_{\alpha,\lambda}^{\text{min}}) \leq \mathbf{H}_{\alpha,\lambda}(\bar{x}_\Delta^{n-1}) - \mathbf{H}_{\alpha,\lambda}^{\text{min}}, \quad \text{and (40)}
\]
\[
(1 + 2\tau_n \lambda)(\mathbf{F}_{\alpha,\lambda}(\bar{x}_\Delta^n) - \mathbf{F}_{\alpha,\lambda}^{\text{min}}) \leq \mathbf{F}_{\alpha,\lambda}(\bar{x}_\Delta^{n-1}) - \mathbf{F}_{\alpha,\lambda}^{\text{min}}, \quad \text{for any time step } n \in \mathbb{N}.
\]

**Proof.** Due to (31), the gradient of the information functional \( \mathbf{F}_{\alpha,\lambda} \) is given by
\[
\partial_\delta \mathbf{F}_{\alpha,\lambda}(\bar{x}) = 2\sigma^{-1}(\partial_\delta \mathbf{H}_{\alpha,\lambda}(\bar{x}))^T \partial_\delta \mathbf{H}_{\alpha,\lambda}(\bar{x}) + (2\alpha - 1)\lambda \partial_\delta \mathbf{H}_{\alpha,\lambda}(\bar{x}),
\]
which yields in combination with the \( \Lambda_{\alpha,\lambda} \)-convexity of \( \mathbf{H}_{\alpha,\lambda} \) and (36)
\[
\mathbf{H}_{\alpha,\lambda}(\bar{x}_\Delta^{n-1}) - \mathbf{H}_{\alpha,\lambda}(\bar{x}_\Delta^n) \\
\geq \tau_n \|\nabla_\delta \mathbf{F}_{\alpha,\lambda}(\bar{x}_\Delta^n)\|_\delta \|\nabla_\delta \mathbf{H}_{\alpha,\lambda}(\bar{x}_\Delta^n)\|_\delta \\
\geq \tau_n \|\nabla_\delta \mathbf{H}_{\alpha,\lambda}(\bar{x}_\Delta^n)\|_\delta \|\nabla_\delta \mathbf{H}_{\alpha,\lambda}(\bar{x}_\Delta^n)\|_\delta \\
\geq \tau_n (2\alpha - 1)\lambda \|\nabla_\delta \mathbf{H}_{\alpha,\lambda}(\bar{x}_\Delta^n)\|^2 \\
\geq \tau_n (2\alpha - 1)\lambda \|\nabla_\delta \mathbf{H}_{\alpha,\lambda}(\bar{x}_\Delta^n)\|^2 \\
\geq \tau_n (2\alpha + 1)\lambda \|\nabla_\delta \mathbf{H}_{\alpha,\lambda}(\bar{x}_\Delta^n)\|^2.
\]
Using inequality (30), we conclude in

\[(1 + 2\tau_n(2\alpha + 1)\Lambda_{\alpha,\lambda}^2) (H_{\alpha,\lambda}(x^{n}_{\delta}) - H_{\alpha,\lambda}(x^{n}_{\delta})) \leq H_{\alpha,\lambda}(x^{n-1}_{\delta}) - H_{\alpha,\lambda}(x^{\min}_{\delta})\]

for any \(n \in \mathbb{N}\). Since \((2\alpha + 1)\Lambda_{\alpha,\lambda}^2 = \lambda\), this shows (40). To prove (41), we first note that a Cauchy-Schwarz inequality and (36) imply

\[
\langle \nabla \delta F_{\alpha,\lambda}(x^{n}_{\delta}), \nabla \delta H_{\alpha,\lambda}(x^{n}_{\delta}) \rangle_{\delta} \leq \frac{1}{\tau_n} \| \nabla \delta H_{\alpha,\lambda}(x^{n}_{\delta}) \|_{\delta} \| x^{n}_{\delta} - x^{n-1}_{\delta} \|_{\delta},
\]

and thanks to (42) further

\[
\tau_n(2\alpha + 1)\Lambda_{\alpha,\lambda} \| \nabla \delta H_{\alpha,\lambda}(x^{n}_{\delta}) \|_{\delta}^2 \leq \| \nabla \delta H_{\alpha,\lambda}(x^{n}_{\delta}) \|_{\delta} \| x^{n}_{\delta} - x^{n-1}_{\delta} \|_{\delta}.
\]

As a consequence, we get two types of inequalities, namely

\[
\tau_n(2\alpha + 1)\Lambda_{\alpha,\lambda} \| \nabla \delta H_{\alpha,\lambda}(x^{n}_{\delta}) \|_{\delta}^2 \leq \| x^{n}_{\delta} - x^{n-1}_{\delta} \|_{\delta} \quad \text{and}
\]

\[
2\alpha^2\Lambda_{\alpha,\lambda} \| F_{\alpha,\lambda}(x^{n}_{\delta}) - F_{\alpha,\lambda}(x^{\min}_{\delta}) \|_{\delta} \leq \| x^{n}_{\delta} - x^{n-1}_{\delta} \|_{\delta}^2,
\]

where we used \(\Lambda_{\alpha,\lambda} = \sqrt{\lambda/(2\alpha + 1)}\) and (35). To get the desired estimate, we denote for fixed \(n \geq 1\) and arbitrary \(\sigma \in (0, \tau_n]\) by \(x^{n}_{\sigma}\) a minimizer of \(x \mapsto \frac{1}{\tau_n} \| x - x^{n-1}_{\delta} \|_{\delta}^2 + F_{\alpha,\lambda}(x^{n}_{\delta})\). Then \(x^{n}_{\sigma}\) connects \(x^{n-1}_{\delta}\) and \(x^{n}_{\delta}\), and the monotonicity of \(\sigma \mapsto F_{\alpha,\lambda}(x^{n}_{\delta})\) and (43) yields for any \(\sigma \in (0, \tau]\)

\[
2\alpha^2\Lambda_{\alpha,\lambda} \| F_{\alpha,\lambda}(x^{n}_{\delta}) - F_{\alpha,\lambda}(x^{\min}_{\delta}) \|_{\delta} \leq 2\alpha^2\Lambda_{\alpha,\lambda} \| F_{\alpha,\lambda}(x^{n}_{\delta}) - F_{\alpha,\lambda}(x^{\min}_{\delta}) \|_{\delta} \leq \| x^{n}_{\delta} - x^{n-1}_{\delta} \|_{\delta}^2.
\]

Due to Lemma A.1 from the appendix, one attains

\[
F_{\alpha,\lambda}(x^{n}_{\delta}) + \frac{\| x^{n}_{\delta} - x^{n-1}_{\delta} \|_{\delta}^2}{2\tau_n} + \int_0^{\tau_n} \frac{\| x^{n}_{\sigma} - x^{n-1}_{\delta} \|_{\delta}^2}{2\sigma^2} d\sigma = F_{\alpha,\lambda}(x^{n-1}_{\delta}).
\]

Inserting (44) in the above equation then finally yields

\[(1 + 2\tau_n\lambda) \left( F_{\alpha,\lambda}(x^{n}_{\delta}) - F_{\alpha,\lambda}(x^{\min}_{\delta}) \right) \leq F_{\alpha,\lambda}(x^{n-1}_{\delta}) - F_{\alpha,\lambda}(x^{\min}_{\delta}),\]

and the claim is proven.

The proof is a special case of [11 Theorem 3.1.4]

\(\square\)

**Remark 4.** In the continuous situation, the analogue proofs of (40) and (41) require a more deeper understanding of variational techniques. An essential tool in this context is the flow interchange lemma, see e.g. [11 Theorem 3.2]. Although one can easily proof a discrete counterpart of the flow interchange lemma, it is not essential in the above proof, since the smoothness of \(x \mapsto H_{\alpha,\lambda}(x)\) allows an explicit calculation of its gradient and hessian.

Lemma 3.1 paves the way for the exponential decay rates of Theorem 1.2. Effectively, (18) and (19) are just applications of the following version of the discrete Gronwall lemma: Assume \(\{c_n\}_{n \in \mathbb{N}}\) and \(\{y_n\}_{n \in \mathbb{N}}\) to be sequences with values in \(\mathbb{R}_+\), satisfying \((1 + c_n)y_n \leq y_{n-1}\) for any \(n \in \mathbb{N}\), then

\[y_n \leq y_0 e^{-\sum_{k=0}^{n-1} \frac{c_k}{1+c_k}} \quad \text{for any } n \in \mathbb{N}.
\]

This statement can be easily proven by induction. Furthermore, inequality (20) is then a corollary of (18) and a Csiszar-Kullback inequality, see [13 Theorem 30].
3.1.1. Convergence towards Barenblatt profiles and Gaussians. This section is devoted to the proof of Theorem 1.1. Hence let us assume again \( \lambda > 0 \). To prove the statement of this theorem we are going to show that the sequence of functionals \( \mathcal{H}_{\alpha, \lambda}^\delta : \mathcal{P}(\mathbb{R}) \to (-\infty, +\infty] \) given by
\[
\mathcal{H}_{\alpha, \lambda}^\delta(u) := \begin{cases} 
\mathcal{H}_{\alpha, \lambda}(u), & u \in \mathcal{P}_\delta(\mathbb{R}) \\
+\infty, & \text{else} 
\end{cases}
\]
\( \Gamma \)-converges towards \( \mathcal{H}_{\alpha, \lambda} \). More detailed, the following points are satisfied for any \( u \in \mathcal{P}(\mathbb{R}) \):

(i) \( \liminf_{\delta \to 0} \mathcal{H}_{\alpha, \lambda}^\delta(u_\delta) \geq \mathcal{H}_{\alpha, \lambda}(u) \) for any sequence \( u_\delta \) with \( \lim\delta \to 0 W_2(u_\delta, u) = 0 \).

(ii) There exists a recovery sequence \( u_\delta \) of \( u \), i.e. \( \limsup_{\delta \to 0} \mathcal{H}_{\alpha, \lambda}^\delta(u_\delta) \leq \mathcal{H}_{\alpha, \lambda}(u) \) and \( \lim\delta \to 0 W_2(u_\delta, u) = 0 \).

The \( \Gamma \)-convergence of \( \mathcal{H}_{\alpha, \lambda}^\delta(\cdot) \) towards \( \mathcal{H}_{\alpha, \lambda} \) is a powerful property, since it implies convergence of the sequence of minimizers \( u_\delta^{\text{min}} = u_\delta^{[\mathcal{H}_{\alpha, \lambda}^\delta]} \) towards \( b_{\alpha, \lambda} \) or \( b_{1/2, \lambda} \), respectively, with respect to the \( L^2 \)-Wasserstein metric, see [8]. To conclude even strong convergence of \( u_\delta^{\text{min}} \) at least in \( L^p(\mathbb{R}) \) for arbitrary \( p \geq 1 \), we proceed similar as in [45] Proposition 18. Necessary for that, recall that the total variation of a function \( f \in L^1(\mathbb{R}) \) is given by
\[
\{f\}_{TV} := \sup \left\{ \int_{\mathbb{R}} f(x)\varphi'(x) \, dx \mid \varphi \text{ is Lipschitz continuous with compact support,} \right. \\
\left. \sup_{x \in \mathbb{R}} |\varphi(x)| \leq 1 \right\},
\]
we refer to [30], Definition 1.1. If \( f \) is a piecewise constant function with compact support \( [x_0, x_K] \), taking values \( f_{k-\frac{1}{2}} \) on intervals \( (x_{k-1}, x_k) \), then the integral in (45) amounts to
\[
\int_{\mathbb{R}} f(x)\varphi'(x) \, dx = \sum_{k=1}^{K} \left[ f(x)\varphi(x) \right]_{x=x_{k-1}+0}^{x_k-0} \\
= \sum_{k=1}^{K-1} (f_{k-\frac{1}{2}} - f_{k+\frac{1}{2}}) \varphi(x_k) + f_{\frac{1}{2}} \varphi(x_0) - f_{K-\frac{1}{2}} \varphi(x_K).
\]
Consequently, for such \( f \), the supremum in (45) equals
\[
\{f\}_{TV} = \sum_{k=1}^{K-1} |f_{k+\frac{1}{2}} - f_{k-\frac{1}{2}}| + |f_{\frac{1}{2}}| + |f_{K-\frac{1}{2}}| 
\]
and is attained for every Lipschitz continuous function \( \varphi \) with \( \varphi(x_k) = \text{sgn}(f_k - f_{k+1}) \) at \( k = 1, \ldots, K-1 \), \( \varphi(x_0) = \text{sgn}(f_{\frac{1}{2}}) \) and \( \varphi(x_K) = -\text{sgn}(f_{K-\frac{1}{2}}) \).

**Lemma 3.2.** For any \( \alpha \in [\frac{1}{2}, 1] \), assume \( \bar{x}_\delta^{\text{min}} \in \mathcal{P}_\delta \) to be the unique minimizer of \( \mathcal{H}_{\alpha, \lambda}^\delta \) and declare the sequence of functions \( u_\delta^{\text{min}} = u_\delta^{[\mathcal{H}_{\alpha, \lambda}^\delta]} \). Then
\[
u^{\text{min}} \xrightarrow{\delta \to 0} b_{\alpha, \lambda} \text{ strongly in } L^p(\mathbb{R}) \text{ for any } p \geq 1, \\
\hat{u}_\delta^{\text{min}} \xrightarrow{\delta \to 0} b_{\alpha, \lambda} \text{ uniformly on } \mathbb{R},
\]
where \( \hat{u}_3^{\text{min}} : \mathbb{R} \to \mathbb{R} \) is a local affine interpolation of \( u_3^{\text{min}} \) on \( \mathbb{R} \), such that for any \( \kappa \in \mathbb{I}^{1/2}_K \cup \mathbb{I}^0_K \)

\[
(\hat{u}_3^{\text{min}} \circ X_{\delta}[\xi]) (\xi) = \begin{cases} z_{\kappa}, & \kappa \in \mathbb{I}^{1/2}_K \\ \frac{1}{2} (z_{\kappa + \frac{1}{2}} + z_{\kappa - \frac{1}{2}}), & \kappa \in \mathbb{I}^0_K \end{cases}.
\]

Proof. We will first prove the \( \Gamma \)-convergence of \( \mathcal{H}^{\delta}_{\alpha, \lambda} \) towards \( \mathcal{H}_{\alpha, \lambda} \). The first requirement (i) is a conclusion of the lower semi-continuity of \( \mathcal{H}_{\alpha, \lambda} \).

For the second point (ii), we fix \( u \in \mathcal{P}(\mathbb{R}) \) and assume \( X : [0, M] \to [-\infty, +\infty] \) to be the Lagrangian map of \( u \). In addition assume for the moment, that \( \int_0^M u(x) \, dx < M \) for any compact subset \( K \subseteq \mathbb{R} \), hence there is no bounded interval such that the whole mass of \( u \) is concentrated therein. Further assume without loss of generality that the center of mass is at \( x = 0 \), i.e. \( \int_{-\infty}^0 u(x) \, dx = M/2 \). Then one can find for any \( \varepsilon > 0 \) a compact set of the form \( K = [L_1, L_2] \) with \( L_1 < 0 < L_2 \), and an integer \( K \in \mathbb{N} \), such that

\[
\int_{\mathbb{R}\setminus K} |x|^2 u(x) \, dx < \varepsilon, \quad \text{and} \quad \int_{-\infty}^{L_1} u(x) \, dx = \int_{L_2}^{+\infty} u(x) \, dx = \delta := MK^{-1} \quad (49)
\]

The first statement is valid due to the boundedness of the second momentum of \( u \), and the last one is satisfied since one can choose \( K \in \mathbb{N} \) arbitrarily large. An immediate consequence of the above choices is that \( 2\delta L^2 < \varepsilon \) for \( L = \max\{|L_1|, |L_2|\} \) due to

\[
2(L_1^2 + L_2^2)\delta = L_1^2 \int_{-\infty}^{L_1} u(x) \, dx + L_2^2 \int_{L_2}^{+\infty} u(x) \, dx \leq \int_{\mathbb{R}\setminus K} |x|^2 u(x) \, dx < \varepsilon.
\]

Using \( \delta = MK^{-1} \) we define an equidistant decomposition \( \delta \) of the mass domain \( [0, M] \). We furthermore declare \( x_0 = -2L, x_K = 2L \) and \( x_k = X(\xi_k) \) for any \( k = 1, \ldots, K-1 \) and introduce the locally constant density \( u_\delta \in \mathcal{P}(\mathbb{R}) \) that corresponds to the Lagrangian map \( X_\delta[\xi] \). This procedure defines a sequence of densities \( u_\delta \), since \( \varepsilon > 0 \) was arbitrary, and we are going to prove that \( u_\delta \) is the right choice for the recovery sequence.

To prove the convergence in the \( L^2 \)-Wasserstein distance, we fix \( \varepsilon > 0 \). Then the last property of \( u \) in (49) yields especially that \( x_1 = L_1 \) and \( x_{K-1} = L_2 \). Furthermore, since \( X \) and \( X_\delta[\xi] \) are monotonically increasing, one obtains for any \( \xi \in [\xi_1, \xi_{K-1}] \)

\[
|X(\xi) - X_\delta[\xi](\xi)| \leq (X(\xi_k) - X(\xi_{k-1})) \leq 2L
\]

with \( \xi \in [\xi_{k-1}, \xi_k], k = 2, \ldots, K-1 \). Therefore

\[
\|X - X_\delta[\xi]\|_{L^2([-\epsilon, \epsilon])} \leq 2\delta L \sum_{k=2}^{K-1} (X(\xi_k) - X(\xi_{k-1})) \leq 2\delta L^2 < \varepsilon, \quad (50)
\]

where we used \( x_1 = L_1 \) and \( x_{K-1} = L_2 \). As a next step, we note that \( |x_1|, |x_{K-1}| \leq L \) and \( |x_0| = |x_K| = 2L \), which yields

\[
\|X_\delta[\xi]\|_{L^2([0, M]\setminus[\xi_1, \xi_{K-1}])} = \int_{[x_0, x_K]\setminus[x_1, x_{K-1}]} |x|^2 u_\delta(x) \, dx
\]

\begin{align*}
&= \frac{\delta}{x_K - x_{K-1}} \int_{x_{K-1}}^{x_K} |x|^2 \, dx + \frac{\delta}{x_1 - x_0} \int_{x_0}^{x_1} |x|^2 \, dx \\
&\leq \frac{2\delta}{3} (x_K^2 + x_{K-1}^2 + x_1^2 + x_0^2) \leq \frac{40}{3} \delta L^2 < 7\varepsilon, \quad (51)
\end{align*}
where we used the elementary equality \((a^3 - b^3) = (a - b)(a^2 + b^2 + ab)\). Combining (49), (50), and (51), and the fact that \(X((0, M] \setminus [\xi_1, \xi_{K-1}]) = \mathbb{R} \setminus \mathcal{K}\), we finally conclude

\[
\mathcal{W}_2(u, u_\delta) = \|X - X_\delta[\bar{x}]\|_{L^2([0, M])} \\
\leq \|X - X_\delta[\bar{x}]\|_{L^2([\xi_1, \xi_{K-1}])} + \|X - X_\delta[\bar{x}]\|_{L^2([0, M] \setminus [\xi_1, \xi_{K-1}])} \\
\leq \sqrt{\varepsilon} + \|X_\delta[\bar{x}]\|_{L^2([0, M] \setminus [\xi_1, \xi_{K-1}])} + \|X\|_{L^2([0, M] \setminus [\xi_1, \xi_{K-1}])} \\
\leq \sqrt{\varepsilon} + \sqrt{7\varepsilon} + \left(\int_{\mathbb{R} \setminus \mathcal{K}} |x|^2 u(x) \, dx\right)^{1/2} < 4\sqrt{\varepsilon}.
\]

This shows \(u_\delta \to u\) in \(L^2\)-Wasserstein as \(\delta \to 0\). The second point in (ii) easily follows by using Jensen’s inequality,

\[
\mathcal{H}^\delta_{\alpha, \lambda}(u_\delta) = \mathcal{H}_{\alpha, \lambda}(u_\delta) = \sum_{\kappa \in \mathbb{N}^{1/2}} \int_{x_{\kappa - 1/2}^+}^{x_{\kappa + 1/2}^-} \varphi_\alpha \left( \frac{\delta}{x_{\kappa + 1/2}^- - x_{\kappa - 1/2}^+} \right) \, dx
\]

\[
= \sum_{\kappa \in \mathbb{N}^{1/2}} (x_{\kappa + 1/2}^- - x_{\kappa - 1/2}^+) \varphi_\alpha \left( \frac{1}{x_{\kappa + 1/2}^- - x_{\kappa - 1/2}^+} \int_{x_{\kappa - 1/2}^-}^{x_{\kappa + 1/2}^+} u(s) \, ds \right)
\]

\[
\leq \sum_{\kappa \in \mathbb{N}^{1/2}} \int_{x_{\kappa - 1/2}^-}^{x_{\kappa + 1/2}^+} \varphi_\alpha (u(s)) \, ds = \int_{x_0}^{x_K} \varphi_\alpha (u(s)) \, ds.
\]

Adding the drift-term and taking the limes superior on both sides proves

\[
\limsup_{\delta \to 0} \mathcal{H}^\delta_{\alpha, \lambda}(u_\delta) \leq \mathcal{H}_{\alpha, \lambda}(u).
\]

Since \(\mathcal{H}_{\alpha, \lambda}\) is lower semi-continuous, we obtain \(\lim_{\delta \to 0} \mathcal{H}^\delta_{\alpha, \lambda}(u_\delta) = \mathcal{H}_{\alpha, \lambda}(u)\).

In the simpler case that \(\int K u \, dx = M\) for a compact subset \(\mathcal{K} \subseteq \mathbb{R}\), we choose \(X_\delta[\bar{x}]\) as the local affine function with \(X(\xi_k) = X(\xi_k)\) for any \(k = 0, \ldots, K\), and take the corresponding sequence \(u_\delta\) of local constant density functions. Then the above calculation can be achieved analogously.

To conclude the convergence of \(u_\delta^{\min}\) towards \(b_{\alpha, \lambda}\) with respect to \(\mathcal{W}_2\) and the convergence of \(\mathcal{H}_{\alpha, \lambda}(x_\delta^{\min})\) towards \(\mathcal{H}_{\alpha, \lambda}(b_{\alpha, \lambda})\), we invoke [8] Theorem 1.21. Note that \(\inf_{u \in P(\mathbb{R})} \mathcal{H}^\delta_{\alpha, \lambda}(u) = \mathcal{H}_{\alpha, \lambda}(x_\delta^{\min})\) by definition of \(\mathcal{H}^\delta_{\alpha, \lambda}\), hence the minimum of \(\mathcal{H}^\delta_{\alpha, \lambda}\) is \(u_\delta^{\min}\). Furthermore, each functional \(\mathcal{H}^\delta_{\alpha, \lambda}\) has precompact sublevels which is a consequence of \(\lambda > 0\) and Prokhorov’s Theorem, see for instance [11] Theorem 5.1.3. Since \(\mathcal{H}^\delta_{\alpha, \lambda}\) \(\Gamma\)-converges towards \(\mathcal{H}_{\alpha, \lambda}\), all requirements for [8] Theorem 1.21 are satisfied.

Let us now prove (47). The convergence of \(\mathcal{H}_{\alpha, \lambda}(x_\delta^{\min})\) to \(\mathcal{H}_{\alpha, \lambda}(b_{\alpha, \lambda})\) yields on the one hand the uniform boundedness of \(\mathcal{H}_{\alpha, \lambda}(x_\delta^{\min})\) with respect to the spatial discretization parameter \(\delta\), and on the other hand the uniform boundedness of \(\mathcal{F}_{\alpha, \lambda}(x_\delta^{\min})\), which is a conclusion of (51) and \(\nabla_\delta \mathcal{H}_{\alpha, \lambda}(x_\delta^{\min}) = 0\). Similar to [25] Proposition 18, one can easily prove that the term \(\mathcal{F}_{\alpha, \lambda}(x_\delta^{\min})\) is an upper bound on the total variation of \(P_\alpha(u_\delta^{\min})\) with \(P_\alpha(s) = \Theta_\alpha s^{\alpha + 1/2}\). Take an arbitrary \(\bar{y} \in \mathbb{R}^{K+1}\) with \(\|\bar{y}\|_\infty \leq 1\). Then

\[
\langle \nabla_\delta \mathcal{H}_{\alpha, 0}(x_\delta^{\min}), \bar{y} \rangle_\delta = \langle \nabla_\delta \mathcal{H}_{\alpha, \lambda}(x_\delta^{\min}), \bar{y} \rangle_\delta - \Lambda_{\alpha, \lambda}(x_\delta^{\min}, \bar{y})_\delta, \quad (52)
\]
and the left-hand side can be reformulated, using (28),

\[
\langle \nabla \delta H_{\alpha,0}(S_{\delta}^{\text{min}}), \mathbf{y} \rangle_\delta = \sum_{k=1}^{K} P_{\alpha}(z_k)(y_k - \frac{1}{\delta})
\]

\[
= \sum_{k=1}^{K-1} (P_{\alpha}(z_{k+\frac{1}{2}}) - P_{\alpha}(z_{k-\frac{1}{2}}))y_k + P_{\alpha}(z_{\frac{1}{2}})y_0 + P_{\alpha}(z_{K-\frac{1}{2}})y_K.
\]

Respecting that \(\|y\|_\delta \leq M\|y\|_\infty\), we can take the supremum over all \(y\) with \(\|y\|_\infty \leq 1\) in (52). Then the Cauchy-Schwarz inequality and the representation of \(\{\cdot\}_{TV}\) in (46) yields

\[
\{P_{\alpha}(\hat{u}_{\delta}^{\text{min}})\}_{TV} \leq M \|\nabla \delta H_{\alpha,\lambda}(S_{\delta}^{\text{min}})\|_\delta + \Lambda_{\alpha,\lambda} M \|\hat{u}_{\delta}^{\text{min}}\|_\delta,
\]

which is uniformly bounded from above due to (31) and the uniform boundedness of \(F_{\alpha,\lambda}(S_{\delta}^{\text{min}})\). This proves the uniform boundedness of \(\{P_{\alpha}(\hat{u}_{\delta}^{\text{min}})\}_{TV}\) and using the superlinear growth of \(s \mapsto P_{\alpha}(s)\) and [30] Proposition 1.19, we conclude in (47).

To proof (48), we show that the \(H^1(\mathbb{R})\)-norm of \(\hat{u}_{\delta}^{\text{min}}\) is bounded by the information functional. This was already done in [46] for \(\alpha = \frac{1}{2}\), where we also showed

\[
\|\hat{u}_{\delta}^{\text{min}}\|_{H^1(\mathbb{R})}^2 = \delta \sum_{k \in \mathbb{N}_K} \frac{z_{k+\frac{1}{2}} + z_{k-\frac{1}{2}}}{2} \left(\frac{z_{k+\frac{1}{2}} - z_{k-\frac{1}{2}}}{\delta}\right)^2.
\]

So assume \(\alpha \in (\frac{1}{2}, 1]\), then the concavity of the mapping \(s \mapsto s^{\alpha-1/2}\) yields for any values \(b > a > 0\) the validity of

\[
b^{\alpha+\frac{1}{2}} - a^{\alpha+\frac{1}{2}} = (\alpha + \frac{1}{2}) \int_a^b s^{\alpha-\frac{1}{2}} \, ds \geq (\alpha + \frac{1}{2}) \frac{b^{\alpha+\frac{1}{2}} - a^{\alpha+\frac{1}{2}}}{2} (b - a),
\]

and further \((b^{\alpha+\frac{1}{2}} - a^{\alpha+\frac{1}{2}})^2 \geq (\alpha + \frac{1}{2})^2 \frac{b^{2\alpha-1} - a^{2\alpha-1}}{4} (b - a)^2\). Therefore

\[
\|\hat{u}_{\delta}^{\text{min}}\|_{H^1(\mathbb{R})} \leq \delta \sum_{k \in \mathbb{N}_K} \frac{z_{k+\frac{1}{2}} + z_{k-\frac{1}{2}}}{2} \left(\frac{z_{k+\frac{1}{2}} - z_{k-\frac{1}{2}}}{\delta}\right)^2
\]

\[
\leq \|\hat{u}_{\delta}^{\text{min}}\|_{L^\infty(\mathbb{R})} \delta \sum_{k \in \mathbb{N}_K} \frac{z_{2\alpha-1}}{2} \left(\frac{z_{k+\frac{1}{2}} - z_{k-\frac{1}{2}}}{\delta}\right)^2
\]

\[
\leq \frac{2\|\hat{u}_{\delta}^{\text{min}}\|_{L^\infty(\mathbb{R})}}{(\alpha + \frac{1}{2})^2} \delta \sum_{k \in \mathbb{N}_K} \left(\frac{z_{\alpha+\frac{1}{2}}}{z_{k+\frac{1}{2}} - z_{k-\frac{1}{2}}}\right)^2
\]

\[
\leq \|\hat{u}_{\delta}^{\text{min}}\|_{L^\infty(\mathbb{R})} \frac{2(1-\alpha)}{(\alpha + \frac{1}{2})^2} F_{\alpha,\lambda}(S_{\delta}^{\text{min}}).
\]

\[\square\]

3.2. The case of zero confinement \(\lambda = 0\). This section is essentially devoted to the proof of Theorem 1.3. Let us therefore consider (1) in case of vanishing confinement \(\lambda = 0\), hence

\[
\partial_t u = -(u^{\alpha-1} u_{xx})_x \quad \text{for} \ (t, x) \in (0, +\infty) \times \mathbb{R},
\]

and \(u(0) = u^0\) for arbitrary initial density \(u^0 \in \mathcal{P}(\mathbb{R})\). From the continuous theory, it is known that solutions to (53) or (21) with \(\Lambda_{\alpha,\lambda} = 0\) branch out over the whole set of real numbers, hence converge towards zero at a.e. point. This matter of fact
makes rigorous analysis of the long-time behavior of solutions to (53) more difficult as in the case of positive confinement. However, the unperturbed functionals $H_{\alpha,0}$ and $F_{\alpha,0}$ satisfy the scaling property, see again [44],

$$H_{\alpha,0}(\delta_r u) = r^{-(2\alpha-1)/2}H_{\alpha,0}(u), \quad \text{and} \quad F_{\alpha,0}(\delta_r u) = r^{-(2\alpha+1)}F_{\alpha,0}(u),$$

(54)

for any $r > 0$ and $\delta_r u(x) := r^{-1}u(r^{-1}x)$ with $u \in \mathcal{P}(\mathbb{R})$. Due to this, it is possible to find weak solutions to a rescaled version of (53) by solving problem (1) with $\lambda = 1$. More detailed, the following lemma is valid:

**Lemma 3.3.** A function $u \in L^2_{\text{loc}}((0,T);W^{2,2}(\mathbb{R}))$ is a weak solution of (1) with $\lambda = 1$, iff

$$w(t,\cdot) = \delta_{R(t)}u(\log(R(t),\cdot) \quad \text{with} \quad R(t) := (1 + (2\alpha + 3)t)^{1/(2\alpha+3)}$$

(55)

is a weak solution to (53).

A consequence of the above lemma is that one can describe how solutions $w$ to (53) vanish asymptotically as $t \to \infty$, although the gained information is not very strong and useful: In fact the first observation (without studying local asymptotics in more detail) is, that $w$ decays to zero with the same rate as the rescaled (time-dependent) Barenblatt-profile $b_{\alpha,0}^*$ defined by $b_{\alpha,0}^*(t,\cdot) := \delta_{R(t)}b_{\alpha,1}$ with $R(t) w(0) = H_{\alpha,0}(w(0))$ with

$$||w(t,\cdot) - b_{\alpha,0}^*(t,\cdot)||_{L^1(\mathbb{R})} \leq CR(t)^{-1}$$

(56)

for any $t > 0$. In [44], this behavior was described using weak solutions constructed by minimizing movements. We will adopt this methodes to derive a discrete analogue of (56) for the discrete solutions $\tilde{x}_\Delta$ of (53).

First of all, we reformulate the scaling operator $\delta_r$ for fixed $r > 0$ in the setting of monotonically increasing vectors $\bar{x} \in \mathfrak{I}_\delta$. Since $\delta_r u(x) := r^{-1}u(r^{-1}x)$ for arbitrary density in $\mathcal{P}(\mathbb{R})$, the same can be done for $u_\delta = \bar{u}_\delta[\bar{x}]$, hence

$$\delta_r u_\delta(x) = \sum_{k=1}^K r^{-1}\delta \Pi(x_{x_{k-1}}x_{x_{k+1}})(r^{-1}x) = \sum_{k=1}^K \frac{\delta}{r x_{k-1} - r x_{k+1}} \Pi(r^{-1}x)$$

$$= \bar{u}_\delta(\bar{x})$$

for any $x \in \mathbb{R}$. The natural extension of $\delta_r$ to the set $\mathfrak{I}_\delta$ is hence

$$\delta_r \bar{x} := r \bar{x} \quad \text{with corresponding} \quad \delta_r \bar{z} = \bar{z}_\delta[\delta_r \bar{x}] = r^{-1} \bar{z}.$$ 

As a consequence of this definition, note the validity of the discrete scaling property for $H_{\alpha,0}$ and $F_{\alpha,0}$, i.e. for any $r > 0$ and $\bar{x} \in \mathfrak{I}_\delta$,

$$H_{\alpha,0}(\delta_r \bar{x}) = r^{-(2\alpha-1)/2}H_{\alpha,0}(\bar{x}), \quad \text{and} \quad F_{\alpha,0}(\delta_r \bar{x}) = r^{-(2\alpha+1)}F_{\alpha,0}(\bar{x}).$$

(57)

The first equality is satisfied due to $H_{\alpha,0}(\bar{x}) = H_{\alpha,0}(\bar{u}_\delta[\bar{x}])$ and the scaling property (54) of the continuous entropy functions. The analogue claim for $F_{\alpha,0}$ in (57) follows by inserting $\delta_r \bar{x}$ into $\partial_\delta H_{\alpha,0}$ and using $\delta_r \bar{z} = r^{-1} \bar{z}$, then

$$\partial_\delta H_{\alpha,0}(\delta_r \bar{x}) = \Theta_\delta \delta \sum_{\kappa \in \mathbb{Z}} (\delta_r \bar{z}_\kappa)^{\alpha+1/2} \frac{\delta \kappa}{\delta} = r^{-(\alpha+1/2)} \partial_\delta H_{\alpha,0}(\bar{x})$$

$$\implies F_{\alpha,0}(\delta_r \bar{x}) = ||\nabla_\delta H_{\alpha,0}(\delta_r \bar{x})||_{\delta}^2 = r^{-(2\alpha+1)} ||\nabla_\delta H_{\alpha,0}(\bar{x})||_{\delta}^2 = r^{-(2\alpha+1)} F_{\alpha,0}(\bar{x}).$$

This scaling properties can now be used to build a bridge between solutions of discrete minimizing movement schemes with $\lambda = 0$ to those with positive confinement.
The following lemma is based on the proof of [41, Theorem 5.6], but nevertheless, it is an impressive example for the powerful structure-preservation of our numerical scheme.

**Lemma 3.4.** Assume $\tilde{x}^* \in \mathfrak{F}_\delta$ with $F(\tilde{x}^*) < \infty$. Further fix $\tau > 0$ and $R > S > 0$. Then $\tilde{x} \in \mathfrak{F}_\delta$ is a minimizer of

$$
\tilde{y} \mapsto F(\lambda, \tau, \tilde{y}, \tilde{x}^*) = \frac{1}{2\tau} ||\tilde{y} - \tilde{x}^*||^2 + F(\tilde{y}) + \frac{\lambda}{2} ||\tilde{y}||^2,
$$

if and only if $\partial_R \tilde{x} \in \mathfrak{F}_\delta$ minimizes the functional

$$
\tilde{w} \mapsto F(\lambda, \tau, \tilde{w}, \partial_S \tilde{x}^*) = \frac{1}{2\tau} ||\tilde{w} - \partial_S \tilde{x}^*||^2 + F(\tilde{w}) + \frac{\lambda}{2} ||\tilde{w}||^2 \quad \text{with}
$$

$$
\tau = \tau \sigma R^{2\alpha+2}, \quad \lambda = \frac{S(1 + \lambda \tau) - R}{\tau R}.
$$

It is not difficult to see that this lemma can be formulated for any functional mapping $\mathfrak{F}_\delta$ to $\mathbb{R}$ with the same scaling property as $F(\alpha, 0)$ in (57). We refer to the appendix for the very technical proof of Lemma 3.4.

Before we prove the claim of Theorem 1.3 let us introduce the rescaled discrete Barenblatt-profile. Define inductively

$$
S_n^0 := 1, \quad S_n^\tau = (1 + \tau_n)S_n^{\tau-1}
$$

for $n \in \mathbb{N}$. Further take the minimizer $\hat{x}_n^{\min} \in \mathfrak{F}_\delta$ of the functional $\hat{x} \mapsto H_{\alpha, 0}(\hat{x})$. Then denote the scaled vector $\hat{b}_{\alpha, 0} := \partial S_n^\tau \hat{x}_n^{\min}$ and define its corresponding density function $b_{\alpha, 0} \hat{b}_{\alpha, 0} = u_{\delta}[b_{\alpha, 0}^{\min}]$. This function can be interpreted as a self-similar solution of (59) with initial density $u_{\delta}[b_{\alpha, 0}^{\min}]$, $\hat{\lambda} = 0$ and with time steps $\tau$ inductively defined by

$$
\hat{\tau}_n := \tau_n S_n^{\tau-1} S_n^{2\alpha+2}.
$$

**Proof of Theorem 1.3.** As already mentioned above, we define a sequence of functions $S_n^\tau$ inductively through (60) and declare a new partition of the time scale $[0, +\infty)$ by

$$
\{0 = \hat{s}_0 < \hat{s}_2 < \ldots < \hat{s}_n < \ldots\}, \quad \text{where} \quad \hat{s}_n := \sum_{k=1}^n \hat{\tau}_k \quad \text{and} \quad \hat{\tau}_k := \tau_k S_k^{\tau-1}(S_k^\tau)^{2\alpha+2},
$$

(61)

and we write $\tau = (\hat{\tau}_1, \hat{\tau}_2, \ldots)$. As a first consequence of the iterative character of the above object, we note that $(1 + x) \leq e^x$ causes $S_n^\tau \leq e^{\hat{t}_n}$ for any $n \in \mathbb{N}$. This further implies

$$
\hat{s}_n = \sum_{k=1}^n \tau_k S_k^{\tau-1}(S_k^\tau)^{2\alpha+2} = \sum_{k=1}^n \tau_k (1 + \tau_k)^{2\alpha+2}(S_k^{\tau-1})^{2\alpha+3} \leq (1 + \tau)^{2\alpha+2} \sum_{k=1}^n \tau_k e^{(2\alpha+3)t_k-1}.
$$

This estimate is beneficial, insofar as the right hand side is a lower sum of the integral $(1 + \tau)^{2\alpha+2} \int_0^{\hat{t}_n} e^{(2\alpha+3)t} \, ds$, hence

$$
\hat{s}_n \leq (1 + \tau)^{2\alpha+2}(2\alpha + 3)^{-1} [e^{(2\alpha+3)t_n - 1}] \Rightarrow e^{-\hat{t}_n} \leq (1 + \alpha \hat{s}_n (2\alpha + 3))^{-1/(2\alpha+3)},
$$

(62)
with $a_{\tau} = (1 + \tau)^{(2\alpha + 2)}$ converging to 1 as $\tau \to 0$. For a given solution $\hat{x}_\Delta$ of (58) with $\lambda = 1$ and fixed discretization $\Delta = (\tau; \delta)$, it is a trivial task to check that the recursively defined sequence of vectors $\delta_{S\tau} \hat{x}_\Delta^n$ is a solution to (59) for $S = S_{\tau}^{\alpha - 1}$, $R = S_{\tau}^\alpha$, $\tilde{\lambda} = 0$ and $\tilde{\tau} = \tau_n$ defined in (61). Henceforth, we write $\hat{x}_\Delta^n = \delta_{S\tau} \hat{x}_\Delta^n$ with the discretization $\hat{\Delta} = (\tau; \delta)$. We can hence use the discrete scaling property of $H_{\alpha, \lambda}$ and invoke (40) of Lemma 3.1 which results in

$$(1 + 2\tau_n)(S_{\tau}^n)^{\frac{2a_\alpha - 1}{\alpha}}(H_{\alpha, 1}(\hat{x}_\Delta^n) - H_{\alpha, 1}(\hat{b}_\Delta^n))$$

$$\leq (S_{\tau}^{n-1})^{\frac{2a_\alpha - 1}{\alpha}}(H_{\alpha, 1}(\hat{x}_\Delta^{n-1}) - H_{\alpha, 1}(\hat{b}_\Delta^{n-1}))$$

$$\Rightarrow (1 + 2\tau_n)(1 + \tau_n)^{\frac{2a_\alpha - 1}{\alpha}}(H_{\alpha, 1}(\hat{x}_\Delta^n) - H_{\alpha, 1}(\hat{b}_\Delta^n)) \leq H_{\alpha, 1}(\hat{x}_\Delta^{n-1}) - H_{\alpha, 1}(\hat{b}_\Delta^{n-1})$$

$$\Rightarrow (1 + 2\tau_n)(H_{\alpha, 1}(\hat{x}_\Delta^n) - H_{\alpha, 1}(\hat{b}_\Delta^n)) \leq H_{\alpha, 1}(\hat{x}_\Delta^{n-1}) - H_{\alpha, 1}(\hat{b}_\Delta^{n-1})$$

$$(63)$$

where we used in the last step $1 + \tau_n > 1$. As before in the proof of (18) of Theorem 1.2 this yields for any $n \in \mathbb{N}$, due to (62)

$$(H_{\alpha, 1}(\hat{x}_\Delta^n) - H_{\alpha, 1}(\hat{b}_\Delta^n)) \leq (H_{\alpha, 1}(\hat{x}_\Delta^0) - H_{\alpha, 1}(\hat{x}_\Delta^{\min}))e^{-\frac{2a_\alpha - 1}{\alpha}\tau_n}$$

$$\leq (H_{\alpha, 1}(\hat{x}_\Delta^0) - H_{\alpha, 1}(\hat{x}_\Delta^{\min}))\left(1 + a_{\tau}(2\alpha + 3)\right)^{-\frac{\tau_n^2}{2\tau_n}}$$

with $b_{\tau} = 1 + 2\tau$. Therefore, Theorem 1.3 is a conclusion of the above calculation, the identity $b_{\Delta, \alpha, 0} = u_{\delta}(\hat{b}_{\Delta, \alpha, 0})$, and a Csiszar-Kullback inequality, see [15, Theorem 30].

4. Numerical experiments.

4.1. Non-uniform meshes. An equidistant mass grid — as used in the analysis above — leads to a good spatial resolution of regions where the value of $u^0$ is large, but provides a very poor resolution in regions where $u^0$ is small. Since we are interested in regions of low density, and especially in the evolution of supports, it is natural to use a non-equidistant mass grid with an adapted spatial resolution, like the one defined as follows: The mass discretization of $[0, M]$ is determined by a vector $\delta = (\xi_0, \xi_1, \ldots, \xi_{K-1}, \xi_K)$, with $0 = \xi_0 < \xi_1 < \cdots < \xi_{K-1} < \xi_K = M$ and we introduce accordingly the distances (note the convention $\xi_{-1} = \xi_{K+1} = 0$)

$$\delta_{\kappa} = \xi_{\kappa} + \frac{1}{2} - \xi_{\kappa - \frac{1}{2}}, \quad \delta_{\kappa} = \frac{1}{2}(\delta_{\kappa + \frac{1}{2}} + \delta_{\kappa - \frac{1}{2}}) = \frac{1}{2}(\xi_{k+1} - \xi_{k-1})$$

for $\kappa \in \mathbb{Z}^{K+1}$ and $k \in \mathbb{Z}^0$, respectively. The piecewise constant density function $u \in \mathcal{P}_\delta(\mathbb{R})$ corresponding to a vector $\hat{x} \in \mathbb{R}^{K-1}$ is now given by

$$u(x) = z_{\kappa} \quad \text{for} \quad x_{\kappa - \frac{1}{2}} < x < x_{\kappa + \frac{1}{2}} \quad \text{with} \quad z_{\kappa} = \frac{\delta_{\kappa}}{x_{\kappa + \frac{1}{2}} - x_{\kappa - \frac{1}{2}}}.$$

The Wasserstein-like metric (and its corresponding norm) needs to be adapted as well: The scalar product $\langle \cdot, \cdot \rangle_\delta$ is replaced by

$$\langle \hat{v}, \hat{w} \rangle_\delta = \sum_{k \in \mathbb{Z}^K} \delta_k v_k w_k,$$
and we set \(\|v\|_{\delta} = \sqrt{\langle v, v \rangle_{\delta}}\). Hence the metric gradient \(\nabla_{\delta} f(\vec{x}) \in \mathbb{R}^{K+1}\) of a function \(f: \mathcal{F}_{\delta} \to \mathbb{R}\) at \(\vec{x} \in \mathcal{F}_{\delta}\) is given by

\[
[\nabla_{\delta} f(\vec{x})]_k = \frac{1}{\delta_k} \partial_{x_k} f(\vec{x}).
\]

Otherwise, we proceed as before: The entropy is discretized by restriction, and the discretized information functional is the self-dissipation of the discretized entropy. Explicitly, the resulting fully discrete gradient flow equation

\[
\frac{\vec{x}^n_{\Delta} - \vec{x}^{n-1}_{\Delta}}{\tau} = -\nabla_{\delta} \mathcal{F}_{\alpha,\lambda}(\vec{x}^n_{\Delta})
\]

attains the form

\[
\frac{x^n_k - x^{n-1}_k}{\tau} = \frac{2\alpha}{(2\alpha + 1)^2 \delta_k} \left[ (z^n_{k+\frac{1}{2}})^{\alpha+\frac{1}{2}} [D^2_{\delta}(z^n)^{\alpha+\frac{1}{2}}]_{k+\frac{1}{2}} - (z^n_{k-\frac{1}{2}})^{\alpha+\frac{1}{2}} [D^2_{\delta}(z^n)^{\alpha+\frac{1}{2}}]_{k-\frac{1}{2}} \right] + \lambda x^n_k
\]

with \([D^2_{\delta}(z^n)^{\alpha+\frac{1}{2}}]_{k-\frac{1}{2}} := (z^{\alpha+\frac{1}{2}}_{k+\frac{1}{2}} - 2z^{\alpha+\frac{1}{2}}_{k-\frac{1}{2}} + z^{\alpha+\frac{1}{2}}_{k-\frac{1}{2}})/\delta^2_k\).

4.2. Implementation. To guarantee the existence of an initial vector \(\vec{x}^0_{\Delta} \in \mathcal{F}_{\delta}\), which “reaches” any mass point of \(u^0\), i.e., \([x^0, x^0_K] \subseteq \text{supp}(u^0)\), one has to consider initial density functions \(u^0\) with compact support.

Starting from the initial condition \(\vec{x}^0_{\Delta}\), the fully discrete solution is calculated inductively by solving the implicit Euler scheme (64) for \(\vec{x}^n_{\Delta}\), given \(\vec{x}^{n-1}_{\Delta}\). In each time step, a damped Newton iteration is performed, with the solution from the previous time step as initial guess. More precisely, for given \(\vec{x}^{n-1}_{\Delta}\), we calculate \(\vec{x}^n_{\Delta}\) by means of the following algorithm:

\[
\vec{x} := \vec{x}^{n-1}_{\Delta}
\]

repeat

\[
d\vec{x} := -\left(\partial_{\vec{x}}^2 \mathcal{F}_{\alpha,\lambda}(\vec{x})\right)^{-1} \partial_{\vec{x}} \mathcal{F}_{\alpha,\lambda}(\vec{x})
\]

while \(\vec{x} + d\vec{x}\) not monotone

\[
d\vec{x} := d\vec{x}/2
\]

end

\[
\vec{x} := \vec{x} + d\vec{x}
\]

until \(\|d\vec{x}\|_1 < \text{tol}\) and \(\|\partial_{\vec{x}} \mathcal{F}_{\alpha,\lambda}(\vec{x})\|_1 < \text{tol}\)

\[
\vec{x}^n_{\Delta} := \vec{x}.
\]

In all of our experiments, we use \(\text{tol} = 10^{-8}\). Relatively slow convergence of the Newton iteration has been observed in situations where the density \(u^{n-1}_{\Delta}\) has steep gradients and/or intervals of very low values. In the experiments that follow, the number of Newton iterations is between ten at the very first time iterations and decreases up to one iteration at later times.
Figure 1. Left: Numerically observed decay of $H_{\alpha,\lambda}(t) - H_{\alpha,\lambda}^{\text{min}}$ and $F_{\alpha,\lambda}(t) - F_{\alpha,\lambda}^{\text{min}}$ along a time period of $t \in [0, 0.8]$, using $K = 25, 50, 100, 200$, in comparison to the upper bounds $(H_{\alpha,\lambda}(u^0) - H_{\alpha,\lambda}(b_{\alpha,\lambda})) \exp(-2\lambda t)$ and $(F_{\alpha,\lambda}(u^0) - F_{\alpha,\lambda}(b_{\alpha,\lambda})) \exp(-2\lambda t)$, respectively. Right: Convergence of discrete minimizers $u_{\Delta}^{\text{min}}$ with a rate of $K^{-1.5}$.

4.3. Experiment I – Exponential decay rates.

Figure 2. Evolution of a discrete solution $u_{\Delta}$, evaluated at different times $t = 0, 0.05, 0.1, 0.15, 0.175, 0.25$ (from top left to bottom right). The red line is the corresponding Barenblatt-profile $b_{\alpha,\lambda}$.

As a first numerical experiment, we want to analyse the rate of decay in case of positive confinement $\lambda = 5$, using $\alpha = 1$. For that purpose, consider the initial
density function
\[ u^0 = \begin{cases} 0.25|\sin(x)| \cdot (0.5 + I_{x>0}(x)) & , x \in [-\pi, \pi], \\ 0 & , \text{else} \end{cases} \] (65)

Figure 2 shows the evolution of \( u_\Delta \) at \( t = 0.05, 0.1, 0.15, 0.175, 0.225 \), using \( K = 200 \).

The two initially separated clusters quickly merge, and finally changes the shape towards a Barenblatt-profile (red line).

The exponential decay of the entropies \( H_{\alpha,\lambda} \) and \( F_{\alpha,\lambda} \) along the solution can be seen in Figure 1 left for \( K = 25, 50, 100, 200 \), where we observed the evolution for \( t \in [0,0.8] \). Note that we write \( H_{\alpha,\lambda}(t) = H_{\alpha,\lambda}(\vec{x}_\Delta^0) \) and \( F_{\alpha,\lambda}(t) = F_{\alpha,\lambda}(\vec{x}_\Delta^0) \) for \( t \in (t_n-1, t_n) \), and set \( H_{\alpha,\lambda}(0) = H_{\alpha,\lambda}(\vec{x}_\Delta^0) \) and \( F_{\alpha,\lambda}(0) = F_{\alpha,\lambda}(\vec{x}_\Delta^0) \). As the picture shows, the rate of decay does not really depend on the choice of \( K \), in fact the curves lie de facto on the top of each other. Furthermore, the curves are bounded from above by \( (H_{\alpha,\lambda}(u_0) - H_{\alpha,\lambda}(b_{\alpha,\lambda})) \exp(-2\lambda t) \) and \( (F_{\alpha,\lambda}(u_0) - F_{\alpha,\lambda}(b_{\alpha,\lambda})) \exp(-2\lambda t) \) at any time, respectively, as \([18]\) and \([19]\) from Theorem \([12]\) postulate. One can even recognize that the decay rates are even bigger at the beginning, until the moment \( t = 0.175 \) when \( u_\Delta \) finishes its “fusion” to one single Barenblatt-like curve. After that, the solution’s evolution mainly consists of a transversal shift towards the stationary solution \( b_{\alpha,\lambda} \), which is reflected by a henceforth constant rate of approximately \(-2\lambda\).

Moreover, Figure 2 right pictures the convergence of \( u_0^\min \) towards \( b_{\alpha,\lambda} \). We used several values for the spatial discretization, \( K = 25, 50, 100, 200, 400, 800 \), and plotted the \( L^2 \)-error. The observed rate of convergence is \( K^{-1.5} \).

4.4. Experiment II – Self-similar solutions. A very interesting consequence of Section 3.2 is, that the existence of self-similar solutions bequeath from the continuous to the discrete case. To visualise this feature of the numerical scheme, we are going to plot continuous and discrete solutions at certain times simultaneously.

To this end, set \( \lambda = 0 \) and define for \( t \in [0, +\infty) \)
\[ b^*_{\alpha,0}(t, \cdot) := \partial_R(t) b_{\alpha,1} \quad \text{with} \quad R(t) := (1 + (2\alpha + 3)t)^{1/(2\alpha + 3)}, \] (66)
then \( b^*_{\alpha,0} \) is a solution of the continuous problem (53) with \( u^0 = b^*_{\alpha,0}(0, \cdot) \). In the discrete setting, solutions to (64) with \( \lambda = 0 \) are inductively given by an initial vector \( b^n_{0,\alpha,0} \) with corresponding density \( u_\Delta^n = u^n_{\vec{x}}(b^n_{0,\alpha,0}) \) that approaches \( b^*_{\alpha,0}(0, \cdot) \) and \( \tilde{b}_{\Delta,\alpha,0} = \partial_S b_{\Delta,\alpha,0} \) with \( S^n_{\alpha} \) defined as in \([60]\) for further \( n \in \mathbb{N} \). As Figure 3 shows, the resulting sequence of densities \( u_\Delta \) (black lines) approaches the continuous solution \( b^*_{\alpha,0} \) of (66) (red lines) astonishingly well, even if the discretization parameters are chosen quite rough. In this specific case we used the values \( \alpha = 1, K = 50, \tau = 10^{-3} \). The discrete and continuous solutions are evaluated at times \( t = 0, 0.1, 1, 10, 100 \).

Appendix A. A lemma concerning De Giorgi’s variational interpolation.

Lemma A.1. Fix \( \vec{x} \in \mathfrak{g}_\delta, \tau > 0, \) and let \( V : \mathfrak{g}_\delta \rightarrow (-\infty, +\infty] \). Furthermore, assume for any \( \sigma \in (0, \tau] \) the existence of a minimizer \( \vec{x}_\sigma \) of \( \vec{x} \mapsto \frac{1}{2\sigma} \| \vec{x} - \vec{x}_\sigma \|_\delta^2 + V(\vec{x}) \). Then \( \vec{x}_\sigma \) connecting \( \vec{x} \) and \( \vec{x}_\tau \) satisfies
\[ V(\vec{x}_\sigma) + \frac{\| \vec{x}_\sigma - \vec{x} \|_\delta^2}{2\sigma} + \int_{0}^{\sigma} \frac{\| \vec{x}_r - \vec{x}_\sigma \|_\delta^2}{2r^2} \, dr = V(\vec{x}) \] (67)
for any \( \sigma \in (0, \tau] \).
Figure 3. Snapshots of the densities $b_{\alpha,0}^*(t, \cdot)$ (red lines) and $u_{\Delta}$ (black lines) for the initial condition $b_{\alpha,0}^*(0, \cdot)$ at times $t = 0$ and $t = 0.1 \cdot 10^4$, $i = 0, \ldots, 3$, using $K = 50$ grid points and the time step size $\tau = 10^{-3}$.

Proof. This is a special case of [1, Theorem 3.1.4].

Appendix B. Technical proofs from Section 2.2

Proof of Lemma 2.2. If we prove (29), then the existence of the unique minimizer is a consequence [45, Proposition 10]. By definition and a change of variables, we get for $\alpha \in (\frac{1}{2}, 1]$ $H_{\alpha,0}(\vec{x}) = H_{\alpha,0}(u_{\delta}^{\vec{x}}) = \int_0^M \psi_\alpha (X_{\delta}[\vec{x}]_\xi) \, d\xi$

with $\psi_\alpha(s) = \begin{cases} \Theta_\alpha s^{\frac{1}{2} - \alpha}, & \alpha \in (\frac{1}{2}, 1] \\ -\Theta_{1/2} \ln(s), & \alpha = \frac{1}{2} \end{cases}$, hence $\vec{x} \mapsto H_{\alpha,0}(\vec{x})$ is convex. Since the functional $\vec{x} \mapsto ||\vec{x}||^2_\delta$ trivially satisfies $||\vec{x} + s\vec{y}||^2_\delta \leq (1 - s) ||\vec{x}||^2_\delta + s ||\vec{y}||^2_\delta - (1 - s)s ||\vec{x} - \vec{y}||^2_\delta$

for any $\vec{x}, \vec{y} \in R^d$ and $s \in (0, 1)$, the functionals $H_{\alpha,\lambda}(\vec{x}) = H_{\alpha,0}(\vec{x}) + \frac{\Lambda_{\alpha,\lambda}}{2} ||\vec{x}||^2_\delta$ fulfill (29).

Dividing (29) by $s > 0$ and passing to the limit as $s \downarrow 0$ yields $H_{\alpha,\lambda}(\vec{x}) - H_{\alpha,\lambda}(\vec{y}) \leq \partial_2 H_{\alpha,\lambda}(\vec{x})(\vec{x} - \vec{y}) - \frac{\Lambda_{\alpha,\lambda}}{2} ||\vec{x} - \vec{y}||^2_\delta$.

The second inequality of (30) easily follows from Young’s inequality $|ab| \leq \varepsilon |a|^2 + (2\varepsilon)^{-1} |b|^2$ with $\varepsilon = (25\Lambda_{\alpha,\lambda})^{-1}$, and even stays valid for arbitrary $\vec{y} \in R^d$.

To get the first inequality of (30), we set $\vec{x} = \vec{x}^{\min}_{\delta}$ and again divide (29) by $s > 0$, then $H_{\alpha,\lambda}((1 - s)\vec{x}^{\min}_{\delta} + s\vec{y}) - H_{\alpha,\lambda}(\vec{x}^{\min}_{\delta})$

$s \leq H_{\alpha,\lambda}(\vec{y}) - H_{\alpha,\lambda}(\vec{x}^{\min}_{\delta}) - \frac{\Lambda_{\alpha,\lambda}}{2} (1 - s) ||\vec{x}^{\min}_{\delta} - \vec{y}||^2_\delta$,.
where the left hand side is obviously nonnegative for any \( s > 0 \). Since \( s > 0 \) was arbitrary, the statement is proven. \( \square \)

**Proof of Corollary**\([\square]\). Let us first assume \( \alpha \in (\frac{1}{2}, 1] \). A straightforward calculation using the definition of \( \| \cdot \|_\delta \) and \( \partial_\delta H_{\alpha, \lambda} \) in (28) yields

\[
\| \nabla_\delta H_{\alpha, \lambda}(\bar{x}) \|_\delta^2 = \delta^{-1} \langle \partial_\delta H_{\alpha, \lambda}(\bar{x}), \partial_\delta H_{\alpha, \lambda}(\bar{x}) \rangle = \| \nabla_\delta H_{\alpha, 0}(\bar{x}) \|_\delta^2 - 2\Theta_\alpha \Lambda_{\alpha, \lambda} \delta \sum_{\kappa \in I_{1/2}^2} z_\kappa^{\alpha - 1/2} + \Lambda_{\alpha, \lambda}^2 \delta \sum_{k \in \mathbb{Z}} |x_k|^2. \quad (68)
\]

Here, we used the explicit representation of \( \partial_\delta H_{\alpha, \lambda}(\bar{x}) \), remember (28),

\[
\partial_\delta H_{\alpha, \lambda}(\bar{x}) = \Theta_\alpha \delta \sum_{\kappa \in I_{1/2}^2} z_\kappa^{\alpha + 1/2} \frac{e_\kappa - \frac{1}{\delta}}{\delta} - \frac{e_\kappa}{\delta} + \Lambda_{\alpha, \lambda} \delta \sum_{k \in \mathbb{Z}} x_k e_k,
\]

and especially the definition of (7), which yields

\[
\delta^{-1} \left( \Theta_\alpha \delta \sum_{\kappa \in I_{1/2}^2} z_\kappa^{\alpha + 1/2} \frac{e_\kappa - \frac{1}{\delta}}{\delta} - \frac{e_\kappa}{\delta} + \Lambda_{\alpha, \lambda} \delta \sum_{k \in \mathbb{Z}} x_k e_k \right)
\]

\[
= \Theta_\alpha \Lambda_{\alpha, \lambda} \delta \sum_{\kappa \in I_{1/2}^2} z_\kappa^{\alpha + 1/2} \frac{e_\kappa - \frac{1}{\delta}}{\delta} - \frac{e_\kappa}{\delta} = -\Theta_\alpha \Lambda_{\alpha, \lambda} \delta \sum_{\kappa \in I_{1/2}^2} z_\kappa^{\alpha - 1/2}.
\]

Since \( \alpha \neq \frac{1}{2} \), we can write \( 2\Theta_\alpha = (2\alpha - 1) \frac{\Theta_\alpha}{\alpha - 1/2} \). Further note that the relation \( \Lambda_{\alpha, \lambda} = \sqrt{\lambda/(2\alpha + 1)} \) yields

\[
\Lambda_{\alpha, \lambda}^2 = \frac{\lambda}{2 \alpha + 1} = \frac{\lambda}{2} \left( \frac{1}{\alpha + 1/2} \right) = \frac{\lambda}{2} \left( 1 - \frac{\alpha - 1/2}{\alpha + 1/2} \right) = \frac{\lambda}{2} \left( 1 - \frac{2\alpha - 1}{2\alpha + 1} \right).
\]

Using this information and the definition of \( H_{\alpha, 0} \), we proceed in the above calculations by

\[
\| \nabla_\delta H_{\alpha, \lambda}(\bar{x}) \|_\delta^2 = F_{\alpha, 0}(\bar{x}) - (2\alpha - 1)\Lambda_{\alpha, \lambda} H_{\alpha, 0}(\bar{x}) + \frac{\lambda}{2} \left( 1 - \frac{2\alpha - 1}{2\alpha + 1} \right) \delta \sum_{k \in \mathbb{Z}} |x_k|^2
\]

\[
= F_{\alpha, 0}(\bar{x}) - (2\alpha - 1)\Lambda_{\alpha, \lambda} H_{\alpha, 0}(\bar{x}) + \frac{\lambda}{2} \delta \sum_{k \in \mathbb{Z}} |x_k|^2
\]

\[
- (2\alpha - 1) \frac{\Lambda_{\alpha, \lambda}^2}{2} \delta \sum_{k \in \mathbb{Z}} |x_k|^2
\]

\[
= F_{\alpha, \lambda}(\bar{x}) - (2\alpha - 1)\Lambda_{\alpha, \lambda} H_{\alpha, \lambda}(\bar{x}).
\]

In case of \( \alpha = \frac{1}{2} \), we see that \( \Theta_{1/2} = \frac{1}{2} \), and \( \Lambda_{1/2, \lambda} = \sqrt{\lambda/2} \). We hence conclude in (68) that

\[
\| \nabla_\delta H_{1/2, \lambda}(\bar{x}) \|_\delta^2 = \| \nabla_\delta H_{1/2, 0}(\bar{x}) \|_\delta^2 - \Lambda_{1/2, \lambda} \delta \sum_{\kappa \in I_{1/2}^2} z_\kappa^0 + \frac{\lambda}{2} \delta \sum_{k \in \mathbb{Z}} |x_k|^2
\]

\[
= F_{\alpha, \lambda}(\bar{x}) - \Lambda_{1/2, \lambda}.
\]

\( \square \)
Appendix C. Technical proofs from Section 3.2

Proof of Lemma 3.4. To simplify the proof, we show first that we can assume $S = 1$ without loss of generality, which is because of the following calculation: If for $R > S > 0$ the vector $\partial_R \vec{x}$ minimizes (59), then the linearity of $\| \cdot \|_\delta$ and (57) yield

$$F_\alpha(\hat{\lambda}, \hat{\tau}, \partial_R \vec{x}, \partial_S \vec{x}^*)$$

$$\begin{align*}
&= \frac{S^2}{2\hat{\tau}} \|S^{-1} \partial_R \vec{x} - \vec{x}^*\|_\delta^2 + S^{-2\alpha+1} F_{\alpha,0}(S^{-1} \partial_R \vec{x}) + S^2 \hat{\lambda} \frac{1}{2} \|S^{-1} \partial_R \vec{x}\|_\delta^2 \\
&= S^{-2\alpha+1} \left( \frac{1}{2\hat{\tau} S^{-(2\alpha+3)}} \| \partial_R \vec{x} - \vec{x}^*\|_\delta^2 + F_{\alpha,0}(\partial_R \vec{x}) + S^2 \alpha \frac{1}{2} \| \partial_R \vec{x}\|_\delta^2 \right) \\
&= S^{-2\alpha+1} F_{\alpha}(\hat{\lambda}, \hat{\tau}, \partial_R \vec{x}, \vec{x}^*)
\end{align*}$$

with $\hat{R} = \frac{R}{S} > 1 > 0$ and the new constants

$$\hat{\tau} = \tau S R^{2\alpha+2} S^{-(2\alpha+3)} = \tau \hat{R}^{2\alpha+3},$$

$$\text{and} \quad \hat{\lambda} = S^{2\alpha+3} \frac{(1 + \lambda \tau) - R/S}{\hat{\tau} R/S} = \frac{(1 + \lambda \tau) - \hat{R}}{\hat{\tau}}.$$ 

hence $\partial_R \vec{x}$ minimizes $F_{\alpha}(\hat{\lambda}, \hat{\tau}, \partial_R \vec{x}, \vec{x}^*)$.

So assume $S = 1$ and $R > 1$ in (59) by now. Further introduce the functional $g : \mathfrak{F}_\delta \times \mathbb{R} \to \mathbb{R}$ through

$$g(\vec{y}, r) := \frac{1}{2} \| \partial_r \vec{y} - \vec{x}^*\|_\delta^2 + r F_{\alpha,0}(\vec{y}) + \frac{1}{2} (1 + \lambda \tau - r) \| \vec{y}\|_\delta^2,$$

then by definition

$$\tau^{-1} g(\vec{y}, 1) = F_{\alpha}(\lambda, \tau, \vec{y}, \vec{x}^*), \quad \text{and} \quad (\tau R^{2\alpha+2})^{-1} g(\vec{y}, R) = F_{\alpha}(\hat{\lambda}, \hat{\tau}, \partial_R \vec{y}, \vec{x}^*) \quad \text{(69)}.$$

For fixed $\vec{y} \in \mathfrak{F}_\delta$, a straight-forward calculation shows that the derivative of $r \mapsto g(\vec{y}, r)$ fulfills

$$\partial_r g(\vec{y}, r) = \langle \partial_r \vec{y} - \vec{x}^*, \vec{y}\rangle_\delta + F_{\alpha,0}(\vec{y}) - \frac{r}{2} \| \vec{y}\|_\delta^2 + \frac{1}{2} (1 + \lambda \tau - r) \| \vec{y}\|_\delta^2$$

$$= - \langle \vec{x}^*, \vec{y}\rangle_\delta + F_{\alpha,0}(\vec{y}) + \frac{1}{2} (1 + \lambda \tau) \| \vec{y}\|_\delta^2$$

$$= \frac{1}{2} \| \vec{y} - \vec{x}^*\|_\delta^2 - \frac{1}{2} \| \vec{x}^*\|_\delta^2 + F_{\alpha,0}(\vec{y}) + \frac{\lambda \tau}{2} \| \vec{y}\|_\delta^2$$

$$= g(\vec{y}, 1) - \frac{1}{2} \| \vec{x}^*\|_\delta^2.$$

Hence, if $\vec{x}$ minimizes (58), then the same vector minimizes $\vec{y} \mapsto g(\vec{y}, 1)$ and further $\vec{y} \mapsto \partial_r g(\vec{y}, r)$ for any $r > 0$. By integration

$$g(\vec{y}, r) - g(\vec{y}, 1) = \int_1^r \partial_s g(\vec{y}, s) \, ds = (r - 1)(g(\vec{y}, 1) - \frac{1}{2} \| \vec{x}^*\|_\delta^2)$$

$$\implies g(\vec{y}, r) = rg(\vec{y}, 1) - (r - 1) \frac{1}{2} \| \vec{x}^*\|_\delta^2$$

for any $r > 1$ and $\vec{y} \in \mathfrak{F}_\delta$. This means especially that for arbitrary $r > 1$, the function $g(\vec{y}, r)$ is minimal if and only if $g(\vec{y}, 1)$ is so. This proves in combination with (69) that $\partial_R \vec{x}$ is a minimizer of (59). By integration of $\partial_s g(\vec{y}, s)$ over $[r^{-1}, 1]$, $r > 1$, one can analogously prove that if $\vec{x} \in \mathfrak{F}_\delta$ is a minimizer of (59), the rescaled vector $\partial_{R^{-1}} \vec{x}$ has to be a minimizer of (58). □
REFERENCES

[1] L. Ambrosio, N. Gigli and G. Savaré, *Gradient Flows in Metric Spaces and in the Space of Probability Measures*, Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, Basel, 2005.

[2] L. Ambrosio, S. Lisini and G. Savaré, Stability of flows associated to gradient vector fields and convergence of iterated transport maps, *Manuscripta Mathematica*, 121 (2006), 1–50.

[3] J. Becker and G. Grün, The thin-film equation: Recent advances and some new perspectives, *Journal of Physics: Condensed Matter*, 17 (2015), 291–307. Available from: http://iopscience.iop.org/0953-8984/17/9/002/pdf/0953-8984_17_9_002.pdf

[4] F. Bernis and F. Avner, Higher order nonlinear degenerate parabolic equations, *Journal of Differential Equations*, 83 (1990), 179–206. Available from: http://dx.doi.org/10.1016/0022-0396(90)90074-Y

[5] M. Bertsch, R. Dal Passo, H. Garcke and G. Grün, The thin viscous flow equation in higher space dimensions, *Journal of Differential Equations*, 3 (1998), 417–440.

[6] A. Blanchet, V. Calvez and J. A. Carrillo, Convergence of the mass-transport steepest descent scheme for the subcritical Patlak-Keller-Segel model, *SIAM Journal on Numerical Analysis*, 46 (2008), 691–721.

[7] P. M. Bleher, J. L. Lebowitz and E. R. Speer, Existence and positivity of solutions of a fourth-order nonlinear PDE describing interface fluctuations, *Communications on Pure and Applied Mathematics*, 47 (1994), 923–942.

[8] A. Braides, *Γ-convergence for Beginners*, Oxford Lecture Series in Mathematics and its Applications, Oxford University Press, Oxford, 2002.

[9] C. J. Budd, G. J. Collins, W. Z. Huang and R. D. Russell, Self-similar numerical solutions of the porous-medium equation using moving mesh methods, *The Royal Society of London. Philosophical Transactions. Series A. Mathematical, Physical and Engineering Sciences*, 357 (1999), 1047–1077.

[10] M. Bukal, E. Emmrich and A. Jüngel, Entropy-stable and entropy-dissipative approximations of a fourth-order quantum diffusion equation, *Numerische Mathematik*, 127 (2014), 365–396.

[11] M. Burger, J. A. Carrillo and M.-T. Wolfram, A mixed finite element method for nonlinear diffusion equations, *Kinetic and Related Models*, 3 (2010), 59–83.

[12] M. J. Cáceres, J. A. Carrillo and G. Toscani, Long-time behavior for a nonlinear fourth-order parabolic equation, *Transactions of the American Mathematical Society*, 357 (2005), 1161–1175.

[13] E. A. Carlen and S. Ulusoy, Asymptotic equipartition and long time behavior of solutions of a thin-film equation, *Journal of Differential Equations*, 241 (2007), 279–292.

[14] J. A. Carrillo, J. Dolbeault, I. Gentil and A. Jüngel, Entropy-energy inequalities and improved convergence rates for nonlinear parabolic equations, *Discrete and Continuous Dynamical Systems. Series B. A Journal Bridging Mathematics and Sciences*, 6 (2006), 1027–1050.

[15] J. A. Carrillo, A. Jüngel, P. A. Markowich, G. Toscani and A. Unterreiter, Entropy dissipation methods for degenerate parabolic problems and generalized Sobolev inequalities, *Monatshefte für Mathematik*, 133 (2001), 1–82.

[16] J. A. Carrillo, A. Jüngel and S. Tang, Positive entropic schemes for a nonlinear fourth-order parabolic equation, *Discrete and Continuous Dynamical Systems. Series B. A Journal Bridging Mathematics and Sciences*, 3 (2003), 1–20.

[17] J. A. Carrillo and J. S. Moll, Numerical simulation of diffusive and aggregation phenomena in nonlinear continuity equations by evolving diffeomorphisms, *SIAM Journal on Scientific Computing*, 31 (2009/10), 4305–4329.

[18] J. A. Carrillo and G. Toscani, Long-time asymptotics for strong solutions of the thin film equation, *Communications in Mathematical Physics*, 225 (2002), 551–571.

[19] J. A. Carrillo and M.-T. Wolfram, A finite element method for nonlinear continuity equations in Lagrangian coordinates, work in progress.

[20] F. Cavalli and G. Naldi, A Wasserstein approach to the numerical solution of the one-dimensional Cahn-Hilliard equation, *Kinetic and Related Models*, 3 (2010), 123–142.

[21] R. Dal Passo, H. Garcke and G. Grün, On a fourth-order degenerate parabolic equation: Global entropy estimates, existence, and qualitative behavior of solutions, *SIAM Journal on Mathematical Analysis*, 29 (1998), 321–342.

[22] J. Denzler and R. J. McCann, Nonlinear diffusion from a delocalized source: affine self-similarity, time reversal, & nonradial focusing geometries, *Annales de l’Institut Henri Poincaré. Analyse Non Linéaire*, 25 (2008), 865–888.
B. Derrida, J. L. Lebowitz, E. R. Speer and H. Spohn. Dynamics of an anchored Toom interface. *Journal of Physics. A. Mathematical and General*, 24 (1991), 4895–4894.

B. Derrida, J. L. Lebowitz, E. R. Speer and H. Spohn. Fluctuations of a stationary nonequilibrium interface. *Physical Review Letters*, 67 (1991), 165–168.

B. Düring, D. Matthes and J. P. Milišić. A gradient flow scheme for nonlinear fourth order equations. *Discrete and Continuous Dynamical Systems. Series B. A Journal Bridging Mathematics and Sciences*, 14 (2010), 935–959.

L. C. Evans, O. Savin and W. Gangbo. Diffeomorphisms and nonlinear heat flows. *SIAM Journal on Mathematical Analysis*, 37 (2005), 737–751.

J. Fischer. Uniqueness of solutions of the Derrida-Lebowitz-Speer-Spohn equation and quantum drift-diffusion models. *SIAM Journal on Mathematical Analysis*, 38 (2013), 2004–2047.

L. Giacomelli and F. Otto. Variational formulation for the lubrication approximation of the Hele-Shaw flow. *Calculus of Variations and Partial Differential Equations*, 13 (2001), 377–403.

U. Gianazza, G. Savaré and G. Toscani. The Wasserstein gradient flow of the Fisher information and the quantum drift-diffusion equation. *Archive for Rational Mechanics and Analysis*, 194 (2009), 133–220.

E. Giusti. *Minimal Surfaces and Functions of Bounded Variation*. Monographs in Mathematics, Birkhäuser Verlag, Basel, 1984.

L. Gosse and G. Toscani. Identification of asymptotic decay to self-similarity for one-dimensional filtration equations. *SIAM Journal on Numerical Analysis*, 43 (2006), 2590–2606.

L. Gosse and G. Toscani. Lagrangian numerical approximations to one-dimensional convolution-diffusion equations. *SIAM Journal on Scientific Computing*, 28 (2006), 1203–1227.

G. Grün. Droplet spreading under weak slippage—existence for the Cauchy problem. *Communications in Partial Differential Equations*, 29 (2004), 1697–1744.

M. P. Gualdani, A. Jüngel and G. Toscani. A nonlinear fourth-order parabolic equation with nonhomogeneous boundary conditions. *SIAM Journal on Mathematical Analysis*, 37 (2006), 1761–1779.

R. Jordan, D. Kinderlehrer and F. Otto. The variational formulation of the Fokker-Planck equation. *SIAM Journal on Mathematical Analysis*, 29 (1998), 1–17.

A. Jüngel and D. Matthes. An algorithmic construction of entropies in higher-order nonlinear PDEs. *Nonlinearity*, 19 (2006), 633–659.

A. Jüngel and D. Matthes. The Derrida-Lebowitz-Speer-Spohn equation: Existence, nonuniqueness, and decay rates of the solutions. *SIAM Journal on Mathematical Analysis*, 39 (2008), 1996–2015.

A. Jüngel and R. Pinnau. Global nonnegative solutions of a nonlinear fourth-order parabolic equation for quantum systems. *SIAM Journal on Mathematical Analysis*, 32 (2000), 760–777.

A. Jüngel and R. Pinnau. A positivity-preserving numerical scheme for a nonlinear fourth order parabolic system. *SIAM Journal on Numerical Analysis*, 39 (2001), 385–406.

A. Jüngel and G. Toscani. Exponential time decay of solutions to a nonlinear fourth-order parabolic equation. *Journal of Applied Mathematics and Physics*, 54 (2003), 377–386.

A. Jüngel and I. Violet. First-order entropies for the Derrida-Lebowitz-Speer-Spohn equation. *Discrete and Continuous Dynamical Systems. Series B. A Journal Bridging Mathematics and Sciences*, 8 (2007), 861–877.

D. Kinderlehrer and N. J. Walkington. Approximation of parabolic equations using the Wasserstein metric. *M2AN. Mathematical Modelling and Numerical Analysis*, 33 (1999), 837–852.

R. C. MacCamy and E. Socolovsky. A numerical procedure for the porous media equation. *Computers & Mathematics with Applications. An International Journal*, 11 (1985), 315–319.

D. Matthes, R. J. McCann and G. Savaré. A family of nonlinear fourth order equations of gradient flow type. *Communications in Partial Differential Equations*, 34 (2009), 1352–1397.

D. Matthes and H. Osberger. Convergence of a variational Lagrangian scheme for a nonlinear drift diffusion equation. *ESAIM. Mathematical Modelling and Numerical Analysis*, 48 (2014), 697–726.

D. Matthes and H. Osberger. A convergent Lagrangian discretization for a nonlinear fourth-order equation. *Foundations of Computational Mathematics*, 17 (2015), 1–54.
[47] A. Oron, S. H. Davis and S. G. Bankoff, Long-scale evolution of thin liquid films, American Physical Society, 69 (1997), 931–980.

[48] G. Russo, Deterministic diffusion of particles, Communications on Pure and Applied Mathematics, 43 (1990), 697–733.

[49] C. Villani, Topics in Optimal Transportation, American Mathematical Society, Providence, RI, 2003.

Received January 2015; revised August 2016.

E-mail address: osberger@ma.tum.de