A Convergent Lagrangian Discretization for a Nonlinear Fourth-Order Equation

Daniel Matthes · Horst Osberger

Abstract A fully discrete Lagrangian scheme for numerical solution of the nonlinear fourth-order DLSS equation in one space dimension is analyzed. The discretization is based on the equation’s gradient flow structure in the $L^2$-Wasserstein metric. By construction, the discrete solutions are strictly positive and mass conserving. A further key property is that they dissipate both the Fisher information and the logarithmic entropy. Our main result is a proof of convergence of fully discrete to weak solutions in the limit of vanishing mesh size. Convergence is obtained for arbitrary nonnegative, possibly discontinuous initial data with finite entropy, without any CFL-type condition. The key estimates in the proof are derived from the dissipations of the two Lyapunov functionals. Numerical experiments illustrate the practicability of the scheme.

Keywords Lagrangian discretization · Gradient flow · Wasserstein metric · Quantum drift diffusion

Mathematics Subject Classification Primary 65M12; Secondary 35A15 · 35G31

Communicated by Eitan Tadmor.

This research was supported by the DFG Collaborative Research Center TRR 109, “Discretization in Geometry and Dynamics”.

Daniel Matthes
matthes@ma.tum.de
Horst Osberger
osberger@ma.tum.de

1 Zentrum Mathematik, Technische Universität München, Boltzmannstr. 3, 85748 Garching, Germany
1 Introduction

1.1 The Equation and Its Properties

In this paper, we study a full discretization of the following initial boundary value problem on the one-dimensional interval $[a, b]$:

\[
\partial_t u + 2\partial_x \left( u\partial_x \left( \frac{\partial_{xx} \sqrt{u}}{\sqrt{u}} \right) \right) = 0 \quad \text{for } t > 0 \text{ and } x \in [a, b],
\]

(1)

\[
\partial_x u = 0, \quad u\partial_x \left( \frac{\partial_{xx} \sqrt{u}}{\sqrt{u}} \right) = 0 \quad \text{for } t > 0 \text{ and } x \in \{a, b\},
\]

(2)

\[
u = u^0 \text{ at } t = 0.
\]

(3)

Equation (1) is known as the DLSS equation, where the acronym refers to Derrida, Lebowitz, Speer and Spohn, who introduced (1) in [19,20] for studying interface fluctuations in the anchored Toom model. In the context of semiconductor physics, (1) appears as a simplified quantum drift diffusion equation [17,34].

The analytical treatment of (1) is far from trivial: see, e.g., [5,23,26,32,36,37] for results on existence and uniqueness of solutions in various different settings, and [9,10,14,26,36,39,43] for qualitative and quantitative descriptions of the long-time behavior. The main difficulty in the development of the time global well-posedness theory has been that the nonlinear operator in (1) is defined only for positive functions $u$, but there is no comparison principle available, which would provide an a priori positive lower bound on $u$. Ironically, solutions are known to be $C^\infty$-smooth as long as they remain strictly positive [5], but the question if strict positivity of the initial datum $u^0$ is sufficient for that remains open, despite much effort and some recent progress in that direction, see [24]. In order to deal with the general case—allowing arbitrary nonnegative initial data $u^0$ of finite entropy—a theory for nonnegative weak solutions has been developed [26,36] on grounds of the a priori regularity estimate

\[
\sqrt{u} \in L^2_{\text{loc}}(\mathbb{R}^N_+; H^2([a, b])),
\]

(4)

which gives a meaning to (1) in the formally equivalent representation

\[
\partial_t u + \partial_{xxx} u - 4\partial_{xx} (\partial_x \sqrt{u})^2 = 0.
\]

(5)

What eventually paved the way to a rigorous analysis of the DLSS equation are several remarkable structural properties. The most important ones are as follows:

- The evolution is mass preserving,

\[
\int_a^b u(t; x) \, dx = M := \int_a^b u^0(x) \, dx \quad \text{for all } t > 0.
\]
There exist infinitely many (formal) Lyapunov functionals [5, 9, 35]. The two most significant ones are the (logarithmic) entropy,

\[ \mathcal{H}(u) = \int_a^b u \ln u \, dx - \mathcal{H}_0 \quad \text{with} \quad \mathcal{H}_0 = M \ln \left( \frac{M}{b - a} \right), \]  

and the Fisher information,

\[ \mathcal{F}(u) = 2 \int_a^b \left( \partial_x \sqrt{u} \right)^2 \, dx. \]  

The Fisher information is more than just a Lyapunov functional: In [26], it has been shown that (1)–(2) is a gradient flow in the potential landscape of \( \mathcal{F} \) with respect to the \( L^2 \)-Wasserstein metric \( \mathcal{W} \). That is, formally one can write (1)–(2) as

\[ \partial_t u = - \text{grad}_{\mathcal{W}} \mathcal{F}(u). \]  

Also, \( \mathcal{H} \) is not “an arbitrary” Lyapunov functional: the \( L^2 \)-Wasserstein gradient flow of \( \mathcal{H} \) is the heat equation [33],

\[ \partial_s v = - \text{grad}_{\mathcal{W}} \mathcal{H}(v) = \partial_{xx} v, \]  

and the Fisher information \( \mathcal{F} \) equals the dissipation of \( \mathcal{H} \) along its own gradient flow,

\[ \mathcal{F}(v(s)) = - \frac{1}{2} \frac{d}{ds} \mathcal{H}(v(s)). \]  

In view of (8), this relation makes the DLSS equation the “big brother” of the heat equation, see [18, 43] for structural consequences.

1.2 Fully Discrete Approximation

For the numerical approximation of solutions to (1)–(3), it is natural to ask for structure preserving discretizations that inherit at least some of the nice properties listed above. At the very least, the scheme should produce nonnegative (preferably positive) discrete solutions, but there is no reason to expect that behavior from a standard discretization approach. Several (semi-)discretizations for (1)–(3) that guarantee positivity have been proposed in the literature [7, 11, 38, 40]. In all of them, positivity actually appears as a consequence of another, more fundamental feature: each of these schemes also inherits a Lyapunov functional, either a logarithmic/power-type entropy [7, 11, 38], or a variant of the Fisher information [7, 21, 40]. An exception is the discretization from [21], which preserves the Lagrangian representation of (1), see below, and thus enforces positivity by construction. Apparently, at least some structure preservation seems necessary to obtain an acceptable numerical scheme.
Here, we follow further the ansatz from [21], which lead to a discretization with a very rich structure: the scheme there is positivity and mass preserving, it dissipates the Fisher information, it has the same Lagrangian structure as (1), and it even inherits (in a certain sense) the gradient flow structure (8). However, a crucial property of (1) is apparently lost in that discretization, namely the monotonicity of the logarithmic entropy (6). The dissipation of $\mathcal{H}$ is the key ingredient in the derivation of the regularity (4) for (1). Without a properly discretized version of it, it seems unlikely that a rigorous analysis, like the continuous limit, can be carried out.

The main difference of the ansatz followed here to the one from [21] is the discretization of the Fisher information. Instead of the immediate “discretization by restriction,” we take an indirect approach: First, we discretize—by restriction—our auxiliary functional $\mathcal{H}$; then, we define the discretization of the primary functional $\mathcal{F}$ by using (9); see (14) below. This way, we obtain a new scheme which still has all of the aforementioned properties, but in addition also dissipates a discrete variant of the entropy $\mathcal{H}$. Thus, we have two discrete Lyapunov functionals at our disposal, and the interplay between these allows us to give a proof of convergence in the limit of vanishing mesh size.

To the best of our knowledge, our scheme is the first one to preserve the structural properties of (1) to this extent, and in particular the first one that admits simultaneously two Lyapunov functionals. Moreover, it is the only full discretization for which a rigorous convergence analysis is available.

The idea to preserve several Lyapunov functionals under the discretization and to use them for the derivation of estimates on the discrete approximations has been used in the context of finite element (“Eulerian”) schemes for various other evolution equations before. In particular, we mention the convergence results [30,31,48] for the formally similar fourth-order thin film equation. We are not aware of any related results in the context of Lagrangian discretizations.

Below, we give the “pragmatic” definition of our full discretization, which is actually very simple. In Sect. 2, we show how this scheme arises from a structure preserving discretization of the gradient flow structure. The starting point is the Lagrangian representation of (1)–(2). Since each $u(t; \cdot)$ is of mass $M$, there is a Lagrangian map $X(t; \cdot): [0, M] \to [a, b]$—the so-called pseudo-inverse distribution function of $u(t; \cdot)$—such that

$$\xi = \int_0^X u(t; x) \, dx, \quad \text{for each } \xi \in [0, M]. \quad (10)$$

Written in terms of $X$, the Wasserstein gradient flow (8) for $\mathcal{F}$ turns into an $L^2$-gradient flow for

$$\mathbf{F}(X) = \frac{1}{2} \int_0^M \left[ \partial_\xi \left( \frac{1}{\partial_\xi X} \right) \right]^2 \, d\xi. \quad (11)$$
\[ \partial_t X = \partial_\xi \left( Z^2 \partial_\xi Z \right), \quad \text{where} \quad Z(t; \xi) := \frac{1}{\partial_\xi X(t; \xi)} = u(t; X(t; \xi)). \quad (12) \]

At this point, a standard discretization of (12) with parameter \( \Delta = (\tau; \delta) \) is performed: We use the implicit Euler method for time discretization with fixed time step \( \tau > 0 \), and central finite differences for equidistant discretization on the mass space \([0, M]\) with mesh width \( \delta > 0 \). More explicitly: denote by \( \vec{x}_{\Delta}^{n} = (x^n_k) \) a fully discrete solution on the \( \Delta \)-mesh, so that \( x^n_k \) approximates \( X(n\tau; k\delta) \), then the \( x^n_k \) satisfy

\[ \frac{x^n_k - x^n_{k-1}}{\tau} = \frac{1}{\delta} \left[ \left( \frac{z^n_{k+\frac{1}{2}}}{} \right)^2 \left( \frac{z^n_{k+\frac{3}{2}} - 2z^n_{k+\frac{1}{2}} + z^n_{k-\frac{1}{2}}}{\delta^2} \right) \right. \\
\left. - \left( \frac{z^n_{k-\frac{1}{2}}}{} \right)^2 \left( \frac{z^n_{k+\frac{1}{2}} - 2z^n_{k-\frac{1}{2}} + z^n_{k-\frac{3}{2}}}{\delta^2} \right) \right], \quad (13) \]

where the values \( \frac{z^n_{k-\frac{1}{2}}}{} = \delta/(x^n_k - x^n_{k-1}) \) are associated to the midpoints of the mass grid. At each time step \( n \in \mathbb{N} \), \( \vec{x}_{\Delta}^{n} = (x^n_1, \ldots, x^n_{K-1}) \) approximates a Lagrangian map, so we assume that \( \vec{x}_{\Delta}^{n} \) is monotone, i.e., \( x^n_k > x^n_{k-1} \), and in accordance with (10), we associate with \( \vec{x}_{\Delta}^{n} \) a piecewise constant function \( \vec{u}_{\Delta}^{n} : [a, b] \to \mathbb{R}_{>0} \) with

\[ \vec{u}_{\Delta}^{n}(x) = \frac{\delta}{x^n_k - x^n_{k-1}} \quad \text{for} \quad x^n_{k-1} < x \leq x^n_k. \]

As replacements for the entropy \( \mathcal{H} \) and the Fisher information \( \mathcal{F} \), we introduce

\[ \mathcal{H}_\delta(\vec{x}^n) = \delta \sum_{k=1}^{K} \ln \frac{z^n_{k-\frac{1}{2}}}{}, \quad \mathcal{F}_\delta(\vec{x}^n) = \frac{1}{2} \sum_{k=1}^{K-1} \left( \frac{z^n_{k+\frac{1}{2}} - z^n_{k-\frac{1}{2}}}{\delta} \right)^2. \quad (14) \]

These choices are made such that \( \mathcal{H}_\delta \) is the restriction of \( \mathcal{H} \), i.e., \( \mathcal{H}_\delta(\vec{x}_{\Delta}^{n}) = \mathcal{H}(\vec{u}_{\Delta}^{n}) \), and such that \( \mathcal{F}_\delta \) is related to \( \mathcal{H}_\delta \) in the same way (9) as \( \mathcal{F} \) is related to \( \mathcal{H} \); see Sect. 2 for details.

1.3 Results

Our first result is concerned with qualitative properties of the discrete solutions. For the moment, fix a discretization parameter \( \Delta = (\tau; \delta) \).

**Theorem 1** *From any monotone discrete initial datum \( \vec{x}_{\Delta}^{0} \), a sequence of monotone \( \vec{x}_{\Delta}^{n} \) satisfying (13) can be constructed by inductively defining \( \vec{x}_{\Delta}^{n} \) as a global minimizer of \( \mathcal{H}_\delta(\vec{x}^n) + \mathcal{F}_\delta(\vec{x}^n) \) with respect to the mass, subject to the constraint \( \vec{x}_{\Delta}^{n}(a) = \vec{x}_{\Delta}^{0} \).*
of
\[ \vec{x} \mapsto \frac{\delta}{2\tau} \sum_k (x_k - x_{n-1}^k)^2 + F_\delta(\vec{x}). \] (15)

This sequence of vectors \( \vec{x}_\Delta^n \) and the associated densities \( \bar{u}_\Delta^n \) have the following properties:

- **Positivity:** \( \bar{u}_\Delta^n \) is a strictly positive function.
- **Mass conservation:** \( \bar{u}_\Delta^n \) has mass equal to \( M \).
- **Dissipation:** Both the entropy and the discrete Fisher information are dissipated,
  \[ H_\delta(\vec{x}_\Delta^n) \leq H_\delta(\vec{x}_\Delta^{n-1}) \quad \text{and} \quad F_\delta(\vec{x}_\Delta^n) \leq F_\delta(\vec{x}_\Delta^{n-1}). \]
- **Equilibration:** There is a constant \( r > 0 \) only depending on \( b - a \) such that
  \[ H_\delta(\vec{x}_\Delta^n) \leq H_\delta(\vec{x}_0) e^{-r n \tau}. \] (16)

Some of these properties follow immediately from the construction, while others (like the equilibration) are difficult to prove. Note that even well posedness (which involves existence of a monotone minimizer for the functional) is a non-trivial claim.

To state our main result about convergence, we need to introduce the time interpolation \( \{\bar{u}_\Delta\}_\tau : \mathbb{R}_{>0} \times [a, b] \to \mathbb{R}_{>0} \), which is given by
\[ \{\bar{u}_\Delta\}_\tau (t; x) = \bar{u}_\Delta^n (x) \quad \text{for} \quad (n-1)\tau < t \leq n\tau. \]

Further, \( \Delta \) symbolizes a whole sequence of mesh parameters from now on, and we write \( \Delta \to 0 \) to indicate that \( \tau \to 0 \) and \( \delta \to 0 \) simultaneously. We do not assume any relation (CFL-type condition) between \( \tau \) and \( \delta \).

**Theorem 2** Let a nonnegative initial condition \( u^0 \) with \( \mathcal{H}(u^0) < \infty \) be given. Choose initial conditions \( \vec{x}_\Delta^0 \) such that \( \bar{u}_\Delta^0 \) converges to \( u^0 \) weakly as \( \Delta \to 0 \), and
\[ \mathcal{H} := \sup_\Delta H_\delta(\vec{x}_\Delta^0) < \infty \quad \text{and} \quad \lim_{\Delta \to 0} (\tau + \delta) F_\delta(\vec{x}_\Delta^0) = 0. \] (17)

For each \( \Delta \), construct a discrete approximation \( \vec{x}_\Delta \) according to the procedure described in Theorem 1 above. Then, there are a subsequence with \( \Delta \to 0 \) and a limit function \( u_* \in C(\mathbb{R}_{>0} \times [a, b]) \) such that:

- \( \{\bar{u}_\Delta\}_\tau \) converges to \( u_* \) locally uniformly on \( \mathbb{R}_{>0} \times [a, b] \),
- \( \sqrt{\bar{u}_\Delta} \in L^2_{\text{loc}}(\mathbb{R}_{>0}; H^1([a, b])) \),
- there are non-increasing functions \( f, h : \mathbb{R}_{>0} \to \mathbb{R} \) such that \( F(u_*(t)) = f(t) \) and \( \mathcal{H}(u_*(t)) = h(t) \) for a.e. \( t > 0 \), and additionally \( h(t) \leq \mathcal{H} e^{-rt} \) with the constant \( r > 0 \) from (16),

\[ \mathcal{H} \leq \mathcal{H} e^{-rt} \]
• \( u^* \) satisfies the following weak formulation of (1)–(2), see (5):

\[
\int_0^\infty \int_a^b \left[ \partial_t \varphi u^* + \partial_{xx} \varphi \partial_x u^* + 4 \partial_{xx} \varphi \left( \partial_x \sqrt{u^*} \right)^2 \right] dx \, dt \\
+ \int_a^b \varphi(0; x) u^0(x) \, dx = 0
\]  

(18)

for every test function \( \varphi \in C_\infty^\infty(\mathbb{R}_{\geq 0} \times [a, b]) \) satisfying \( \partial_x \varphi(t; a) = \partial_x \varphi(t; b) = 0 \).

**Remark 3**

(1) **Quality of convergence** Since \( \{\bar{u}_\Delta\}_\tau \) is piecewise constant in space and time, uniform convergence is obviously the best kind of convergence that can be achieved.

(2) **Rate of convergence** The scheme (13) is formally consistent of order \( \tau + \delta^2 \), see Proposition 25, and this is also the observed rate of convergence in numerical experiments with smooth initial data \( u^0 \), see Sect. 6.2.5.

(3) **No CFL condition** There is no restriction on the ratio at which \( \tau \) and \( \delta \) tend to zero. Even the second condition in (17) above is rather a restriction on the initial approximation \( \bar{x}_{\Delta x}^0 \) than on \( \tau \). Note, however, that for “large” \( \tau \) (relative to \( \delta \) the minimizer \( \bar{x}_{\Delta x}^n \) in (15) might not be the unique solution to (13); see Proposition 9 for details.

(4) **Equidistant grid** Our convergence result is proven for equidistant grids in time and “mass space.” We believe that it carries over verbatim to non-equidistant grids of uniformly bounded mesh ratio, but the proof will require a significant additional technical effort. Notice that our numerical experiments in Sect. 6 indicate convergence for non-uniform meshes.

(5) **Initial condition** Our only hypothesis on \( u^0 \) is \( \mathcal{H}(u^0) < \infty \), which allows the same general initial conditions as in [26, 36]. If \( \mathcal{F}(u^0) \) happens to be finite, and also \( \sup_{\Delta} \mathcal{F}_{\delta}(\bar{x}_{\Delta x}^0) < \infty \), then the uniform convergence of \( \{\bar{u}_\Delta\}_\tau \) holds up to \( t = 0 \).

(6) **Long-time behavior** By means of the Csiszar–Kullback inequality, the exponential decay of \( \mathcal{H}(u^* \tau)) \) to zero implies exponential convergence of \( u^* \) to the constant function \( u_\infty \equiv M/(b - a) \) in \( L^1([a, b]) \).

(7) **No uniqueness** Our notion of solution is too weak to apply the uniqueness result from [23]; hence, different subsequences of \( \{\bar{u}_\Delta\}_\tau \) might converge to different limits.

Lagrangian discretizations for numerical solution of Wasserstein gradient flows are not new in the literature, see, e.g., [41] for a general treatise. Indeed, several practical schemes have been developed on the basis of a Lagrangian formulation, mainly for second-order diffusion equations [6, 8, 42, 45], but also for chemotaxis systems [4], for non-local aggregation equations [12, 15], and for variants of the Boltzmann equation [29]. For certain nonlinear fourth-order equations, Lagrangian numerical schemes exist as well, e.g., for the Hele–Shaw flow [29] and for the Cahn–Hilliard equation [16].

On the other hand, contributions to the mathematical analysis of stability and convergence of these Lagrangian schemes are rare. There are several results available for
semi-discrete Lagrangian approximations, see, e.g., [2,22]. For fully discrete schemes, however, rigorous results are apparently limited to the case of second-order diffusion in one space dimension, see [28,44] and references therein. A noteworthy exception is the recent work [3], which is also concerned with the rigorous consistency analysis of discretizations for second-order diffusion equations, but where the limitation to one space dimension has been overcome. We are not aware of any previous convergence results for Lagrangian discretizations of fourth-order equations, even in one space dimension.

The primary challenge in our analysis is to carry out all estimates under no additional assumptions on the regularity of the limit solution $u^\ast$. In particular, we do not exclude a priori the formation of zeros—and the induced loss of regularity—in the limit $u^\ast$, since this cannot be excluded by the existing theory. Also, we allow extremely general initial conditions $u^0$. Without sufficient a priori smoothness, we cannot simply use Taylor approximations and the like to estimate the difference between $\{\bar{u}_\Delta\}^\tau$ and $u^\ast$. Instead, we are forced to derive new a priori estimates directly from the scheme, using our two Lyapunov functionals.

On the technical level, the main difficulty is that our scheme is fully discrete, which means that we are working with spatial difference quotients instead of derivatives. Lacking a discrete chain rule, the full equivalence between the original (1) and the Lagrangian (12) formulation of the DLSS equation is destroyed under discretization. This forces us to switch back and forth between the two formulations and thus makes the derivation of the relevant estimates much harder than for the original problem (1)–(3).

We further remark that the convergence of a family of gradient flows to a limiting gradient flow has been thoroughly investigated on an abstract level, see, e.g., in [1,46], using methods of $\Gamma$-convergence. Unfortunately, these appealing abstract results would not simplify our proof significantly (if at all), since the verification of their main hypothesis ($\Gamma$-convergence of the subdifferentials) is essentially equivalent to the derivation of the a priori estimates, which is the main part of our work. Our “hands-on proof” below requires only a couple of basic elements from the general theory of metric gradient flows.

Finally, a comment is in place on the spatial dimension. Among the manifold simplifications that result from working on a one-dimensional interval, the most important for us is that the space of densities is flat with respect to the $L^2$-Wasserstein metric; it is of non-positive curvature in higher dimensions, which makes the numerical approximation of the Wasserstein distance and also the preservation of geodesic convexity under the discretization significantly more difficult. A promising approach for a truly structure preserving discretization in higher space dimensions is the recent work [3] mentioned above. There, a numerical solver for second-order drift diffusion equations with aggregation in multiple space dimensions is introduced that preserves—in addition to the Lagrangian and the gradient flow aspects—also “some geometry” of the optimal transport. These manifold structural properties enable the authors to rigorously perform a (partial) convergence analysis. It is currently unclear whether that approach can be pushed further to deal with fourth-order equations as well.
1.4 Structure of the Paper

We start with a description of our Lagrangian discretization in Sect. 2; the fully discrete scheme is defined in Sect. 2.5. In Sect. 3, we derive various a priori estimates on the fully discrete solutions. This leads to the main convergence results in Propositions 19 and 20, showing the existence of a limit function $u_*$ for $\Delta \to 0$. In Sect. 4, it is verified that $u_*$ is indeed a weak solution to (1)–(3). The formal conclusion of the proofs for Theorems 1 and 2 is contained in the short Sect. 5. Finally, Sect. 6 provides a consistency analysis and results from numerical simulations of (13).

2 Discretization in Space and Time

2.1 Inverse Distribution Functions

Before defining the discrete quantities, let us recall some basic facts from the continuous context. We denote by

$$
P([a, b]) = \left\{ u : [a, b] \to \mathbb{R}_{\geq 0} : \int_a^b u(x) \, dx = M \right\}
$$

the space of densities of total mass $M$ on $[a, b]$, and we endow $P([a, b])$ with the $L^2$-Wasserstein metric $\mathbb{W}$. We refer to [47] for a comprehensive introduction to the topic. For our purposes here, it suffices to know that convergence with respect to $\mathbb{W}$ is equivalent to weak-$\star$ convergence in $L^1([a, b])$ and that the $L^2$-Wasserstein distance on $P([a, b])$ is isometrically equivalent to the usual $L^2$-distance on the space

$$
\mathcal{X} = \left\{ X : [0, M] \to [a, b] : X \text{ is càdlàg and strictly increasing} \right\}
$$

of inverse distribution functions $X$. The isometry is given as follows.

**Lemma 4** Given $u^0, u^1 \in P([a, b])$, introduce their Lagrangian maps $X^0, X^1 \in \mathcal{X}$ such that

$$
\xi = \int_0^{X^j(\xi)} u^j(x) \, dx \quad \text{for all } \xi \in [0, M].
$$

Then,

$$
\mathbb{W}(u^0, u^1) = \|X^0 - X^1\|_{L^2([0, M])}.
$$

Above, the name Lagrangian map is underlined by the following change of variables formula,

$$
\int_a^b \varphi(x) u(x) \, dx = \int_0^M \varphi(X(\xi)) \, d\xi,
$$

(19)
that holds for every bounded and continuous test function $\varphi \in C^0([a, b])$.

### 2.2 Ansatz Space

Fix a discretization parameter $K \in \mathbb{N}$, which is the number of degrees of freedom plus one. We will need both the integers and the half-integers between $0$ and $K$, that is

$$I_0^K = \{1, 2, \ldots, K - 1\}, \quad I^{1/2}_K = \left\{\frac{1}{2}, \frac{3}{2}, \ldots, K - \frac{1}{2}\right\}.$$ 

For discretization of $[0, M]$, introduce the equidistant mass grid $(\xi_0, \ldots, \xi_K)$ with

$$\xi_k = k\delta \quad \text{for} \quad \delta := \frac{M}{K}.$$ 

For discretization of $[a, b]$, we consider (non-equidistant) grids from

$$\xi_\delta = \{\bar{x} = (x_1, \ldots, x_{K-1}) \mid a < x_1 < \cdots < x_{K-1} < b\} \subseteq (a, b)^{K-1}.$$ 

By definition, $\bar{x} \in \xi_\delta$ is a vector with $K - 1$ components, but we shall frequently use the convention that $x_0 = a$ and $x_K = b$. In the convex set $X$ of inverse distribution functions, we single out the $(K - 1)$-dimensional open and convex subset

$$X_\delta = \{X \in X \cap C^0([a, b]) \mid \text{X is affine on each } [\xi_{k-1}, \xi_k], \text{ with } X(0) = a, \ X(M) = b\}.$$ 

Functions $X \in X_\delta$ are called Lagrangian maps, since they map the (fixed reference) mesh $(\xi_0, \xi_1, \ldots, \xi_K)$ to a (variable) mesh $\bar{x} \in \xi_\delta$. There is a one-to-one correspondence between grid vectors $\bar{x} \in \xi_\delta$ and inverse distribution function $X \in X_\delta$, explicitly given by

$$X = X_\delta[\bar{x}] = \sum_{k \in I_0^K} x_k \theta_k,$$ 

where the $\theta_k : [0, M] \to \mathbb{R}$ are the usual affine hat functions, with $\theta_k(\xi_\ell) = \delta_{k, \ell}$. Further, the density function $u_\delta[\bar{x}] \in P([a, b])$ associated with $X_\delta[\bar{x}]$ is

$$u_\delta[\bar{x}](x) = \sum_{\kappa \in I^{1/2}_K} z_\kappa 1_{(x_{\kappa-\frac{1}{2}}, x_{\kappa+\frac{1}{2}}]}(x),$$ 

where the vector

$$\tilde{z} = z_\delta[\bar{x}] = (z_{1/2}, \ldots, z_{K-1/2})$$ 

of weights $z_\kappa = \frac{\delta}{x_{\kappa+\frac{1}{2}} - x_{\kappa-\frac{1}{2}}}, \ \kappa \in I^{1/2}_K,$

is such that each interval $(x_{\kappa-\frac{1}{2}}, x_{\kappa+\frac{1}{2}}]$ contains the same amount $\delta$ of total mass. The following convention reflects the no-flux boundary conditions:
We finally introduce the associated $(K - 1)$-dimensional submanifold $\mathcal{P}_\delta([a, b]) := \mathbf{u}_\delta[\tau_\delta] \subset \mathcal{P}([a, b])$ as the image of the injective map $\mathbf{u}_\delta : \tau_\delta \to \mathcal{P}_\delta([a, b])$.

### 2.3 A Metric on the Ansatz Space

Below, we define a “Wasserstein-like” metric $W_\delta$ on the ansatz space $\mathcal{P}_\delta([a, b])$. For motivation of that definition, observe that $\mathcal{P}_\delta([a, b])$ is a geodesic submanifold of $\mathcal{P}([a, b])$; hence, the restriction $\tilde{W}_\delta$ of the genuine $L^2$-Wasserstein distance $W$ to $\mathcal{P}_\delta([a, b])$ appears as a natural candidate for $W_\delta$. Thanks to the flatness of $W$ in one space dimension, see Lemma 4, the pullback metric of $W$ on $\tau_\delta$ induced by $\mathbf{u}_\delta$ is a homogeneous quadratic form. More precisely,

$$W(u_\delta[\bar{x}^0], u_\delta[\bar{x}^1])^2 = \sum_{k=1}^{K-1} (\bar{x}^1_k - \bar{x}^0_k) \tilde{W}_{k\ell} (\bar{x}^1_\ell - \bar{x}^0_\ell) \text{ for all } \bar{x}^0, \bar{x}^1 \in \tau_\delta, \quad (24)$$

where the positive matrix $\tilde{W} \in \mathbb{R}^{(K-1) \times (K-1)}$ is tridiagonal. This approach has been followed in our previous work [44].

Here, we take a modified approach and use (24) to define a metric $W_\delta$ on $\mathcal{P}_\delta([a, b])$, but with the simpler matrix $\delta \mathbb{I}_{K-1}$ in place of $\tilde{W}$ above. In other words: up to a factor $\delta^{1/2}$, the pullback metric of $W_\delta$ via $\mathbf{u}_\delta$ is the usual Euclidean distance on $\tau_\delta$.

**Remark 5** Our proof of convergence heavily relies on several explicit estimates of quantities with respect to the metric $W_\delta$.

With the rescaled scalar product $\langle \cdot, \cdot \rangle_\delta$ and norm $\| \cdot \|_\delta$ defined for $\bar{v}, \bar{w} \in \mathbb{R}^{K-1}$ by

$$\langle \bar{v}, \bar{w} \rangle_\delta = \delta \sum_{k=1}^{K-1} v_k w_k, \quad \| \bar{v} \|_\delta = \left( \delta \sum_{k=1}^{K-1} v_k^2 \right)^{1/2},$$

the distance $W_\delta$ is conveniently written as

$$W_\delta(u_\delta[\bar{x}^0], u_\delta[\bar{x}^1]) = \| \bar{x}^1 - \bar{x}^0 \|_\delta.$$ 

In [44, Lemma 3.2], we have shown the following.

**Lemma 6** $W_\delta$ is equivalent to the Wasserstein metric restricted to $\mathcal{P}_\delta([a, b])$, uniformly in $K$:

$$\frac{1}{6} W_\delta(u_0, u_1)^2 \leq W(u_0, u_1)^2 \leq W_\delta(u_0, u_1)^2 \text{ for all } u_0, u_1 \in \mathcal{P}_\delta([a, b]). \quad (25)$$
Note that, as a direct consequence of (25), we obtain that
\[ \left\| X_\delta [\bar{x}^1] - X_\delta [\bar{x}^0] \right\|_{L^2([0,M])} \leq \left\| \bar{x}^1 - \bar{x}^0 \right\|_\delta. \]
\[(26)\]

We shall not elaborate further on the point in which sense the thereby defined metric \( \mathbb{W}_\delta \) is a good approximation of the \( L^2 \)-Wasserstein distance on \( P_\delta([a,b]) \). However, Theorem 2 validates our choice a posteriori. For results concerning the \( \Gamma \)-convergence of discretized transport metrics to the Wasserstein distance, see [27].

2.4 Functions on \( P_\delta([a,b]) \)

When discussing functions on \( P_\delta([a,b]) \) in the following, we always assume that these are given in the form \( f : \tau_\delta \to \mathbb{R} \). We denote the first and second derivatives of \( f \) by \( \partial_\bar{x} f : \tau_\delta \to \mathbb{R}^{K-1} \) and by \( \partial_\bar{x}^2 f : \tau_\delta \to \mathbb{R}^{(K-1) \times (K-1)} \), respectively, with components
\[ [\partial_\bar{x} f(\bar{x})]_k = \partial_{x_k} f(\bar{x}) \quad \text{and} \quad [\partial_\bar{x}^2 f(\bar{x})]_{k,l} = \partial_{x_k} \partial_{x_l} f(\bar{x}). \]

Example 7 Each component \( z_k \) of \( \bar{z} = z_\delta[\bar{x}] \) is a function on \( \tau_\delta \), and
\[ \partial_\bar{x} z_k = - z_k^2 \frac{e_k + \frac{1}{2} - e_k - \frac{1}{2}}{\delta}, \]
where \( e_k \in \mathbb{R}^{K-1} \) is the \( k \)th canonical unit vector, with the convention \( e_0 = e_K = 0 \).

We introduce further the gradient
\[ \nabla_\delta f(\bar{x}) = \delta^{-1} \partial_\bar{x} f(\bar{x}), \]
where the scaling by \( \delta^{-1} \) is chosen such that, for arbitrary vectors \( \bar{v} \in \mathbb{R}^{K-1} \),
\[ \langle \bar{v}, \nabla_\delta f(\bar{x}) \rangle_\delta = \sum_{k=1}^{K-1} v_k \partial_{x_k} f(\bar{x}). \]

The gradient flow of a function \( f \) on \( P_\delta([a,b]) \) with respect to \( \mathbb{W}_\delta \) is then defined as the solution \( \bar{x} : [0; \infty) \to \tau_\delta \) for the system of ordinary differential equations
\[ \dot{\bar{x}} = - \nabla_\delta f(\bar{x}), \quad \text{or, more explicitly,} \quad \dot{x}_k = - \delta^{-1} \partial_{x_k} f(\bar{x}), \quad \text{for each} \ k \in \mathbb{N}. \]

2.4.1 The Discretized Boltzmann Entropy

The Boltzmann entropy \( \mathcal{H} \) as defined in (6) is a nonnegative functional on \( P([a,b]) \), which vanishes precisely on the constant function \( u \equiv M/(b-a) \). In analogy to [44],
we introduce a discretization $H_{\delta} : \mathcal{P}_{\delta} \to \mathbb{R}$ of the Boltzmann entropy $\mathcal{H}$ by restriction to $P_{\delta}([a, b])$:

$$H_{\delta}(\bar{x}) := \mathcal{H}(u_{\delta}[\bar{x}]) = \int_{a}^{b} u_{\delta}[\bar{x}] \ln u_{\delta}[\bar{x}] \, dx - \mathcal{H}_{0} = \delta \sum_{\kappa \in I_{1/2}} \ln z_{\kappa} - \mathcal{H}_{0},$$

where $\mathcal{H}_{0}$ was defined in (6), and $\bar{z} = z_{\delta}[\bar{x}]$. Naturally, $H_{\delta}$ inherits nonnegativity and vanishes only for $\bar{x}$ with $x_{k} = a + (b - a)k/K$. For the derivatives, we obtain—using the rule (27)—

$$\partial_{\bar{x}} H_{\delta}(\bar{x}) = -\delta \sum_{\kappa \in I_{1/2}} \frac{e_{\kappa - \frac{1}{2}} - e_{\kappa + \frac{1}{2}}}{\delta} = \delta \sum_{\kappa \in I_{1/2}} \frac{z_{k+\frac{1}{2}} - z_{k-\frac{1}{2}}}{\delta} e_{k}, \quad (28)$$

$$\partial_{\bar{x}}^{2} H_{\delta}(\bar{x}) = \delta \sum_{\kappa \in I_{1/2}} z_{\kappa}^{2} \left( \frac{e_{\kappa - \frac{1}{2}} - e_{\kappa + \frac{1}{2}}}{\delta} \right) \left( \frac{e_{\kappa - \frac{1}{2}} - e_{\kappa + \frac{1}{2}}}{\delta} \right)^{T} \delta. \quad (29)$$

It is obvious that $\partial_{\bar{x}}^{2} H_{\delta}$ is positive semi-definite, i.e., that $H_{\delta}$ is convex.

2.4.2 The Discretized Fisher Information

The discrete Fisher information $F_{\delta} : \mathcal{P}_{\delta} \to \mathbb{R}$ is not defined by restriction of $\mathcal{F}$ from (7). Instead, we mimic (9) and define accordingly

$$F_{\delta}(\bar{x}) = \frac{1}{2} \| \nabla_{\delta} H_{\delta}(\bar{x}) \|_{\delta}^{2} = \frac{\delta}{2} \sum_{\kappa \in I_{1/2}} \left( \frac{z_{k+\frac{1}{2}} - z_{k-\frac{1}{2}}}{\delta} \right)^{2},$$

using (28). The formal similarity between $F_{\delta}$ and the original Fisher information $\mathcal{F}$ is most easily seen from the representation (11), which can also be written as

$$F(X) = \frac{1}{2} \int_{0}^{M} (\partial_{\xi} Z)^{2} d\xi, \quad \text{with} \quad Z = \frac{1}{\partial_{\xi} X}.$$ 

Thanks to the simple structure of $F_{\delta}$, its gradient flow equation has an explicit and compact representation. Using the rule (27), the representation (29) and the convention (23), we obtain with $\bar{z} = z_{\delta}[\bar{x}]$:

$$\nabla_{\delta} F_{\delta}(\bar{x}) = \delta^{-2} \partial_{\bar{x}}^{2} H_{\delta}(\bar{x}) \partial_{\bar{x}} H_{\delta}(\bar{x})$$

$$= \sum_{\kappa \in I_{1/2}, \kappa \in I_{1/2}^{+}} z_{\kappa}^{2} \left( \frac{e_{\kappa+\frac{1}{2}} - e_{\kappa-\frac{1}{2}}}{\delta} \right) \left( \frac{e_{\kappa+\frac{1}{2}} - e_{\kappa-\frac{1}{2}}}{\delta} \right)^{T} e_{k}.$$
\[
\sum_{\kappa \in \mathbb{N}/2} \frac{z_{\kappa}^{2}}{2} \left( \frac{z_{\kappa+1} - 2z_{\kappa} + z_{\kappa-1}}{\delta^{2}} \right) \left( \frac{e_{\kappa}^{rac{1}{2}} - e_{\kappa-rac{1}{2}}}{\delta} \right).
\]

This should be understood as a discretization of the differential operator \(\partial_{\zeta} (Z^{2} \partial^{2}_{\zeta} Z)\) appearing on the right-hand side of (12).

**Remark 8** Without calculating the second derivative \(\partial^{2}_{\zeta} F_{\delta}\) explicitly, we remark that it is unbounded from below on \(X_{\delta}\); hence, \(F_{\delta}\) is not \(\lambda\)-convex for any \(\lambda \in \mathbb{R}\). This is in agreement with the fact that already the original Fisher information \(F\) is not geodesically \(\lambda\)-convex in the Wasserstein metric, see [13].

### 2.5 Time Stepping

For the definition of the fully discrete scheme for solution of (8), we discretize the spatially discrete gradient flow equation

\[
\dot{\bar{\mathbf{x}}} = -\nabla_{\delta} F_{\delta}(\bar{\mathbf{x}})
\]

also in time, using minimizing movements. To this end, fix a time step with \(\tau > 0\); we combine the spatial and temporal mesh widths in a single discretization parameter

\[
\Delta = (\tau; \delta).
\]

For each \(\bar{y} \in \mathbb{R}_{\delta}\), introduce the Yosida-regularized Fisher information \(F_{\Delta}(:; \bar{y}) : \mathbb{R}_{\delta} \rightarrow \mathbb{R}\) by

\[
F_{\Delta}(\bar{x}; \bar{y}) = \frac{1}{2\tau} \|\bar{x} - \bar{y}\|_{\delta}^{2} + F_{\delta}(\bar{x}).
\]

A fully discrete approximation \((\bar{x}_{\Delta}^{n})_{n=0}^{\infty}\) of (31) is now defined inductively from a given initial datum \(\bar{x}_{\Delta}^{0}\) by choosing each \(\bar{x}_{\Delta}^{n}\) as a global minimizer of \(F_{\Delta}(:; \bar{x}_{\Delta}^{n-1})\). Below, we prove that such a minimizer always exists, see Lemma 10.

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

\[
\frac{\bar{x}_{\Delta}^{n} - \bar{x}_{\Delta}^{n-1}}{\tau} = -\nabla_{\delta} F_{\delta}(\bar{x}_{\Delta}^{n}).
\]

Using the explicit representation (30) of \(\partial_{\zeta} F_{\delta}\), it is immediately seen that (32) is indeed the same as (13). Equivalence of (32) and the minimization problem is guaranteed at least for sufficiently small \(\tau > 0\).

**Proposition 9** For each discretization \(\Delta\) and every initial condition \(\bar{x}_{\Delta}^{0} \in \mathbb{R}_{\delta}\), the sequence of equations (32) can be solved inductively. Moreover, if \(\tau > 0\) is sufficiently
small with respect to \( \delta \) and \( F_\delta(\vec{x}_\Delta^0) \), then each equation (32) possesses a unique solution \( \vec{x}_\Delta^n \in \mathcal{G}_\delta \) with \( F_\delta(\vec{x}_\Delta^n) \leq F_\delta(\vec{x}_\Delta^0) \) and that solution is the unique global minimizer of \( F_\Delta(\cdot; \vec{x}_\Delta^n) \).

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

**Lemma 10** Fix a spatial discretization parameter \( \delta \) and a bound \( C > 0 \). Then, for every \( \vec{y} \in \mathcal{G}_\delta \) with \( F_\delta(\vec{y}) \leq C \), the following are true:

- for each \( \tau > 0 \), the function \( F_\Delta(\cdot; \vec{y}) \) possesses at least one global minimizer \( \vec{x}^* \in \mathcal{G}_\delta \);
- there exists a \( \tau_C > 0 \) independent of \( \vec{y} \) such that for each \( \tau \in (0, \tau_C) \), the global minimizer \( \vec{x}^* \in \mathcal{G}_\delta \) is strict and unique, and it is the only critical point of \( F_\Delta(\cdot; \vec{y}) \) with \( F_\delta(\vec{x}) \leq C \).

**Proof** First, observe that the sublevel \( A_C := F_\delta^{-1}([0, C + 1]) \subset \mathcal{G}_\delta \) is a compact subset of \( \mathbb{R}^{K-1} \). Indeed, \( A_C \) is a relatively closed subset of \( \mathcal{G}_\delta \) by continuity of \( F_\delta \). Moreover, thanks to (101) from the “Appendix,” every \( \vec{x} \in A_C \) satisfies \( x_{k+\frac{1}{2}} - x_{k-\frac{1}{2}} \geq \chi \) for all \( k \in [1/2]^K \) with a positive constant \( \chi \) that depends on \( C \) only. Thus, \( A_C \) does not touch the boundary (in the ambient \( \mathbb{R}^{K-1} \) of \( \mathcal{G}_\delta \)). Consequently, \( A_C \) is closed and bounded in \( \mathbb{R}^{K-1} \).

Let \( \vec{y} \in \mathcal{G}_\delta \) with \( F_\delta(\vec{y}) \leq C \) be given. The restriction of the continuous function \( F_\Delta(\cdot; \vec{y}) \) to the compact and non-empty (since it contains \( \vec{y} \)) set \( A_C \) possesses a minimizer \( \vec{x}^* \in A_C \). We clearly have \( F_\delta(\vec{x}^*) \leq F_\delta(\vec{y}) \leq C \), and so \( \vec{x}^* \) lies in the interior of \( A_C \) and therefore is a global minimizer of \( F_\Delta(\cdot; \vec{y}) \). This proves the first claim.

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

\[
\partial^2_x F_\Delta(\vec{x}; \vec{y}) = \partial^2_x F_\delta(\vec{x}) + \frac{\delta}{\tau} 1_{K-1},
\]

which means that \( \vec{x} \mapsto F_\Delta(\vec{x}; \vec{y}) \) is strictly convex on \( A_C \) if

\[
0 < \tau < \frac{\delta}{(-\lambda_C)}.
\]

Consequently, each such \( F_\Delta(\cdot; \vec{y}) \) has at most one critical point \( \vec{x}^* \) in the interior of \( A_C \), and this \( \vec{x}^* \) is necessarily a strict global minimizer. \( \square \)

### 2.6 Spatial Interpolations

Consider a fully discrete solution (\( \vec{x}_\Delta^n \))\( _{n=0}^\infty \). For notational simplification, we write the entries of the vectors \( \vec{x}_\Delta^n \) and \( \vec{z}_\Delta^n = \vec{z}_\delta(\vec{x}_\Delta^n) \) as \( x^n_k \) and \( z^n_k \), assuming that the discretization parameter \( \Delta \) is fixed.

Recall that \( \vec{u}_\Delta = \vec{u}_\delta(\vec{x}_\Delta^n) \in \mathcal{P}_\delta([a, b]) \) defines a sequence of densities on \([a, b]\) which are piecewise constant with respect to the (non-uniform) grid
Throughout this section, we consider a sequence \( x_0, x_1, \ldots, x_{K-1}, b \). To facilitate the study of convergence of weak derivatives, we introduce also piecewise affine interpolations \( \bar{z}_n^\delta : [0, M] \to \mathbb{R} \) and \( \bar{u}_n^\delta : [a, b] \to \mathbb{R} \).

In addition to \( \xi_k = k\delta \) for \( k \in [0, K] \), introduce the intermediate points \( \xi_k = \kappa \delta \) for \( \kappa \in \mathbb{I}_{K}^{1/2} \). Accordingly, introduce the intermediate values for the vectors \( \bar{x}_n^\delta \) and \( \bar{z}_n^\delta \):

\[
x_k^n = \frac{1}{2} \left( x_{k+\frac{1}{2}} + x_{k-\frac{1}{2}} \right) \quad \text{for} \quad k \in \mathbb{I}_{K}^{1/2},
\]

\[
z_k^n = \frac{1}{2} \left( z_{k+\frac{1}{2}} + z_{k-\frac{1}{2}} \right) \quad \text{for} \quad k \in \mathbb{I}_{K}^{1/2}.
\]

Now define

1. \( \bar{z}_n^\delta : [0, M] \to \mathbb{R} \) as the piecewise affine interpolation of the values \( \left( z_n^{\delta/2}, z_n^{\delta}, \ldots, z_n^{M-\delta/2} \right) \) with respect to the equidistant grid \( \left( \frac{\delta}{2}, \frac{\delta}{2}, \ldots, \frac{\delta}{2} \right) \), and

2. \( \bar{u}_n^\delta : [a, b] \to \mathbb{R} \) as the piecewise affine function with

\[
\bar{u}_n^\delta \circ X_n^\delta = \bar{z}_n^\delta.
\]

Our convention is that \( \bar{z}_n^\delta \left( \xi \right) = z_n^{\delta/2} \) for \( 0 \leq \xi \leq 2/\delta \) and \( \bar{z}_n^\delta \left( \xi \right) = z_n^{M-\delta/2} \) for \( M-\delta/2 \leq \xi \leq M \), and accordingly \( \bar{u}_n^\delta \left( x \right) = z_n^{\delta/2} \) for \( x \in [a, x_n^{\delta/2}] \) and \( \bar{u}_n^\delta \left( x \right) = z_n^{M-\delta/2} \) for \( x \in [x_n^{M-\delta/2}, b] \). The definitions have been made such that

\[
x_k^n = X_n^\delta \left( \xi_k \right), \quad z_k^n = \bar{z}_n^\delta \left( \xi_k \right) = \bar{u}_n^\delta \left( x_k^n \right) \quad \text{for all} \quad k \in [0, K] \cup [M/2, M/2].
\]

Notice that \( \bar{u}_n^\delta \) is piecewise affine with respect to the “double grid” \( \left( x_0^n, x_1^n, x_2^n, \ldots, x_{K-1}^{n}, x_K^n \right) \), but in general not with respect to the subgrid \( \left( x_0^n, x_1^n, \ldots, x_K^n \right) \). By direct calculation, we obtain for each \( k \in \mathbb{I}_{K}^{1/2} \) that

\[
\partial_x \bar{u}_n^\delta \bigg|_{\left( x_{k-\frac{1}{2}}, x_{k+\frac{1}{2}} \right)} = \frac{z_k^n - z_{k-\frac{1}{2}}}{x_{k+\frac{1}{2}} - x_{k-\frac{1}{2}}} = \frac{z_k^{n+\frac{1}{2}} - z_k^{n-\frac{1}{2}}}{x_{k+\frac{1}{2}} - x_{k-\frac{1}{2}}} = \frac{z_k^n - z_k^{n+\frac{1}{2}} + z_k^n - z_k^{n-\frac{1}{2}}}{\delta},
\]

\[
\partial_x \bar{u}_n^\delta \bigg|_{\left( x_{k-\frac{1}{2}}, x_{k+\frac{1}{2}} \right)} = \frac{z_{k+\frac{1}{2}} - z_k^n}{x_{k+\frac{1}{2}} - x_k^n} = \frac{z_k^{n+\frac{1}{2}} - z_k^{n-\frac{1}{2}}}{x_{k+\frac{1}{2}} - x_k^n} = \frac{z_k^n - z_k^{n+\frac{1}{2}} + z_k^n - z_k^{n-\frac{1}{2}}}{\delta}. \quad (34)
\]

Trivially, we also have that \( \partial_x \bar{u}_n^\delta \) vanishes identically on the intervals \( (a, x_1^n) \) and \( (x_{K-1}^{n}, b) \).

3 A Priori Estimates and Compactness

Throughout this section, we consider a sequence \( \Delta = (\tau; \delta) \) of discretization parameters such that \( \delta \to 0 \) and \( \tau \to 0 \) in the limit, formally denoted by \( \Delta \to 0 \). We
assume that a fully discrete solution \( (\tilde{x}^n_\Delta)_{n=0}^\infty \) is given for each \( \Delta \)-mesh, defined by inductive minimization of the respective \( F_\Delta \). The sequences \( \tilde{u}_\Delta, \tilde{u}_\Delta, \tilde{z}_\Delta \) and \( X_\Delta \) of spatial interpolations are defined from the respective \( \tilde{x}_\Delta \) accordingly. For the sequence of initial conditions \( \tilde{x}_0 \), we assume that \( \tilde{u}_0 \rightarrow u^0 \) weakly in \( L^1([a,b]) \), that there is some finite \( H \) with

\[
H(\tilde{x}_0^0) \leq H \quad \text{for all } \Delta,
\]

and that

\[
(\tau + \delta)F_\delta(\tilde{x}_0^0) \rightarrow 0 \quad \text{as } \Delta \rightarrow 0.
\]

Further, for any sequence \((q^n)_{n=0}^\infty\) of objects \( q^n \)—typically functions on \([a,b]\)—we use \( \{q\}_\tau \) to denote the constant in time interpolation with step size \( \tau > 0 \), that is

\[
\{q\}_\tau (t) := \begin{cases} q^n & \text{for } t \in ((n-1)\tau, n\tau], \\ q_0 & \text{for } t = 0. \end{cases}
\]

### 3.1 Energy Inequality

The following basic energy estimates are classical for gradient flows.

**Lemma 11** One has that \( F_\delta(\tilde{x}_\Delta) \leq F_\delta(\tilde{x}_\Delta^{n-1}) \), and further:

\[
F_\delta(\tilde{x}_\Delta^n) \leq \frac{1}{2} \tau \|\tilde{x}_\Delta^n - \tilde{x}_\Delta^{n-1}\|_\delta^2 + F_\delta(\tilde{x}_\Delta^n) = F_\Delta(\tilde{x}_\Delta^n; \tilde{x}_\Delta^{n-1}) \leq F_\Delta(\tilde{x}_\Delta^{n-1}; \tilde{x}_\Delta^{n-1}) = F_\delta(\tilde{x}_\Delta^{n-1}).
\]

**Proof** The monotonicity (37) follows (by induction on \( n \)) from the definition of \( \tilde{x}_\Delta^n \) as minimizer of \( F_\Delta(\cdot; \tilde{x}_\Delta^{n-1}) \):

\[
\begin{align*}
F_\delta(\tilde{x}_\Delta^n) &\leq \frac{1}{2} \tau \|\tilde{x}_\Delta^n - \tilde{x}_\Delta^{n-1}\|_\delta^2 + F_\delta(\tilde{x}_\Delta^n) = F_\Delta(\tilde{x}_\Delta^n; \tilde{x}_\Delta^{n-1}) \\
&\leq F_\Delta(\tilde{x}_\Delta^{n-1}; \tilde{x}_\Delta^{n-1}) = F_\delta(\tilde{x}_\Delta^{n-1}).
\end{align*}
\]

Moreover, summation of these inequalities from \( n = n + 1 \) to \( n = \tilde{n} \) yields

\[
\frac{\tau}{2} \sum_{n=n+1}^{\tilde{n}} \left[ \frac{\|\tilde{x}_\Delta^n - \tilde{x}_\Delta^{n-1}\|_\delta}{\tau} \right]^2 \leq F_\delta(\tilde{x}_\Delta^n) - F_\delta(\tilde{x}_\Delta^\tilde{n}) \leq F_\delta(\tilde{x}_\Delta^0).
\]
For \( n = 0 \) and \( \bar{n} \to \infty \), we obtain the first part of (39). The second part follows by (32). If instead we combine the estimate with Jensen’s inequality, we obtain

\[
\| \bar{x}^n_\Delta - \bar{x}^{n-1}_\Delta \|_\delta \leq \tau \sum_{n=\bar{n}+1}^{\bar{n}} \left( \frac{\| \bar{x}^n_\Delta - \bar{x}^{n-1}_\Delta \|_\delta}{\tau} \right)^{1/2} \left( \tau (\bar{n} - n) \right)^{1/2},
\]

which leads to (38).

\[\square\]

### 3.2 Entropy Dissipation

The key to our convergence analysis is a refined a priori estimate, which follows from the dissipation of the entropy \( H_\delta \) along the fully discrete solution.

**Lemma 12** One has that \( H_\delta \) is monotone, i.e., \( H_\delta(\bar{x}^n_\Delta) \leq H_\delta(\bar{x}^{n-1}_\Delta) \), and further:

\[
\tau \sum_{n=1}^{\infty} \delta \sum_{\kappa \in \mathbb{Z}_K} (z^n_\kappa)^2 \left( \frac{\zeta^n_{k+1} - 2\zeta^n_k + \zeta^n_{k-1}}{\delta^2} \right)^2 \leq H_\delta(\bar{x}^0_\Delta).
\]

**Proof** By convexity of \( H_\delta \) and the discrete evolution (32), we have

\[
H_\delta(\bar{x}^{n-1}_\Delta) - H_\delta(\bar{x}^n_\Delta) \geq \left\langle \nabla_\delta H_\delta(\bar{x}^n_\Delta), \bar{x}^{n-1}_\Delta - \bar{x}^n_\Delta \right\rangle = \tau \left\langle \nabla_\delta H_\delta(\bar{x}^n_\Delta), \nabla_\delta F_\delta(\bar{x}^n_\Delta) \right\rangle_\delta
\]

for each \( n = 1, 2, \ldots \) Evaluate the (telescopic) sum with respect to \( n \) and use that \( H_\delta \geq 0 \) to obtain

\[
\tau \sum_{n=1}^{\infty} \left\langle \nabla_\delta H_\delta(\bar{x}^n_\Delta), \nabla_\delta F_\delta(\bar{x}^n_\Delta) \right\rangle_\delta \leq H_\delta(\bar{x}^0_\Delta).
\]

It remains to make the scalar product explicit, using (28) and (30):

\[
\left\langle \nabla_\delta H_\delta(\bar{x}^n_\Delta), \nabla_\delta F_\delta(\bar{x}^n_\Delta) \right\rangle_\delta = \delta \sum_{\kappa \in \mathbb{Z}_K, \kappa \in \mathbb{Z}_K} (z^n_\kappa)^2 \left( \frac{\zeta^n_{k+1} - 2\zeta^n_k + \zeta^n_{k-1}}{\delta^2} \right)^2 \left( \frac{e^n_{k+1} - e^n_{k-1}}{\delta} \right)^T e_k
\]

using that \( z^n_{-\frac{1}{2}} = z^n_\frac{1}{2} \) and \( z^n_{K+\frac{1}{2}} = z^n_{K-\frac{1}{2}} \), according to our convention (23). \[\square\]
We draw several conclusions from (40). The first is an a priori estimate on the \(\xi\)-derivative of the affine functions \(\hat{z}_n/\Delta_1\).

**Lemma 13** One has that

\[
\tau \sum_{n=1}^\infty \| \partial_\xi \hat{z}_n \|_{L^4([0,M])}^4 = \tau \sum_{n=1}^\infty \delta \sum_{k \in I_k^0} \left( \frac{\hat{z}_n^{k+1/2} - \hat{z}_n^{k-1/2}}{\delta} \right)^4 \leq 9\overline{H}.
\] (41)

**Remark 14** Morally, a bound on \(\partial_\xi \hat{z}_n\) in \(L^4([0,M])\) corresponds to a bound on \(\partial_x \sqrt{\hat{u}_n}\) in \(L^4([a,b])\).

**Proof of Lemma 13** Fix \(n \in \mathbb{N}\). Invoking our convention (23), one obtains

\[
\| \partial_\xi \hat{z}_n \|_{L^4([a,b])}^4 = \sum_{k \in I_k^0} \left( \frac{\hat{z}_n^{k+1/2} - \hat{z}_n^{k-1/2}}{\delta} \right)^4 \leq 3\delta^2 \sum_{k \in I_k^1} \left| \left( \frac{\hat{z}_n^{k+1/2} - \hat{z}_n^{k-1/2}}{\delta} \right)^2 \right| \left| \left( \frac{\hat{z}_n^{k+1/2} - \hat{z}_n^{k-1/2}}{\delta} \right)^2 \right|.
\]

Using the elementary identity \((p^3 - q^3) = (p - q)(p^2 + q^2 + pq)\) and Young’s inequality, one obtains further

\[
(A) = -\delta \sum_{k \in I_k^1} \hat{z}_n^k \left( \frac{\hat{z}_n^{k+1/2} - \hat{z}_n^{k-1/2}}{\delta} \right)^2 \times \left( \left( \frac{\hat{z}_n^{k+1/2} - \hat{z}_n^{k-1/2}}{\delta} \right)^2 + \left( \frac{\hat{z}_n^{k+1/2} - \hat{z}_n^{k-1/2}}{\delta} \right)^2 \right) \leq \frac{3\delta^2}{2} \sum_{k \in I_k^1} \left( \frac{\hat{z}_n^{k+1/2} - \hat{z}_n^{k-1/2}}{\delta} \right)^2.
\]

Note that the last sum above is again the \(L^4\)-norm of \(\partial_\xi \hat{z}_n\). Taking the square on both sides, dividing by the \(L^4\)-norm, summing over \(n = 1, 2, \ldots\), and finally applying the entropy dissipation estimate (40), one arrives at (41). \(\square\)

The a priori estimate (41) is the basis for almost all of the further estimates. For instance, the following control on the oscillation of the \(z\)-values at neighboring grid points is a consequence of (41).
Lemma 15 One has
\[
\tau \sum_{n=1}^{\infty} \delta \sum_{k \in \mathbb{N}_+^1} \left[ \left( \frac{z_n^{n+\frac{1}{2}}}{z_n^{n-\frac{1}{2}}} - 1 \right)^4 + \left( \frac{z_n^{n-\frac{1}{2}}}{z_n^{n+\frac{1}{2}}} - 1 \right)^4 \right] \leq 18(b - a)^4 H_\delta(x_\Delta^0). \tag{42}
\]

Moreover, given \( T > 0 \), then for each \( N \in \mathbb{N} \) with \( N \tau \leq T \), one has
\[
\tau \sum_{n=1}^{N} \delta \sum_{k \in \mathbb{N}_+^1} \left[ \left( \frac{z_n^{n+\frac{1}{2}}}{z_n^{n-\frac{1}{2}}} - 1 \right)^2 + \left( \frac{z_n^{n-\frac{1}{2}}}{z_n^{n+\frac{1}{2}}} - 1 \right)^2 \right] \leq 6(b - a)^2 T^{1/2} H_\delta(x_\Delta^0)^{1/2} \delta^{1/2}. \tag{43}
\]

Proof Recall that \( z_\kappa \geq \delta/(b - a) \) for all \( \kappa \), see (100) in the “Appendix.” Consider the first term in the inner summation in (42):
\[
\delta \sum_{k \in \mathbb{N}_+^1} \left( \frac{z_n^{n+\frac{1}{2}}}{z_n^{n-\frac{1}{2}}} - 1 \right)^4 = \delta \sum_{k \in \mathbb{N}_+^1} \left( \frac{\delta}{z_n^{n-\frac{1}{2}}} \right)^4 \left( \frac{z_n^{n-\frac{1}{2}} - z_n^{n+\frac{1}{2}}}{\delta} \right)^4 \leq (b-a)^4 \|z_n^\infty\|^4_{L^4([a,b])}.
\]
The same estimate holds for the second term. The claim (42) is now directly deduced from (41) above. The proof of the second claim (43) is similar, using the Cauchy–Schwarz inequality instead of the modulus estimate:
\[
\delta \sum_{k \in \mathbb{N}_+^1} \left( \frac{z_n^{n+\frac{1}{2}}}{z_n^{n-\frac{1}{2}}} - 1 \right)^2 = \delta \sum_{k \in \mathbb{N}_+^1} \left( \frac{\delta}{z_n^{n-\frac{1}{2}}} \right)^2 \left( \frac{z_n^{n+\frac{1}{2}} - z_n^{n-\frac{1}{2}}}{\delta} \right)^2 \leq \left( \delta \sum_{k \in \mathbb{N}_+^1} \left( \frac{\delta}{z_n^{n-\frac{1}{2}}} \right)^4 \right)^{1/2} \|z_n^\infty\|^2_{L^4([a,b])}.
\]
Use estimate (99) in the “Appendix,” sum over \( n = 1, \ldots, N \), and apply the Cauchy–Schwarz inequality to this second summation. This yields
\[
\tau \sum_{n=1}^{N} \delta \sum_{k \in \mathbb{N}_+^1} \left( \frac{z_n^{n+\frac{1}{2}}}{z_n^{n-\frac{1}{2}}} - 1 \right)^2 \leq \delta^{1/2}(b - a)^2 \left( \tau \sum_{n=1}^{N} 1 \right)^{1/2} \left( \tau \sum_{n=1}^{\infty} \|z_n^\infty\|^4_{L^4([a,b])} \right)^{1/2}.
\]
Invoking again (41), and recalling that \( N \tau \leq T \), we arrive at (43). \( \square \)

We are now going to prove the main consequence from the entropy dissipation (37), namely a control on the total variation of \( \sqrt{u_\Delta} \). This estimate is the key ingredient for obtaining strong compactness in Proposition 20. Recall that several equivalent definitions of the total variation of \( f \in L^1([a, b]) \) exist. Most generally,
TV \[ f \] = \sup \left\{ \int_{a}^{b} f(x) \partial_x \phi(x) \, dx : \phi \in C_{0}^{0,1}([a, b]), \ \max_{x \in [a, b]} |\phi(x)| \leq 1 \right\}, \quad (44)

where \( C_{0}^{0,1}([a, b]) \) are the Lipschitz continuous functions \( \phi : [a, b] \to \mathbb{R} \) with \( \phi(a) = \phi(b) = 0 \).

**Lemma 16** One has

\[
\frac{1}{\tau} \sum_{n=1}^{\infty} \text{TV} \left[ \partial_x \sqrt{\hat{u}_n^\Delta} \right]^2 \leq 10(b - a) \mathcal{H}.
\] (45)

**Proof** Fix \( \Delta \) and \( n \). For brevity, let \( f := \partial_x \sqrt{\hat{u}_n^\Delta} \), which is well-defined except for (potential) jump discontinuities at the points \( x^\Delta_{1/2}, x^\Delta_1, \ldots, x^\Delta_{K-1/2} \). On the intervals in between, \( f \) is bounded (since \( \hat{u}_n^\Delta \) is bounded and strictly positive); hence, \( f \in L^1([a, b]) \). Since \( \hat{u}_n^\Delta \) is piecewise linear, and since \( s \mapsto \sqrt{s} \) is a concave function, \( \partial_x f = \partial_x \sqrt{\hat{u}_n^\Delta} \) is negative in between jumps of \( f \). Further, we have that \( f(x) = 0 \) for all \( x \in (a, x^\Delta_{1/2}) \) and all \( x \in (x^\Delta_{K-1/2}, b) \) due to our implementation of the Neumann boundary conditions for \( \hat{u}_n^\Delta \).

For a given \( \phi \in C_{0}^{0,1}([a, b]) \) with \( \|\phi\| \leq 1 \), an integration by parts in the integral from (44) yields:

\[
\int_{a}^{b} f(x) \partial_x \phi(x) \, dx = \sum_{\ell \in I_{K}^{+} \cup 1/2} \int_{x^\Delta_{\ell-1/2}}^{x^\Delta_{\ell}} f(x) \partial_x \phi(x) \, dx
\]

\[
= \sum_{\ell \in I_{K}^{+} \cup 1/2} \left( \left| f(x) \phi(x) \right|_{x^\Delta_{\ell-1/2}}^{x^\Delta_{\ell}} - \int_{x^\Delta_{\ell-1/2}}^{x^\Delta_{\ell}} \partial_x f(x) \phi(x) \, dx \right)
\]

\[
\leq \sum_{\ell \in I_{K}^{+} \cup 1/2} \left( \left\| f \right\|_{x^\Delta_{\ell}} \left| \phi(x) \right|_{x^\Delta_{\ell-1/2}}^{x^\Delta_{\ell}} + \int_{x^\Delta_{\ell-1/2}}^{x^\Delta_{\ell}} (-\partial_x f(x)) \, dx \right)
\]

\[
\leq 2 \sum_{\ell \in I_{K}^{+} \cup 1/2} \left\| f \right\|_{x^\Delta_{\ell}},
\]

with the usual notation

\[
\left\| f \right\|_{\hat{x}} = f(\hat{x} + 0) - f(\hat{x} - 0) = \lim_{x \downarrow \hat{x}} f(x) - \lim_{x \uparrow \hat{x}} f(x)
\]

for the height of the jump in \( f(x) \)'s value at \( x = \hat{x} \). For the penultimate estimate above, we have used \( \phi \)'s continuity and boundedness and the fact that \( \partial_x f \leq 0 \) on the integration intervals. Switching back from \( f \) to \( \hat{u}_n^\Delta \), we conclude that
\[
\text{TV} \left[ \partial_x \sqrt{\hat{u}^n} \right] \leq 2 \sum_{k \in \mathbb{V}^+_K} \left| \partial_x \sqrt{\hat{u}^n} \right|_{x_k} + 2 \sum_{\kappa \in \mathbb{V}^{1/2}_K} \left| \partial_x \sqrt{\hat{u}^n} \right|_{x_{\kappa}}. \tag{46}
\]

Now, in view of (34), we have that

\[
\left| \partial_x \sqrt{\hat{u}^n} \right|_{x_k} = \frac{1}{2} \sqrt{\frac{z^n_k}{z^n_{k+1} - z^n_{k-1}}} \delta \quad \text{for } k \in \mathbb{V}^+_K,
\]

\[
\left| \partial_x \sqrt{\hat{u}^n} \right|_{x_{\kappa}} = \frac{1}{2} \sqrt{\frac{z^n_{\kappa+1} - 2z^n_{\kappa} + z^n_{\kappa-1}}{z^n_{\kappa}}} \delta \quad \text{for } \kappa \in \mathbb{V}^{1/2}_K.
\]

Accordingly, using that \(1/z^n_k \leq (1/z^n_{k+1/2} + 1/z^n_{k-1/2})/2\) by the arithmetic-harmonic mean inequality,

\[
\sum_{k \in \mathbb{V}^+_K} \left| \partial_x \sqrt{\hat{u}^n} \right|_{x_k} = \frac{\delta}{2} \sum_{k \in \mathbb{V}^+_K} \frac{\left( \frac{z^n_k - z^n_{k+1}}{z^n_{k+1/2}} \right)^2}{\delta^2} \cdot \frac{1}{\sqrt{z^n_k}} \leq \frac{1}{2} \left( \delta \sum_{k \in \mathbb{V}^+_K} \left( \frac{z^n_k - z^n_{k+1}}{z^n_{k+1/2}} \right)^2 \right)^{1/2} \left( \sum_{k \in \mathbb{V}^+_K} \frac{1}{\sqrt{z^n_k}} \right)^{1/2} = \frac{1}{2} \left\| \partial_x \hat{z}^n \right\|_{L^4((a,b))}^2 (b - a)^{1/2}, \tag{47}
\]

and also

\[
\sum_{\kappa \in \mathbb{V}^{1/2}_K} \left| \partial_x \sqrt{\hat{u}^n} \right|_{x_{\kappa}} = \frac{\delta}{2} \sum_{\kappa \in \mathbb{V}^{1/2}_K} \left| \frac{z^n_{\kappa+1} - 2z^n_{\kappa} + z^n_{\kappa-1}}{z^n_{\kappa}} \right| \cdot \frac{1}{\sqrt{z^n_{\kappa}}} \leq \frac{1}{2} \left( \delta \sum_{\kappa \in \mathbb{V}^{1/2}_K} \left( \frac{z^n_{\kappa+1} - 2z^n_{\kappa} + z^n_{\kappa-1}}{z^n_{\kappa}} \right)^2 \right)^{1/2} (b - a)^{1/2}. \tag{48}
\]

Combine (47) with the \(L^4\) bound from (41), and (48) with the entropy dissipation inequality (40). Inserting this into (46) to obtain the claim (45). \(\square\)

### 3.3 Convergence of Time Interpolants

Recall that we require the a priori bound (35) on the initial entropy, but only (36) on the initial Fisher information. This estimate improves over time.
Lemma 17  One has, for every $N \geq 1$,
\[
F_{\delta}(\vec{x}_N^\Delta) \leq \frac{3}{2}(M\overline{T})^{1/2}(N\tau)^{-1/2}.
\]  
(49)

Consequently, $\{F_{\delta}(\vec{x}_N^\Delta)\}_\tau(t)$ is bounded for each $t > 0$, uniformly in $\Delta$.

Proof  Since $F_{\delta}(\vec{x}_n^\Delta)$ is monotonically decreasing in $n$ (for fixed $\Delta$), it follows that
\[
F_{\delta}(\vec{x}_N^\Delta) \leq \frac{1}{N} \sum_{n=1}^{N} F_{\delta}(\vec{x}_n^\Delta) = \frac{1}{2N} \left( \sum_{n=1}^{N} \delta \sum_{k \in l_k^+} \left( \frac{z_{k+1/2}^n - z_{k-1/2}^n}{\delta} \right)^2 \right) \leq \frac{1}{2N} \left( \sum_{n=1}^{\infty} \delta \sum_{k \in l_k^+} \left( \frac{z_{k+1/2}^n - z_{k-1/2}^n}{\delta} \right)^4 \right)^{1/2} \leq \frac{1}{2N} \left( N\tau M \right)^{1/2} (9\overline{T})^{1/2} = \frac{3}{2} (M\overline{T})^{1/2}(N\tau)^{-1/2},
\]
as desired.  

In the following, we use the notation $[t, \tilde{t}] \subset \mathbb{R}_{>0}$ to denote time intervals with $0 < t < \tilde{t} < \infty$.

Lemma 18  One has that, for each $[t, \tilde{t}] \subset \mathbb{R}_{>0}$,
\[
S_{[t, \tilde{t}]} := \sup_{\Delta} \sup_{t \in [t, \tilde{t}]} \| [\hat{u}_\Delta]_\tau(t) \|_{H^1([a, b])} < \infty,
\]  
(50)

and that, as $\Delta \to 0$,
\[
\sup_{t \in \mathbb{R}_{>0}} \| [\hat{u}_\Delta]_\tau(t) - [\bar{u}_\Delta]_\tau(t) \|_{L^\infty([a, b])} \to 0.
\]  
(51)

Proof  For each $n \in \mathbb{N}$,
\[
\| \partial_x \hat{u}_n \|_{L^2([a, b])}^2 = \sum_{k \in l_k^+} \left[ \left( x_{k+1/2}^n - x_k^n \right) \left( \frac{z_{k+1/2}^n - z_k^n}{x_{k+1/2}^n - x_k^n} \right) + \left( x_k^n - x_{k-1/2}^n \right) \left( \frac{z_k^n - z_{k-1/2}^n}{x_k^n - x_{k-1/2}^n} \right) \right]^2 \leq \delta \sum_{k \in l_k^+} \left( \frac{z_{k+1/2}^n - z_k^n}{\delta} \right)^2 \leq F_{\delta}(\vec{x}_N^\Delta) \max_{k \in l_k^+} z_k^n.
\]
Now combine this with the estimates (49) from above and (101) from the “Appendix” to obtain (50). Estimate (51) follows directly from the elementary observation that

$$\sup_{x \in [a, b]} |\tilde{u}_\Delta^n(x) - \bar{u}_\Delta^n(x)|^2 \leq \max_{k \in \mathbb{Z}} \left| z^n_{k+\frac{1}{2}} - z^n_{k-\frac{1}{2}} \right|^2 \leq \delta F_\delta(\bar{\tilde{u}}_\Delta^n) \leq \delta F_\delta(\bar{\bar{u}}_\Delta^n),$$

and an application of (36).

**Proposition 19** There exists a function $u_* : \mathbb{R}_{\geq 0} \times [a, b] \to \mathbb{R}_{\geq 0}$ with

$$u_* \in C^{1/2}_{\text{loc}}(\mathbb{R}_{>0}; \mathcal{P}([a, b])) \cap L^\infty_{\text{loc}}(\mathbb{R}_{>0}; H^1([a, b])), \quad (52)$$

and there exists a subsequence of $\Delta$ (still denoted by $\Delta$), such that, for every $[t, \bar{t}] \in \mathbb{R}_{>0}$, the following are true:

$$\begin{align*}
\{\bar{u}_\Delta\}_\tau(t) &\to u_*(t) \quad \text{with respect to } \mathbb{W}, \text{ uniformly in } t \in [t, \bar{t}], \quad (53) \\
\{\bar{u}_\Delta\}_\tau, \{\bar{\tilde{u}}_\Delta\}_\tau &\to u_* \quad \text{uniformly on } [t, \bar{t}] \times [a, b], \quad (54) \\
\{X_\Delta\}_\tau(t) &\to X^*(t) \quad \text{in } L^2([0, M]), \text{ uniformly with respect to } t \in [t, \bar{t}], \quad (55)
\end{align*}$$

where $X^* \in C^{1/2}_{\text{loc}}(\mathbb{R}_{>0}; L^2([0, M]))$ is the Lagrangian map of $u_*$. 

**Proof** Fix $t > 0$. From the discrete energy inequality (38), the bound on the Fisher information in Lemma 17, and the equivalence (25) of $\mathbb{W}_\delta$ with the usual $L^2$-Wasserstein metric $\mathbb{W}$, it follows by elementary considerations that

$$\mathbb{W}\left(\{\tilde{u}_\Delta\}_\tau(t), \{\bar{u}_\Delta\}_\tau(s)\right)^2 \leq C(t)(|t - s| + \tau), \quad (56)$$

for all $t, s \geq t$. Moreover, since $[a, b]$ is a compact interval, also $\mathcal{P}([a, b])$ is compact. Hence, the generalized version of the Arzela–Ascoli theorem from [1, Proposition 3.3.1] is applicable and yields the convergence of a subsequence of $(\{\bar{u}_\Delta\}_\tau)$ to a limit $u_\tau$ in $\mathcal{P}([a, b])$, locally uniformly with respect to $t \in [t, \infty)$. The Hölder-type estimate (56) implies $u_\tau \in C^{1/2}([t, \infty); \mathcal{P}([a, b]))$. The claim (55) is a consequence of the equivalence between the Wasserstein metric on $\mathcal{P}([a, b])$ and the $L^2$-metric on $X$, see Lemma 4. Clearly, the previous argument applies to every choice of $t > 0$. Using a diagonal argument, one constructs a limit $u_*$ defined on all $\mathbb{R}_{>0}$, such that $u_\tau$ is the restriction of $u_*$ to $[t, \infty)$. This verifies (53).

For the rest of the proof, let some $[t, \bar{t}] \in \mathbb{R}_{>0}$ be fixed. By construction, we have that $[\bar{u}_\Delta]_\tau(t)$ converges to $u_*(t)$ weakly in $L^1([a, b])$. Since the $H^1([a, b])$-norm is lower continuous with respect to weak convergence, it follows from (50) that

$$\sup_{t \in [t, \bar{t}]} \|u_*(t)\|_{H^1([a, b])} \leq S_\tau,$$

which proves $u_* \in L^\infty([t, \bar{t}]; H^1([a, b]))$, and so eventually yields (52). In particular, each $u_*(t)$ with $t \in [t, \bar{t}]$ has a continuous representative.
For proving (54), it suffices to show that $\{\hat{u}_\Delta\}_\tau \to u_*$ uniformly on $[t, \bar{t}] \times [a, b]$: indeed, (51) implies that if $\{\hat{u}_\Delta\}_\tau$ converges uniformly to some limit, so does $\{\hat{u}_\Delta\}_\tau$. As an intermediate step toward proving uniform convergence of $\{\hat{u}_\Delta\}_\tau$, we show that

$$\{\hat{u}_\Delta\}_\tau (t) \longrightarrow u_*(t) \text{ in } L^2([a, b]), \text{ uniformly in } t \in [t, \bar{t}].$$

(58)

For $t \in [t, \bar{t}]$, we expand the $L^2$-norm as follows:

$$\left\| \{\hat{u}_\Delta\}_\tau (t) - u_*(t) \right\|_{L^2([a, b])}^2 = \int_a^b \left[ \left( \{\hat{u}_\Delta\}_\tau - u_* \right) \{\hat{u}_\Delta\}_\tau \right] (t; x) \, dx$$

$$+ \int_a^b \left[ \left( \{\hat{u}_\Delta\}_\tau - u_* \right) \left( \{\hat{u}_\Delta\}_\tau - \{\hat{u}_\Delta\}_\tau \right) \right] (t; x) \, dx$$

$$- \int_a^b \left[ \left( \{\hat{u}_\Delta\}_\tau - u_* \right) u_* \right] (t; x) \, dx.$$

On the one hand, observe that

$$\sup_{t \in [t, \bar{t}]} \int_a^b \left[ \left( \{\hat{u}_\Delta\}_\tau - u_* \right) \left( \{\hat{u}_\Delta\}_\tau - \{\hat{u}_\Delta\}_\tau \right) \right] (t; x) \, dx$$

$$\leq \sup_{t \in [t, \bar{t}]} \left( \| \{\hat{u}_\Delta\}_\tau (t) \|_{L^1([a, b])} + \| u_*(t) \|_{L^1([a, b])} \right) \| \{\hat{u}_\Delta\}_\tau (t) - \{\hat{u}_\Delta\}_\tau (t) \|_{L^\infty([a, b])}$$

$$\leq \sup_{t \in [t, \bar{t}]} \left( \left(2M + (b - a) \| \{\hat{u}_\Delta\}_\tau (t) - \{\hat{u}_\Delta\}_\tau (t) \|_{L^\infty([a, b])} \right) \left( \{\hat{u}_\Delta\}_\tau (t)$$

$$- \{\hat{u}_\Delta\}_\tau (t) \|_{L^\infty([a, b])} \right),$$

which converges to zero as $\Delta \to 0$, using both conclusions from Lemma 18. On the other hand, we can use the change of variables formula (19)—which is applicable since each $u_*(t)$ is continuous by (57)—to write

$$\int_a^b \left[ \left( \{\hat{u}_\Delta\}_\tau - u_* \right) \{\hat{u}_\Delta\}_\tau \right] (t; x) \, dx - \int_a^b \left[ \left( \{\hat{u}_\Delta\}_\tau - u_* \right) u_* \right] (t; x) \, dx$$

$$= \int_0^M \left[ \{\hat{u}_\Delta\}_\tau - u_* \right] (t; X_\Delta (t; \xi)) \, d\xi - \int_0^M \left[ \{\hat{u}_\Delta\}_\tau - u_* \right] (t; X_* (t; \xi)) \, d\xi.$$

We regroup terms under the integrals and use the triangle inequality. For the first term, we obtain

$$\sup_{t \in [t, \bar{t}]} \left| \int_0^M \left[ \{\hat{u}_\Delta\}_\tau (t; X_\Delta (t; \xi)) - \{\hat{u}_\Delta\}_\tau (t; X_* (t; \xi)) \right] \, d\xi \right|$$

$$\leq \sup_{t \in [t, \bar{t}]} \int_0^M \left| \partial_x \{\hat{u}_\Delta\}_\tau \right| (t; y) \, dy \, d\xi.$$
A similar reasoning applies to the integral involving $u_*$ in place of $\{\hat{u}_\Delta\}_\tau$. Together, this proves (58). Now, the Gagliardo–Nirenberg inequality (105) in the “Appendix” provides the estimate

$$\| \{\hat{u}_\Delta\}_\tau (t) \|_{H^1([a,b])} \leq C \| \{\hat{u}_\Delta\}_\tau (t) \|^{2/3}_{H^1([a,b])} \| \{\hat{u}_\Delta\}_\tau (t) - u^*(t) \|_{L^2([a,b])}^{1/3}.$$  

Combining the convergence in $L^2([a,b])$ by (58) with the boundedness in $H^1([a,b])$ from (50), it readily follows that $\hat{u}_\Delta(t) \to u_*(t)$ in $C^{1/6}([a,b])$, uniformly in $t \in [T, \bar{T}]$. This clearly implies that $\{\hat{u}_\Delta\}_\tau \to u_*$ uniformly on $[T, \bar{T}] \times [a,b]$.

**Proposition 20** Under the hypotheses and with the notations of Proposition 19, one has that $\sqrt{u_*} \in L^2_{\text{loc}}(\mathbb{R}_{\geq 0}; H^1([a,b]))$, and

$$\left\{ \sqrt{u_\Delta} \right\}_\tau \to \sqrt{u_*} \text{ strongly in } L^2_{\text{loc}}(\mathbb{R}_{\geq 0}; H^1([a,b]))$$

(59) as $\Delta \to 0$.

Notice that $\partial_x \sqrt{u_*} \in L^2([0, \bar{T}] \times [a,b])$ for each $\bar{T} > 0$, but strong convergence takes place only on each $[T, \bar{T}] \times [a,b]$.

**Proof** Throughout the proof, let $[L, \bar{T}] \Subset \mathbb{R}_{>0}$ be fixed. We begin with an auxiliary interpolation estimate: let $g : [a,b] \to \mathbb{R}$ be Lipschitz continuous and define the linear correction $\ell(x) = px + q$ such that $(g - \ell)(a) = (g - \ell)(b) = 0$. Then, $g - \ell$ is differentiable at a.e. $x \in [a,b]$, with derivative $\partial_x (g - \ell) \in L^\infty([a,b])$. Substituting $f := \partial_x (g - \ell)$ and $\phi := (g - \ell)/\|g - \ell\|_{L^\infty([a,b])}$ in the definition (44) of the total variation, we find that

$$\| \partial_x (g - \ell) \|_{L^2([a,b])}^2 = \int_a^b \left[ \partial_x (g - \ell) (x) \right]^2 dx \leq \|g - \ell\|_{L^\infty([a,b])} TV \left[ \partial_x (g - \ell) \right] \leq 2 \|g\|_{L^\infty([a,b])} TV \left[ \partial_x g \right].$$

Above, we have used that the shift of $\partial_x g$ by the constant function $\partial_x g \equiv p$ does not change its total variation. Now, since $\partial_x (g - \ell)$ is of zero average, and since $|p| \leq 2 \|g\|_{L^\infty([a,b])}/(b-a)$, we further have

$$\| \partial_x g \|_{L^2([a,b])}^2 \leq \| \partial_x (g - \ell) \|_{L^2([a,b])}^2 + (b-a)p^2 \leq 2 \|g\|_{L^\infty([a,b])} TV \left[ \partial_x g \right] + \frac{4}{b-a} \|g\|_{L^2([a,b])}^2.$$  

(60)
We shall use this inequality to estimate the differences \( \partial_x \left\{ \sqrt{u_\Delta} \right\}_\tau \) \(- \partial_x \sqrt{u_*} \) in \( L^2([t, \bar{t}] \times [a, b]) \). To this end, first observe that by (45), and thanks to the lower semi-continuity of the total variation, we have that

\[
\int_0^\infty \text{TV} \left[ \partial_x \sqrt{u_*}(t; \cdot) \right]^2 \, dt \leq \limsup_{\Delta \to 0} \int_0^\infty \text{TV} \left[ \partial_x \left\{ \sqrt{u_\Delta} \right\}_\tau (t; \cdot) \right]^2 \, dt \leq 10(b-a)\mathcal{H}.
\]

(61)

In particular, \( \partial_x \sqrt{u_*}(t; \cdot) \in L^\infty([a, b]) \), and thus \( x \mapsto \sqrt{u_*}(t; x) \) is Lipschitz continuous, for almost every \( t \in [t, \bar{t}] \). Fix such a \( t \). Since also \( x \mapsto \left\{ \sqrt{u_\Delta} \right\}_\tau (t; x) \) is Lipschitz by construction, we can use (60) with

\[
g(x) := \left( \left\{ \sqrt{u_\Delta} \right\}_\tau - \sqrt{u_*} \right)(t; x).
\]

After an integration with respect to \( t \in [t, \bar{t}] \), using the triangle inequality \( \text{TV} [f_1 - f_2] \leq \text{TV} [f_1] + \text{TV} [f_2] \), and an application of Hölder’s inequality, we arrive at

\[
\int_L^\bar{t} \left\| \partial_x \left( \left\{ \sqrt{u_\Delta} \right\}_\tau - \sqrt{u_*} \right)(t; \cdot) \right\|^2_{L^2} \, dt \\
\leq 2 \sup_{t \in [t, \bar{t}]} \left\| \left( \left\{ \sqrt{u_\Delta} \right\}_\tau - \sqrt{u_*} \right)(t; \cdot) \right\|^2_{L^\infty([a, b])} \\
\times \left( 2(\bar{t} - t) \int_L^\bar{t} \left( \text{TV} \left[ \partial_x \left\{ \sqrt{u_\Delta} \right\}_\tau (t; \cdot) \right]^2 + \text{TV} \left[ \partial_x \sqrt{u_*}(t; \cdot) \right]^2 \right) \, dt \right)^{1/2} \\
+ 4 \frac{\bar{t} - t}{b-a} \sup_{t \in [t, \bar{t}]} \left\| \left( \left\{ \sqrt{u_\Delta} \right\}_\tau - \sqrt{u_*} \right)(t; \cdot) \right\|^2_{L^\infty([a, b])}.
\]

Now recall the uniform bounds in (61) above, and the uniform convergence of \( u_\Delta \) to \( u_* \) on \([t, \bar{t}] \times [a, b]\), see (54). It is then clear that \( \partial_x \left\{ \sqrt{u_\Delta} \right\}_\tau \) converges to \( \partial_x \sqrt{u_*} \) in \( L^2([t, \bar{t}] \times [a, b]) \). Invoking the uniform convergence (54) again, the claim (59) follows.

To show square integrability of the limit, fix some \( T > 0 \). Below, \( N \) is always such that \( T < N\tau < T + 1 \). A direct calculation using (33) yields that

\[
4 \int_a^b \left( \partial_x \sqrt{u_\Delta} \right)^2 (x) \, dx = \int_0^M \frac{\partial_\xi \tilde{z}_\Delta^n (\xi)}{\tilde{z}_\Delta^n (\xi) \partial_\xi X_\Delta^n (\xi)} \, d\xi.
\]

From the properties of \( X_\Delta^n \) and \( \tilde{z}_\Delta^n \) as linear interpolations, one easily deduces that

\[
\frac{1}{\tilde{z}_\Delta^n (\xi) \partial_\xi X_\Delta^n (\xi)} \leq \frac{z_{k+\frac{1}{2}}^n}{z_{k-\frac{1}{2}}^n} + \frac{z_{k-\frac{1}{2}}^n}{z_{k+\frac{1}{2}}^n}.
\]
for all \( \xi \in (\xi_{k-\frac{1}{2}}, \xi_{k+\frac{1}{2}}) \). Therefore,

\[
4 \int_0^T \left( \int_a^b \{\sqrt{u_\Delta}\}_o^2 (t; x) \, dx \right) \, dt \\
\leq \tau \sum_{n=1}^N \delta \sum_{k \in \mathbb{Z}} \left( \frac{z^n_{k+\frac{1}{2}} - z^n_{k-\frac{1}{2}}}{\delta} \right) \left( \frac{z^n_{k+\frac{1}{2}}}{z^n_{k-\frac{1}{2}}} + \frac{z^n_{k-\frac{1}{2}}}{z^n_{k+\frac{1}{2}}} \right) \\
\leq 2 \left( \tau \sum_{n=1}^N \delta \sum_{k \in \mathbb{Z}} \left[ \frac{z^n_{k+\frac{1}{2}} - z^n_{k-\frac{1}{2}}}{\delta} \right]^4 \right)^{1/2} \left( \tau \sum_{n=1}^N \delta \sum_{k \in \mathbb{Z}} \left( \frac{z^n_{k+\frac{1}{2}}}{z^n_{k-\frac{1}{2}}} \right)^2 \right)^{1/2} \\
+ \left( \frac{z^n_{k+\frac{1}{2}}}{z^n_{k-\frac{1}{2}}} \right)^2 \right)^{1/2}.
\]

The two sums are \( \Delta \)-uniformly bounded, thanks to the estimates (41) and (43). By lower semi-continuity of norms, \( \sqrt{u_*} \) obeys the same bound. \( \square \)

### 4 Weak Formulation of the Limit Equation

To finish our discussion of convergence, we verify that the limit \( u_* \) obtained in the previous section is indeed a weak solution to (1). From now on, \( (\tilde{x}_n^\Delta)_{n=0}^\infty \) with its derived functions \( \tilde{u}_\Delta, \tilde{u}_\Delta^* \), \( X_\Delta \) is a (sub)sequence for which the convergence results stated in Propositions 19 and 20 hold. We continue to assume (35) and (36). The goal of this section is to prove the following.

**Proposition 21** For every \( \rho \in C^\infty([a, b]) \) with \( \rho'(a) = \rho'(b) = 0 \), and for every \( \psi \in C^\infty_c([0, T)) \),

\[
\int_0^\infty \psi'(t) \left( \int_a^b \rho(x) u_*(t; x) \, dx \right) \, dt + \psi(0) \int_a^b \rho(x) u^0(x) \, dx \\
+ \int_0^\infty \psi(t) \left( \int_a^b \left[ \rho''(x) \partial_x u_*(t; x) + 4 \rho''(x) \partial_x \sqrt{u_*(t; x)^2} \right] \, dx \right) \, dt = 0 \tag{62}
\]

For definiteness, fix a spatial test function \( \rho \in C^\infty([a, b]) \) with \( \rho'(a) = \rho'(b) = 0 \), and a temporal test function \( \psi \in C^\infty_c([0, T]) \) with supp \( \psi \subset [0, T) \) for a suitable \( T > 0 \). Let \( B > 0 \) be chosen such that

\[
\|\rho\|_{C^4([a, b])} \leq B, \quad \|\psi\|_{C^1_c([0, T])} \leq B. \tag{63}
\]
For convenience, we assume $\delta < 1$ and $\tau < 1$. Further, we introduce the short-hand notation

$$
\rho'(\bar{x}_n^\tau) = (\rho'(x_1^n), \ldots, \rho'(x_{K-1}^n)) \in \mathbb{R}^{K-1}.
$$

(64)

In the estimates that follow, the non-explicit constants possibly depend on $(b - a)$, $T$, $B$, and $\mathcal{H}$, but not $\Delta$. The two main steps in the proof of Proposition 21 are to establish the following estimates, respectively:

$$
e_{1,\Delta} := \left| \int_0^T \left( \psi'(t) \int_a^b \rho(x) \{\bar{u}_\tau(\cdot; x)\} \, dx + \psi(t) \left\{ \{\rho'(\bar{x}_\Delta), \nabla F_\delta(\bar{x}_\Delta)\}_{\tau} \right\} \right) \, dt
$$

$$
+ \psi(0) \int_a^b \rho(x) \bar{u}_\Delta^0(x) \, dx \right| \leq C((\delta F_\delta(\bar{x}_\Delta))^1/2 + (\tau F_\delta(\bar{x}_\Delta))),
$$

(65)

and

$$
e_{2,\Delta} := \left| \int_0^T \psi(t) \left( \int_a^b \left[ \rho''(x) \partial_x \{\bar{u}_\tau(\cdot; x)\} + 4\rho''(x) \partial_x \{\sqrt{\bar{u}_\Delta}(t; x)^2\} \right] \, dx
$$

$$
- \left\{ \{\rho'(\bar{x}_\Delta), \nabla F_\delta(\bar{x}_\Delta)\}_{\tau} \right\} \right) \, dt \right| \leq C\delta^{1/4}.
$$

(66)

We proceed by proving (65) and (66). At the end of this section, it is shown how the claim (62) follows from (65) and (66) on basis of the convergence for $\{\bar{u}_\tau\}_{\tau}$ obtained previously.

Proof of (65) Choose $N_{\tau} \in \mathbb{N}$ such that $N_{\tau} \tau \in (T, T+1)$. Then, using that $\psi(N_{\tau} \tau) = 0$, we obtain after “summation by parts”:

$$
- \int_0^T \psi'(t) \left( \int_a^b \rho(x) \{\bar{u}_\tau(\cdot; x)\} \, dx \right) \, dt
$$

$$
= - \sum_{m=1}^{N_{\tau}} \left( \int_{(m-1)\tau}^{m\tau} \psi'(t) \, dt \int_a^b \rho(x) \bar{u}_\tau^m(x) \, dx \right)
$$

$$
= - \tau \sum_{m=1}^{N_{\tau}} \left( \psi((m \tau) - \psi((m - 1)\tau) \int_0^M \rho \circ X^n_{\tau}(\xi) \, d\xi \right)
$$

$$
= \tau \sum_{n=1}^{N_{\tau}} \left( \psi((n \tau) - \psi((n - 1)\tau) \int_0^M \rho \circ X^n_{\tau} - \rho \circ X^{n-1}_{\tau}(\xi) \, d\xi \right)
$$

$$
+ \psi(0) \int_0^M \rho \circ X^0_{\tau}(\xi) \, d\xi.
$$

(67)
A Taylor expansion of the term in the inner integral yields

\[
\frac{\rho \circ X^n_{\Delta} - \rho \circ X^{n-1}_{\Delta}}{\tau} = \rho' \circ X^n_{\Delta} \left( \frac{X^n_{\Delta} - X^{n-1}_{\Delta}}{\tau} \right) + \frac{\tau}{2} \rho'' \circ \tilde{X} \left( \frac{X^n_{\Delta} - X^{n-1}_{\Delta}}{\tau} \right)^2.
\]

(68)

where \(\tilde{X}\) symbolizes suitable “intermediate values” in \([0, M]\). We analyze the first term on the right-hand side of (68): using the representation (20) of \(X_{\Delta}\) in terms of hat functions \(\theta_k\), we can write its integral as follows,

\[
\int_0^M \rho' \circ X^n_{\Delta} \left( \frac{X^n_{\Delta} - X^{n-1}_{\Delta}}{\tau} \right) d\xi = \sum_{k \in I^+} \left( \frac{x^n_k - x^{n-1}_k}{\tau} \right) \int_{\xi_{k-1}}^{\xi_{k+1}} \rho' \circ X^n_{\theta_k} d\xi.
\]

(69)

On the other hand, since

\[
\int_{\xi_{k-1}}^{\xi_{k+1}} \theta_k d\xi = \delta,
\]

(70)

the discrete evolution equation (32) yields that

\[
- \langle \rho'(\bar{x}^n_{\Delta}), \nabla_\delta F_\delta(\bar{x}^n_{\Delta}) \rangle = \left( \rho'(\bar{x}^n_{\Delta}), \frac{\bar{x}^n_{\Delta} - \bar{x}^{n-1}_{\Delta}}{\delta} \right)
\]

\[
= \sum_{k \in I^+} \left( \frac{x^n_k - x^{n-1}_k}{\tau} \right) \int_{\xi_{k-1}}^{\xi_{k+1}} \rho(x^n_k) \theta_k(\xi) d\xi.
\]

(71)

Finally, observing that

\[
|X^n_{\Delta}(\xi) - x^n_k| \leq (x^n_{k+1} - x^n_{k-1}) \quad \text{for each } \xi \in (\xi_{k-1}, \xi_{k+1}),
\]

we can estimate the difference of the terms in (69) and (71) with the help of the bound (63) on \(\rho\) as follows:

\[
\left| \int_0^M \rho' \circ X^n_{\Delta}(\xi) \left( \frac{X^n_{\Delta} - X^{n-1}_{\Delta}}{\tau} \right)(\xi) d\xi - \langle \rho'(\bar{x}^n_{\Delta}), \nabla_\delta F_\delta(\bar{x}^n_{\Delta}) \rangle \right|
\]

\[
\leq \sum_{k \in I^+} \left| \frac{x^n_k - x^{n-1}_k}{\tau} \right| \int_{\xi_{k-1}}^{\xi_{k+1}} |\rho' \circ X^n_{\Delta}(\xi) - \rho'(x^n_k)| \theta_k(\xi) d\xi
\]

\[
\leq B\delta \sum_{k \in I^+} \left| \frac{x^n_k - x^{n-1}_k}{\tau} \right| (x^n_{k+1} - x^n_{k-1}).
\]

(72)
As a final preparation for the proof of (65), observe that

\[ R' := \int_0^T \psi(t) \left\{ [\rho'(\bar{x}_\Delta), \nabla_\delta F_\delta(\bar{x}_\Delta)]_{\bar{\delta}}\right\}_\tau (t) \, dt \]

\[ - \tau \sum_{n=1}^{N_\tau} \psi((n-1)\tau) \langle \rho'(\bar{x}_\Delta^n), \nabla_\delta F_\delta(\bar{x}_\Delta^n) \rangle_{\bar{\delta}} \]

\[ \leq \left( \tau \sum_{n=1}^{N_\tau} \frac{1}{(n-1)\tau} \int_{(n-1)\tau}^{n\tau} \psi(t) \, dt - \psi((n-1)\tau) \right)^{1/2} \left( \tau \sum_{n=1}^{\infty} B^2 \| \nabla_\delta F_\delta(\bar{x}_\Delta^n) \|_{\bar{\delta}}^2 \right)^{1/2} \]

\[ \leq \left( (T+1)B^2 \tau^2 \right)^{1/2} (2B^2F_\delta(\bar{x}_\Delta^0))^{1/2} = C' F_\delta(\bar{x}_\Delta^0)^{1/2} \tau, \]

using the energy estimate (39). We are now ready to estimate \( e_{1,\Delta} \) in (65):

\[ e_{1,\Delta} \leq R' + B \tau \sum_{n=1}^{N_\tau} \left( \int_0^M \rho \circ X^n_{\Delta} - \rho \circ X^{n-1}_{\Delta} (\xi) \, d\xi \right) \]

\[ - \langle \rho'(\bar{x}_\Delta^n), \nabla_\delta F_\delta(\bar{x}_\Delta^n) \rangle_{\bar{\delta}} \]

\[ \leq R' + B \tau \sum_{n=1}^{N_\tau} \left( \int_0^M \rho' \circ X^n_{\Delta}(\xi) \left( \frac{X^n_{\Delta} - X^{n-1}_{\Delta}}{\tau} \right)(\xi) \, d\xi \right) \]

\[ - \langle \rho'(\bar{x}_\Delta^n), \nabla_\delta F_\delta(\bar{x}_\Delta^n) \rangle_{\bar{\delta}} + \frac{B \tau}{2} \int_0^M \left( \frac{X^n_{\Delta} - X^{n-1}_{\Delta}}{\tau} \right)^2 (\xi) \, d\xi \]

\[ \leq R' + B^2 \left( \tau \sum_{n=1}^{\infty} \delta \sum_{k \in \mathbb{I}^+_{n \Delta}} \left( \frac{x^n_k - x^{n-1}_k}{\tau} \right)^2 \right)^{1/2} \left( \tau \sum_{n=1}^{N_\tau} \delta \sum_{k \in \mathbb{I}_k^+} (x^n_{k+1} - x^n_{k-1})^2 \right)^{1/2} \]

\[ + \frac{B^2 \tau}{2} \sum_{n=1}^{\infty} \left\| \frac{X^n_{\Delta} - X^{n-1}_{\Delta}}{\tau} \right\|_{L^2([0,M])}^2 \]

\[ \leq C' \tau F_\delta(\bar{x}_\Delta^0)^{1/2} + B^2 (2(b-a)T)^{1/2} (\delta F_\delta(\bar{x}_\Delta^0))^{1/2} + B^2 \tau F_\delta(\bar{x}_\Delta^0), \]

where we have used the energy estimate (39), the equivalence (26) of metrics, and the bound (99) from the “Appendix.” □

The proof of (66) requires more calculations, which are distributed in a series of lemmata below. The first step is to derive a fully discrete weak formulation from (32).

**Lemma 22** With (64), one has that

\[ - \langle \rho'(\bar{x}_\Delta^n), \nabla_\delta F_\delta(\bar{x}_\Delta^n) \rangle_{\bar{\delta}} = A_1^n - A_2^n + A_3^n + A_4^n, \]  

(73)
where

\[
A^n_1 = \delta \sum_{k \in I_k^+} \left( \frac{z_{k+\frac{1}{2}} - z_{k-\frac{1}{2}}}{\delta} \right)^2 \left( \frac{z_{k+\frac{1}{2}} + z_{k-\frac{1}{2}}}{2} \right) \left( \frac{\rho'(x_{k+1}^n) - \rho'(x_{k-1}^n)}{\delta} \right),
\]

\[
A^n_2 = \delta \sum_{k \in I_k^+} \left( \frac{z_{k+\frac{1}{2}} - z_{k-\frac{1}{2}}}{\delta} \right)^2 \left( \frac{\left( \frac{z_{k+\frac{1}{2}}}{2} + \frac{z_{k-\frac{1}{2}}}{2} \right)^2}{\delta} \right) \rho''(x_k^n),
\]

\[
A^n_3 = \delta \sum_{k \in I_k^+} \left( \frac{z_{k+\frac{1}{2}} - z_{k-\frac{1}{2}}}{\delta} \right)^2 \left( \frac{\left( \frac{z_{k+\frac{1}{2}}}{2} + \frac{z_{k-\frac{1}{2}}}{2} \right)^2}{\delta} \right) \left( \frac{\rho'(x_{k+1}^n) - \rho'(x_{k}^n)}{\delta} - \rho'(x_{k-1}^n) \right) \rho''(x_k^n).
\]

\[
A^n_4 = \delta \sum_{k \in I_k^+} \left( \frac{z_{k+\frac{1}{2}} - z_{k-\frac{1}{2}}}{\delta} \right)^2 \left( \frac{\left( \frac{z_{k+\frac{1}{2}}}{2} + \frac{z_{k-\frac{1}{2}}}{2} \right)^2}{\delta} \right) \left( \frac{\rho'(x_{k-1}^n) - \rho'(x_{k}^n)}{\delta} - \rho'(x_{k}^n) \right) \rho''(x_k^n).
\]

**Proof** Fix some time index \( n \in \mathbb{N} \). Recall the representation of \( \nabla_\delta F_\delta \) from (30). By definition of \( \rho'(x_k^n) \), it follows via a “summation by parts” that

\[
- \langle \nabla_\delta F_\delta(\bar{x}_k^n), \rho'(\bar{x}_k^n) \rangle = -\delta \sum_{\kappa \in I_k^{1/2}} \left( \frac{z_{\kappa+\frac{1}{2}}}{2} - \frac{z_{\kappa-\frac{1}{2}}}{2} \right) \left( \frac{\rho'(x_{\kappa+\frac{1}{2}}^n) - \rho'(x_{\kappa-\frac{1}{2}}^n)}{\delta} \right)
\]

\[
= \delta \sum_{k \in I_k^+} \left( \frac{z_{k+\frac{1}{2}}}{2} - \frac{z_{k-\frac{1}{2}}}{2} \right) \left( \frac{\rho'(x_{k+1}^n) - \rho'(x_{k}^n)}{\delta} \right)
\]

Using the elementary identity (for arbitrary numbers \( \alpha_\pm \) and \( \beta_\pm \))

\[
\alpha_+ \beta_+ - \alpha_- \beta_- = \frac{(\alpha_+ - \alpha_-) \beta_+ + \beta_-}{2} + \frac{\alpha_+ + \alpha_-}{2}(\beta_+ - \beta_-),
\]
we obtain further:

\[ -\langle \nabla_\delta F(\bar{x}_n^A), \rho'(\bar{x}_n^A) \rangle_\delta \]

\[ = \delta \sum_{k \in \mathbb{I}_K^+} \left( \frac{z_{k+1}^n - z_{k-1}^n}{\delta} \right) \left( \frac{\left( \frac{z_{k+1}^n}{2} \right) - \left( \frac{z_{k-1}^n}{2} \right)}{2\delta} \right) \left( \frac{\rho'(x_{k+1}^n) - \rho'(x_{k-1}^n)}{\delta} \right) \]

\[ + \delta \sum_{k \in \mathbb{I}_K^+} \left( \frac{z_{k+1}^n - z_{k-1}^n}{\delta} \right) \left( \frac{\left( \frac{z_{k+1}^n}{2} \right) + \left( \frac{z_{k-1}^n}{2} \right)}{2} \right) \]

\[ \times \left( \frac{\rho'(x_{k+1}^n) - 2\rho'(x_k^n) + \rho'(x_{k-1}^n)}{\delta^2} \right). \]  

(74)

The sum in (74) is equal to \( A_1^n \). In order to see that the sum in (75) is equal to \(-A_2^n + A_3^n + A_4^n\), simply observe that the identity

\[ \frac{x_{k+1}^n - x_k^n}{\delta} + \frac{x_{k-1}^n - x_k^n}{\delta} = \frac{1}{z_{k+1}^n} - \frac{1}{z_{k-1}^n} = -\frac{z_{k+1}^n - z_{k-1}^n}{z_{k+1}^n z_{k-1}^n}, \]

makes the coefficient of \( \rho''(x_k^n) \) vanish. \( \square \)

**Lemma 23** There are positive constants \( C_1 \) to \( C_4 \)—expressible in \( (b - a) \), \( T \), \( B \) and \( \tilde{T} \)—such that for each \( N \) with \( N \tau < T \), one has

\[ R_1 := \tau \sum_{n=1}^{N} | A_1^n - 2 \int_0^M [\partial_\xi \bar{z}_A^n(\xi)]^2 \rho'' \circ X_A^n(\xi) \ d\xi | \leq C_1 \delta^{1/4}, \]

(76)

\[ R_2 := \tau \sum_{n=1}^{N} | A_2^n - \int_0^M \left[ \partial_\xi \bar{z}_A^n(\xi) \right]^2 \rho'' \circ X_A^n(\xi) \ d\xi | \leq C_2 \delta^{1/4}, \]

(77)

\[ R_3 := \tau \sum_{n=1}^{N} | A_3^n - \frac{1}{2} \int_0^M \partial_\xi \bar{z}_A^n(\xi) \rho''' \circ X_A^n(\xi) \ d\xi | \leq C_3 \delta^{1/4}, \]

(78)

\[ R_4 := \tau \sum_{n=1}^{N} | A_4^n - \frac{1}{2} \int_0^M \partial_\xi \bar{z}_A^n(\xi) \rho''' \circ X_A^n(\xi) \ d\xi | \leq C_4 \delta^{1/4}. \]

(79)

The proof of Lemma 23 can be found in the “Appendix.”

It remains to identify the integral expressions inside \( R_1 \) to \( R_4 \) with those in the weak formulation (62).
Lemma 24 One has that

\[ \int_0^M \partial_\xi \hat{z}_n^\Delta (\xi) \rho''' \circ X_n^\Delta (\xi) \ d\xi = \int_a^b \partial_x \hat{u}_n^\Delta (x) \rho''' (x) \ dx, \]  

(80)

\[ R_5 := \tau \sum_{n=1}^N \left| \int_0^M \left[ \partial_\xi \hat{z}_n^\Delta (\xi) \right]^2 \rho'' \circ X_n^\Delta (\xi) \ d\xi \right| - 4 \int_a^b \left( \partial_x \sqrt{\hat{u}_n^\Delta} \right)^2 (x) \rho'' (x) \ dx \leq C_5 \delta^{1/4}, \]  

(81)

where (81) holds for each \( N \) with \( N \tau \leq T \).

Proof The starting point is relation (33), that is

\[ \hat{z}_n^\Delta (\xi) = \hat{u}_n^\Delta \circ X_n^\Delta (\xi) \]  

(82)

for all \( \xi \in [0, M] \). Both sides of this equation are Lipschitz continuous in \( \xi \) and are differentiable except possibly at \( \xi \frac{1}{2}, \xi_1, \ldots, \xi_{K-\frac{1}{2}} \). At points \( \xi \) of differentiability, we have that

\[ \partial_\xi \hat{z}_n^\Delta (\xi) = \partial_x \hat{u}_n^\Delta \circ X_n^\Delta (\xi) \partial_\xi X_n^\Delta (\xi). \]

Substitute this expression for \( \partial_\xi \hat{z}_n^\Delta (\xi) \) into the left-hand side of (80) and perform a change of variables \( x = X_n^\Delta (\xi) \) to obtain the integral on the right.

Next, take the square root in (82) before differentiation, then calculate the square and divide by \( \partial_\xi X_n^\Delta (\xi) \) afterward:

\[ \frac{[\partial_\xi \hat{z}_n^\Delta (\xi)]^2}{4 \hat{z}_n^\Delta (\xi) \partial_\xi X_n^\Delta (\xi)} = \left( \partial_x \sqrt{\hat{u}_n^\Delta} \right)^2 \circ X_n^\Delta (\xi) \partial_\xi X_n^\Delta (\xi). \]

Performing the same change of variables as before, this proves that

\[ \int_0^M \frac{[\partial_\xi \hat{z}_n^\Delta (\xi)]^2}{\hat{z}_n^\Delta (\xi) \partial_\xi X_n^\Delta (\xi)} \rho'' \circ X_n^\Delta (\xi) \ d\xi = 4 \int_a^b \left( \partial_x \sqrt{\hat{u}_n^\Delta} \right)^2 (x) \rho'' (x) \ dx. \]  

(83)

It remains to estimate the difference between the \( \xi \)-integrals in (81) and in (83), respectively. To this end, observe that for each \( \xi \in (\xi_k, \xi_{k+\frac{1}{2}}) \) with some \( k \in \mathbb{N}_K^+ \), one has \( \partial_\xi X_n^\Delta (\xi) = 1/z_{k+\frac{1}{2}}^n \) and \( \hat{z}_n^\Delta (\xi) \in [z_{k-\frac{1}{2}}^n, z_{k+\frac{1}{2}}^n] \). Hence, for those \( \xi \),

\[ \left| 1 - \frac{1}{\hat{z}_n^\Delta (\xi) \partial_\xi X_n^\Delta (\xi)} \right| \leq \left| 1 - \frac{z_{k+\frac{1}{2}}^n}{z_{k-\frac{1}{2}}^n} \right|. \]
If instead \( \xi \in (\xi_{k-1}^2, \xi_k) \), then this estimate holds with the roles of \( z_n^{k+1} \) and \( z_n^{k-1} \) interchanged. Consequently, using once again (41) and (43),

\[
N \sum_{n=1}^{N} \left| \int_0^M \frac{\partial \xi z_n^{\Delta} (\xi)}{z_n^{\Delta} (\xi)} \rho'' \circ X_n^{\Delta} (\xi) \mathrm{d}\xi - \int_0^M \frac{\partial \xi z_n^{\Delta} (\xi)^2}{z_n^{\Delta} (\xi)} \rho'' \circ X_n^{\Delta} (\xi) \mathrm{d}\xi \right|
\]

\[
\leq B N \sum_{n=1}^{N} \left| \frac{1}{z_n^{\Delta} (\xi)} \frac{\partial \xi z_n^{\Delta} (\xi)}{\rho'' \circ X_n^{\Delta} (\xi)} \right| \left| 1 - \frac{1}{z_n^{\Delta} (\xi)} \frac{\partial \xi z_n^{\Delta} (\xi)}{\rho'' \circ X_n^{\Delta} (\xi)} \right| \mathrm{d}\xi
\]

\[
\leq B \left( \tau \sum_{n=1}^{\infty} \| \partial \xi z_n^{\Delta} \|_{L^4}^4 \right)^{1/2} \left( \tau \sum_{n=1}^{N} \delta \sum_{k \in \mathbb{N}_k} \left[ \left( 1 - \frac{z_n^{k+1}}{z_n^{k-1}} \right) \left( 1 - \frac{z_n^{k+1}}{z_n^{k-1}} \right) \right] \right)^{1/2}
\]

\[
\leq 3H^{1/2} (6(b-a)^2 T^{-1/2} H^{1/2} \delta^{1/2})^{1/2},
\]

since \( N \tau \leq T \) by hypothesis. This shows (81).

**Proof of (66)** Again, let \( N_\tau \in \mathbb{N} \) be such that \( N_\tau \tau \in (T, T + 1) \). Combining the discrete weak formulation (73), the change of variables formulae (80)–(81), and the definitions of \( R_1 \) to \( R_5 \), it follows that

\[
e_{2,\Delta} \leq BR_5 + B \tau \sum_{n=1}^{N_\tau} \left| \int_0^M \left[ \partial \xi z_n^{\Delta} \rho'' \circ X_n^{\Delta} (\xi) + \partial \xi z_n^{\Delta} (\xi)^2 \rho'' \circ X_n^{\Delta} (\xi) \right] \mathrm{d}\xi
\]

\[
- \left( A_1^n - A_2^n + A_3^n + A_4^n \right)
\]

\[
\leq B(R_1 + R_2 + R_3 + R_4 + R_5) \leq B(C_1 + C_2 + C_3 + C_4 + C_5) \delta^{1/4}.
\]

This implies the desired inequality (66).

We are now going to finish the proof of this section’s main result.

**Proof of Proposition 21** Thanks to (65)–(66), we know that

\[
\left| \int_0^T \psi' (t) \int_a^b \rho (x) \{ \bar{u}_\Delta \} \tau (t; x) \mathrm{d}x \mathrm{d}t + \psi (0) \int_a^b \rho (x) \bar{u}^0_\Delta (x) \mathrm{d}x
\]

\[
+ \int_0^T \psi (t) \int_a^b \left[ \rho'' (x) \partial_x \{ \bar{u}_\Delta \} \tau (t; x) + 4 \rho'' (x) \partial_x \left\{ \sqrt{\bar{u}_\Delta} \right\} (t; x) x \right] \mathrm{d}x \mathrm{d}t
\]

\[
\leq e_{1,\Delta} + e_{2,\Delta} \leq C \left( (\tau F_\delta (\bar{x}^0_\Delta)) + (\delta F_\delta (\bar{x}^0_\Delta))^{1/2} + \delta^{1/4} \right).
\]

By our assumption (36) on \( F_\delta (\bar{x}^0_\Delta) \), the expression on the right-hand side vanishes as \( \Delta \to 0 \). To obtain (62) in the limit \( \Delta \to 0 \), we still need to show the convergence of the integrals to their respective limits.
A technical tool is the observation that, for each $p \in [1, 4]$, $Q_p := \sup_{\Delta} \frac{N_{\tau}}{\delta} \sum_{n=1}^{N_{\tau}} (z_{n}^{\Delta})^p < \infty$, thanks to the key estimate (41) and to inequality (100) from the “Appendix.” For the first integral, we use that $\{\bar{u}\}_\tau$ converges to $u_*$ w.r.t. $W$, locally uniformly with respect to $t \in (0, T)$. Thus, clearly

$$\int_a^b \rho(x) \{\bar{u}\}_\tau (t; x) \, dx \to \int_a^b \rho(x) u_* (t; x) \, dx$$

for each $t \in (0, T)$. In order to pass to the limit with the time integral, we apply Vitali’s theorem. To this end, observe that

$$\int_0^T \left| \psi'(t) \int_a^b \rho(x) \{\bar{u}\}_\tau (t; x) \, dx \right|^2 \, dt \leq B^2 (b-a) \tau \sum_{n=1}^{N_{\tau}} \int_a^b \bar{u}_n^\Delta (x)^2 \, dx$$

$$= B^2 (b-a) \tau \sum_{n=1}^{N_{\tau}} \delta \sum_{\kappa \in \{1\}^1/2} z_{n,\kappa}^n \leq Q_1 B^2 (b-a).$$

Next, using the strong convergence from (59), it follows that

$$\partial_x \{\bar{u}\}_\tau = 2 \left\{ \sqrt{\bar{u}_\Delta} \right\}_\tau \partial_x \sqrt{\bar{u}_\Delta} \to 2 \sqrt{u_*} \partial_x \sqrt{u_*} = \partial_x u_*$$

strongly in $L^1([a, b])$, for almost every $t \in (0, T)$. Again, we apply Vitali’s theorem to conclude convergence of the time integral, on grounds of the following estimate:

$$\int_0^T \left| \psi(t) \int_a^b \rho'''(x) \partial_x \{\bar{u}\}_\tau \, dx \right|^2 \, dt \leq B^2 (b-a) \tau \sum_{n=1}^{N_{\tau}} \int_a^b \left( \partial_x \bar{u}_n^\Delta (x) \right)^2 \, dx$$

$$= B^2 (b-a) \tau \sum_{n=1}^{N_{\tau}} \delta \sum_{\kappa \in \{1\}^{1/2}} \left( \frac{z_{n,\kappa}^{n+1/2} - z_{n,\kappa-1/2}^{n-1/2}}{\delta} \right) ^2 \left( \frac{z_{n,\kappa}^{n+1/2} + z_{n,\kappa-1/2}^{n-1/2}}{2} \right) ^2$$

$$\leq B^2 (b-a) \left( \tau \sum_{n=1}^{\infty} \sum_{k \in \{1\}^{1/2}} \left( \frac{z_{n,k+1/2}^{n} - z_{n,k-1/2}^{n}}{\delta} \right) ^4 \right)^{1/2} \left( \tau \sum_{n=1}^{N_{\tau}} \delta \sum_{\kappa \in \{1\}^{1/2}} (z_{n}\kappa)^2 \right)^{1/2}$$

$$\leq 3 \tilde{H}^{1/2} Q_2^{1/2} B^2 (b-a),$$
where we have used (41). Finally, the strong convergence (59) also implies that
\[
\left( \partial_x \left[ \bar{u}_\Delta \right]_\tau \right)^2 \rightarrow \left( \partial_x \sqrt{u_*} \right)^2
\]
strongly in \( L^1([a, b]) \), for almost every \( t \in (0, T) \). One more time, we invoke Vitali’s theorem, using that
\[
\int_0^T \psi(t) \int_a^b \rho''(x) \partial_x \left\{ \sqrt{\hat{u}_\Delta} \right\}_\tau^2 (t; x) \, dx \, dt \leq B^2 \tau \sum_{n=1}^{N_\tau} \int_a^b \left( \partial_x \sqrt{\bar{u}_\Delta} \right)^4 (x) \, dx
\]
\[
\leq \frac{1}{2} B^2 \tau \sum_{n=1}^{N_\tau} \sum_{k \in I_\tau^+} \left( \frac{z_{k+\frac{1}{2}}^n - z_{k-\frac{1}{2}}^n}{\delta} \right)^2 \left( 1 - \frac{z_{k+\frac{1}{2}}^n}{z_{k-\frac{1}{2}}^n} \right)^2 \left( 1 - \frac{z_{k-\frac{1}{2}}^n}{z_{k+\frac{1}{2}}^n} \right)^2.
\]
\[
\leq B^2 \left( \tau \sum_{n=1}^{\infty} \sum_{k \in I_\tau^+} \left( \frac{z_{k+\frac{1}{2}}^n - z_{k-\frac{1}{2}}^n}{\delta} \right)^4 \right)^{1/2} \left( \tau \sum_{n=1}^{\infty} \sum_{k \in I_\tau^+} \left( 1 - \frac{z_{k+\frac{1}{2}}^n}{z_{k-\frac{1}{2}}^n} \right)^4 \right)^{1/2}.
\]

The two terms in the last line are uniformly controlled in view of (41) and (43), respectively.

\[\Box\]

5 Proof of Theorems 1 and 2

Below, we collect the results derived up to here to formally conclude the proofs of our main theorems.

**Proof of Theorem 1** Well posedness of the discrete scheme follows from Proposition 9. Positivity and mass conservation are immediate consequences of the construction: recall that \( \bar{u}_\Delta = u_\delta \{ \bar{x}_\Delta \} \), with \( u_\delta \) defined in (21). The monotonicity of \( H_\delta \) and \( F_\delta \) has been obtained in Lemma 12 and 11, respectively.

It remains to show the exponential decay (16) of \( H_\delta \). From (the proof of) Lemma 16, it follows for each \( n = 1, 2, \ldots \) that
\[
\mathcal{H}(\bar{u}_\Delta^n) - \mathcal{H}(\bar{u}_\Delta^{n-1}) = H_\delta(\bar{x}_\Delta^n) - H_\delta(\bar{x}_\Delta^{n-1}) \leq -\frac{\tau}{10(b-a)} TV \left[ \partial_x \sqrt{\bar{u}_\Delta^n} \right]^2. \tag{84}
\]
By the logarithmic Sobolev inequality on \([a, b]\), see, e.g., [25], and thanks to the fact that \( \partial_x \sqrt{\bar{u}_\Delta^n}(0) = 0 \), we further have that
\[
\mathcal{H}(\bar{u}_\Delta^n) \leq 2 \left( \frac{b-a}{\pi} \right)^2 \int_a^b \left( \partial_x \sqrt{\bar{u}_\Delta^n} \right)^2 \, dx \leq \frac{(b-a)^3}{2\pi^2} TV \left[ \partial_x \sqrt{\bar{u}_\Delta^n} \right]^2. \tag{85}
\]
Now combine (84) and (85) with the estimate (103) from the “Appendix” to conclude that

\[
\left(1 + \frac{\pi^2 \tau}{5(b-a)^4}\right) \mathcal{H}(\bar{u}_\Delta^n) \leq \mathcal{H}(\bar{u}_{\Delta}^{n-1}).
\]

From here, the claim (16) is obtained by induction on \(n\). \(\square\)

**Proof of Theorem 2** Local uniform convergence of the \([\bar{u}_\Delta]_\tau\) to a continuous limit function \(u_*\) is part of the conclusions of Proposition 19, see (52) and (54). The regularity \(\sqrt{u_*} \in L^2(\mathbb{R}_{\geq 0}; H^1([a, b]))\) has been observed in Proposition 20. The strong convergence stated in the same proposition implies that \(F(u_*)\) is “almost monotone”; indeed, thanks to (59) we may assume—passing to a further subsequence with \(\Delta \to 0\) if necessary—that

\[
\left\{\sqrt{\bar{u}_\Delta}\right\}_\tau(t) \to \sqrt{u_*(t)} \quad \text{strongly in } H^1([a, b]), \quad \text{for a.e. } t > 0,
\]

and therefore also

\[
2 \int_a^b \left\{\partial_x \sqrt{\bar{u}_\Delta}\right\}_\tau^2(t; x) \, dx \to F(u_*(t)), \quad \text{for a.e. } t > 0.
\]

On the other hand, arguing just like in the proof of (81), it follows that

\[
\int_0^\infty \left|\{F_\delta(\bar{u}_\Delta)\}_\tau - 2 \int_a^b \left\{\partial_x \sqrt{\bar{u}_\Delta}\right\}_\tau^2(t; x) \, dx\right| \, dt \leq C \delta^{1/4}.
\]

Now combine (86) and (87) with the fact that \(\{F_\delta(\bar{u}_\Delta)\}_\tau\) is decreasing in \(t\), for each \(\Delta\), and is \(\Delta\)-uniformly bounded above according to (49). By Helly’s selection principle, there exists a monotone \(f : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}\) such that \(F(u_*(t)) = f(t)\) for a.e. \(t > 0\). The proof of monotonicity for \(t \mapsto \mathcal{H}(u_*)\) is similar, but easier: here it suffices to use the local uniform convergence from (54).

Finally, the weak formulation (18) has been shown in Proposition 21. Simply observe that any \(\varphi \in C_c^\infty(\mathbb{R}_{\geq 0} \times [a, b])\) can be approximated by linear combinations of products \(\psi(t)\rho(x)\) with functions \(\psi \in C^\infty(\mathbb{R}_{\geq 0})\) and \(\rho \in C^\infty([a, b])\). \(\square\)

### 6 Numerical Results and Order of Consistency

The proof of convergence for our discretization given above is purely qualitative. In this last section, we study quantitative aspects of the convergence. First, we calculate the order of consistency for approximation of smooth and strictly positive solutions. Second, we report on the observed order of convergence in several numerical experiments.
6.1 Order of Consistency

The following proposition shows that our scheme is (formally) of first order in time and of second order in space.

Proposition 25 Suppose that \( X \in C^\infty(\mathbb{R}_{\geq 0} \times [0, M]) \) is a classical solution to

\[
\partial_t X = \partial_\xi (Z^2 \partial_\xi^2 Z),
\]

which is further such that \( Z = 1/\partial_\xi X \) is smooth and strictly positive. Let \( \Delta = (\tau; \delta) \) be a family of discretization parameters. Then, the corresponding restrictions \( \tilde{x}_\Delta \) of \( X \) to the respective meshes, given by \( x_n^k := X(n\tau; \xi_k) \) for \( n \in \mathbb{N} \) and \( k \in \{1, \ldots, K\} \), satisfy (13) with an error \( O(\delta^2) + O(\tau) \) as \( \Delta \to 0 \).

Proof Given \( \Delta = (\tau; \delta) \), introduce \( \tilde{Z} : \mathbb{R}_{\geq 0} \times [\delta/2, M - \delta/2] \to \mathbb{R}_{>0} \) by

\[
\tilde{Z}(t; \xi) = \frac{\delta}{X(t; \xi + \delta/2) - X(t; \xi - \delta/2)},
\]

which is a smooth and strictly positive function, thanks to the properties of \( X \). It is immediately seen that

\[
\partial_\xi^m \tilde{Z}(t; \xi) = \partial_\xi^m Z(t; \xi) + O(\delta^2),
\]

for each \( m \in \mathbb{N} \) and locally uniformly in \((t; \xi)\) as \( \Delta \to 0 \). By definition of \( \tilde{x}_\Delta \) as restriction of \( X \) to the grid with parameters \( \Delta \), one has

\[
z^n_k = \frac{\delta}{x^n_{k+\frac{1}{2}} - x^n_{k-\frac{1}{2}}} = \tilde{Z}(n\tau; \xi_k).
\]

Fix indices \( n \in \mathbb{N} \) and \( k \in \{1, \ldots, K - 1\} \). In the following, we abbreviate

\[
z_* = \tilde{Z}(n\tau; \xi_k), \quad z'_* = \partial_\xi Z(n\tau; \xi_k), \quad \ldots, \quad z''_* = \partial_t Z(n\tau; \xi_k).
\]

Relation (90) and a standard Taylor expansion of \( \tilde{Z} \) around \( \xi = \xi_k \) yield

\[
(z^n_{k+\frac{1}{2}})^2 \left( \frac{z^n_{k+\frac{1}{2}} - 2z^n_{k+\frac{1}{2}} + z^n_{k-\frac{1}{2}}}{\delta^2} \right) = (z_*^2 + \delta z_* z'_* + O(\delta^2)) \left( z''_* + \frac{\delta}{2} z'''_* + O(\delta^2) \right)
\]

\[
= z_*^2 z''_* + \frac{\delta}{2} (2z_* z'_* z''_* + z''_* z'''_* + O(\delta^2)).
\]
The same expansion—with \((-\delta)\) in place of \(\delta\)—is obtained for \(z_k^{n+1}\), \(z_k^{-}\) and \(z_k^{-}\), respectively. Therefore,

\[
\frac{1}{\delta} \left[ \left( \frac{z_k^{n+1} - 2z_k^n + z_k^{-}}{\delta^2} \right) - \left( \frac{z_k^{-} - 2z_k^n + z_k^{n-1}}{\delta^2} \right) \right] = 2z_k z'_k z''_k + z_k^2 z'''_k + O(\delta).
\] (91)

Next, observe that the expression on the left-hand side remains invariant under the simultaneous exchange of \(z_k^{n+1}\) with \(z_k^{-}\) and of \(z_k^{-}\) with \(z_k^{n-1}\). It follows that the odd terms in the Taylor expansion must vanish; thus, the approximation error on the right-hand side is actually of order \(O(\delta^2)\) rather than \(O(\delta)\). Further, using (90) and (89), the term of order \(\delta^0\) can be written as

\[
2z_k z'_k z''_k + z_k^2 z'''_k = \partial \tilde{Z}(n\tau; \xi_k) \frac{\partial^2}{\partial \xi^2} \tilde{Z}(n\tau; \xi_k) = \partial \tilde{Z}(n\tau; \xi_k) \frac{\partial^2}{\partial \xi^2} Z(n\tau; \xi_k) + O(\delta^2).
\]

On the left-hand side of (13), we obtain

\[
\frac{x_k^n - x_k^{n-1}}{\tau} = \frac{1}{\tau} \left( X(n\tau; \xi_k) - X((n-1)\tau; \xi_k) \right) = \partial \tau X(n\tau; \xi_k) + O(\tau),
\] (92)

thanks to the smoothness of \(X\) in time. Combining (91) and (92) with the continuous equation (88), we arrive at (13), with an error of \(O(\tau) + O(\delta^2)\). \(\square\)

6.2 Numerical Experiments

6.2.1 Non-uniform Meshes

In order to make our discretization more flexible, we allow non-equidistant mass grids in the following. That is, the mass discretization of \([0, M]\) is determined by \(K + 1\) points \(\xi_0, \xi_1, \xi_2, \ldots, \xi_{K-1}, \xi_K\), with

\[0 = \xi_0 < \xi_1 < \cdots < \xi_{K-1} < \xi_K = M,\]

and we replace accordingly the spatial mesh width \(\delta\) by the vector \(\vec{\delta} = (\delta_1, \delta_2, \ldots, \delta_{K-1/2})\) of distances \(\delta_k = \xi_{k+1/2} - \xi_{k-1/2}\) for \(k \in \mathbb{N}_{K/2}\). The piecewise constant density function \(\bar{u} \in \mathcal{P}_\delta([a, b])\) corresponding to a vector \(\bar{x} \in \mathbb{R}^{K-1}\) is now given by

\[\bar{u}(x) = z_k \quad \text{for} \quad x_{k-1/2} < x \leq x_{k+1/2}, \quad \text{with} \quad z_k = \frac{\delta_k}{x_{k+1/2} - x_{k-1/2}}.\]
Clearly, the metric structure changes as well: the scalar product $\langle \cdot, \cdot \rangle_\delta$ and the associated gradient $\nabla_\delta f$ of a function $f : \mathbb{R}^k \to \mathbb{R}$ are replaced by

$$\langle \vec{v}, \vec{w} \rangle_\delta = \sum_{k \in \mathbb{I}_+^k} \frac{\delta_{k+\frac{1}{2}} + \delta_{k-\frac{1}{2}}}{2} v_k w_k,$$

and

$$[\nabla_\delta f(\vec{x})]_k = \frac{2}{\delta_{k+\frac{1}{2}} + \delta_{k-\frac{1}{2}}} \partial x_k f(\vec{x}).$$

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

$$\frac{\vec{x}_n^\Delta - \vec{x}_{n-1}^\Delta}{\tau} = -\nabla_\delta \mathcal{F}_\delta(\vec{x}_n^\Delta)$$

attains the form

$$\frac{x_k^n - x_k^{n-1}}{\tau} = \frac{1}{\delta_k} \left[ \left( \frac{z_{k+\frac{1}{2}}^n - z_{k-\frac{1}{2}}^n}{\delta_{k+\frac{1}{2}}} \right)^2 \left( \frac{z_{k+\frac{3}{2}}^n - z_{k+\frac{1}{2}}^n}{\delta_{k+\frac{1}{2}}} - \frac{z_{k+\frac{1}{2}}^n - z_{k-\frac{1}{2}}^n}{\delta_k} \right) \right] - \frac{1}{\delta_{k-\frac{1}{2}}^n} \left( \frac{z_{k-\frac{1}{2}}^n - z_{k-\frac{3}{2}}^n}{\delta_{k-\frac{1}{2}}} - \frac{z_{k-\frac{1}{2}}^n - z_{k-\frac{1}{2}}^n}{\delta_{k-\frac{1}{2}}} \right). \tag{93}$$

### 6.2.2 Initial Condition

Our main experiments are carried out using the by now classical test case from [5], that is

$$u^0_0(x) = \epsilon + \cos^{16}(\pi x), \text{ on } [0, 1], \tag{94}$$

with $\epsilon = 10^{-3}$. The mass grid $\vec{\delta}$ is chosen in such a way that $u^0_0$ is a piecewise constant approximation of $u^0$ with respect to a **spatially uniform** grid. That is, we choose $\vec{\delta}$ such that the initial condition $\vec{x}_0^\Delta$ for $\vec{x}$ attains the simple form

$$x_k^0 = a + \frac{b - a}{K} k.$$

To construct $\vec{\delta}$, we first calculate the cumulative distribution function $U^0 : [0, 1] \to [0, M]$ by numerical integration of $u^0$,

$$U^0(x) = \int_a^x u^0(y) \, dy,$$

and then define $\xi_k := U^0(x_k^0)$, for $k = 0, 1, \ldots, K$. 

\[ Springer \]
**Remark 26** An equidistant mass grid 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 the evolution of the zones with low density is of particular interest in numerical studies of the DLSS equation, it is natural to use a non-uniform mass grid with an adapted spatial resolution, like the one defined above.

6.2.3 **Implementation**

From the initial condition $\bar{x}_n^0$, the fully discrete solution is calculated inductively by solving the implicit Euler scheme (93) for $\bar{x}_n^0$, given $\bar{x}_{n-1}^0$. 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 $\bar{x}_{n-1}^0$, we calculate $\bar{x}_n^0$ by means of the following algorithm:

\[
\begin{align*}
\bar{x} & := \bar{x}_{n-1}^0 \\
\text{repeat} & \\
\quad d\bar{x} & := -\left(\partial_x^2 F_\delta(\bar{x})\right)^{-1}\partial_x F_\delta(\bar{x}) \\
\quad \text{while } \bar{x} + d\bar{x} \text{ not monotone} & \\
\quad \quad d\bar{x} & := d\bar{x}/2 \\
\text{end} & \\
\bar{x} & := \bar{x} + d\bar{x} \\
\text{until } ||d\bar{x}||_{l^1} < \text{tol} & \text{ and } ||\partial_x F_\delta(\bar{x})||_{l^1} < \text{tol} \\
\bar{x}_n^0 & := \bar{x}.
\end{align*}
\]

In all of our experiments, we use $\text{tol} = 10^{-8}$. The slow convergence of the Newton iteration has been observed in situations where the density $\bar{u}_{n-1}^\Delta$ has steep gradients and/or intervals of very low values.

6.2.4 **Reference Solution**

For numerical estimation of the convergence rate, we use reference solutions that are calculated by means of a structurally different discretization of (1). Specifically, we employ the positivity-preserving scheme described in [11]. There, the authors introduce the auxiliary variable $v = \ln u$ and rewrite Eq. (1) in the particular form

\[
\partial_t v = -\frac{\partial_{xx}(u \partial_{xx} v)}{u}.
\] (95)

Formally, the boundary conditions (2) are equivalent to

\[
\partial_x v(t; x) = \partial_{xxx} v(t; x) = 0 \quad \text{for } x \in \{a, b\}.
\] (96)

The numerical discretization of (95)–(96) is performed by a semi-implicit finite difference approximation with step sizes $\tau_{\text{ref}}$ and $h_{\text{ref}} = (b - a)/K_{\text{ref}}$ in the $t$- and...
x-directions, respectively. More precisely, with $t_n := n\tau_{\text{ref}}$, and with $x_k$ for $k = 0, \ldots, K_{\text{ref}}$ being the $K_{\text{ref}} + 1$ equidistant grid points in $[a, b]$, the numerical approximations $u^n_k \approx u(t_n; x_k)$ and $v^n_k \approx v(t_n; x_k)$ are obtained—inductively with respect to $n$—for given vectors $u^{n-1} = (u_0^{n-1}, \ldots, u_{K_{\text{ref}}}^{n-1})$ and $v^{n-1} = (v_0^{n-1}, \ldots, v_{K_{\text{ref}}}^{n-1})$ by

1. solving the semi-implicit difference equation

$$\frac{v^n - v^{n-1}}{\tau_{\text{ref}}} = -\frac{\Delta h_{\text{ref}}}{\Delta_1 h_{\text{ref}}} (u^{n-1} \Delta h_{\text{ref}} v^n),$$

(97)

for the vector $v^n = (v_0^n, \ldots, v_{K_{\text{ref}}}^n)$, and

2. enforcing consistency via $u^n_k := \exp(v^n_k)$ for $k = 0, 1, \ldots, K_{\text{ref}}$.

In (97) above, $\Delta h_{\text{ref}}$ is the standard finite difference approximation of the second derivative with equidistant steps $h_{\text{ref}}$. The boundary conditions (96) are enforced using values $u^n_k$ and $v^n_k$ at “ghost points” in the obvious way, that is

$$u_{-1}^n = u_0^n, \quad u_{-2}^n = u_1^n, \quad u_{K_{\text{ref}}+1}^n = u_{K_{\text{ref}}}^n, \quad u_{K_{\text{ref}}+2}^n = u_{K_{\text{ref}}-1}^n,$$

and likewise for $v^n_k$.

Remark 27 In [11], the scheme is presented as a two-step method where the second step “corrects” the discrete solution if values of $v^n_k$ lie above a threshold. This correcting step is of no importance for the practicability of the scheme, but has only been introduced to avoid the increase in entropy due to numerical fluctuations of the total mass, see [11, Remark 3.1].

To produce reference solutions for the examples discussed below, the scheme above is implemented with $K_{\text{ref}} = 3200$ spatial grid points and a time step $\tau_{\text{ref}} = 5 \times 10^{-9}$. At a given time $T = N\tau_{\text{ref}}$, the respective reference profile $x \mapsto u_{\text{ref}}(T; x)$ is defined via piecewise linear interpolation of the respective values $u^n_k$.

6.2.5 Observed Rate of Convergence

Figure 1 provides a qualitative picture of the evolution with initial condition $u^0$: The plot on the left shows snapshots of the density function $\bar{u}_\Delta$, the plot on the right visualizes the motion of the mesh points $\{x_k\}_\tau$, associated with the Lagrangian maps $X_\Delta$ in time. It is clearly seen that the initial density has a very flat minimum (which is degenerate of order 16) at $x = 1/2$, which bifurcates into two sharper minima at later times and eventually becomes one single minimum again. This behavior underlines that comparison principles do not hold for the DLSS equation. Both figures have been generated using $K = 200$ spatial grid points and the time step size $\tau = 10^{-6}$.

For numerical estimation of the convergence rate, we have carried out two series of experiments. In the first series, we fix the time step size $\tau = 10^{-8}$ and vary the number of spatial grid points, $K = 25, 50, 100, 200$. Figure 2/left shows the $L^2$-norms of the differences between the (piecewise linear) reference profile $u_{\text{ref}}(T; \cdot)$ and the piecewise linear interpolations $\hat{u}_\Delta(T; \cdot)$ of the values $\bar{x}_\Delta$ obtained by our
Fig. 1 Left snapshots of the density profiles $\bar{u}_\Delta(t; \cdot)$ for the initial condition (94) at times $t = 0$ and $t = 10^j$, $j = -6, \ldots, -3$, using $K = 200$ grid points and the time step size $\tau = 10^{-6}$. Right associated particle trajectories

Fig. 2 Comparison of $\hat{u}_\Delta(T; \cdot)$ and $\ln \hat{u}_\Delta(T; \cdot)$ to the respective reference profiles, $u_{\text{ref}}(T; \cdot)$ and $\ln u_{\text{ref}}(T; \cdot)$, for the initial condition (94). The plots show the $L^2$-deviations at $T = 10^{-5}$. Left $K = 25, 50, 100, 200$ spatial grid points for fixed time step $\tau = 10^{-8}$. Right time steps $\tau = 10^{-5}, 5 \times 10^{-6}, 10^{-6}, 5 \times 10^{-7}, 10^{-7}, 5 \times 10^{-8}$ for $K = 800$ grid points

scheme, as defined in Sect. 2.6, evaluated at time $T = 10^{-5}$. It is clearly seen that the error decays almost perfectly with the expected rate of $\delta^2 \propto K^{-2}$. The plot also shows the $L^2$-difference of the logarithms, verifying a similar decay of the relative error, which is very sensitive to density deviations in the flat zone. We remark that a comparison between $u_{\text{ref}}(T; \cdot)$ and the piecewise constant density $\bar{u}_\Delta(T; \cdot)$ would yield a significantly lower convergence rate of only $\delta \propto K$: Apparently, the spatial interpolation error dominates the relevant numerical approximation error.

For the second series of experiments, we keep the spatial discretization parameter $K = 800$ fixed and run our scheme with the time step sizes $\tau = 10^{-5}, 5 \times 10^{-6}, 10^{-6}, 5 \times 10^{-7}, 10^{-7}, 5 \times 10^{-8}$, respectively. The corresponding $L^2$-deviation to the reference solution at $T = 10^{-5}$ is plotted in Fig. 2/Right. Again, $u_{\text{ref}}(T; \cdot)$ is compared to the piecewise linear interpolation $\hat{u}_\Delta(T; \cdot)$. The error is proportional to $\tau$. The same is true for the $L^2$-difference of the logarithms, which measures the relative
error. It is remarkable that our scheme still yields a good approximation if the target time $T$ is reached in a single step. For comparison: in the reference scheme, a time step size of less than $\tau \approx 10^{-8}$ must be used to obtain a profile that is in qualitative agreement (in the logarithmic plot) with the reference solution.

6.2.6 Discontinuous Initial Data

One of the conclusions of Theorem 2 is that the discrete approximation converges also for (a large class of) non-regular initial data $u^0$. For illustration of this feature, we consider the discontinuous initial density function

$$u^0_{\text{discont}} = \begin{cases} 
1 & \text{for } x \in [0, \frac{1}{3}] \cup \left[\frac{2}{3}, 1\right], \\
10^{-3} & \text{for } x \in \left(\frac{1}{3}, \frac{2}{3}\right)
\end{cases} \quad (98)$$

instead of $u^0$ from (94).

Figure 3 provides a qualitative picture of the fully discrete evolution (see below for a discussion of the discretization parameters): snapshots of the discrete density function $\bar{u}_\Delta$ are shown on the left, corresponding snapshots of the logarithmic density are shown on the right. Note that within a very short time, peaks of relatively high amplitudes are generated near the points where $u^0$ is discontinuous. The associated Lagrangian maps are visualized in Fig. 4/left. One observes an extremely fast motion of the moving mesh near the discontinuities.

Again, we wish to estimate the approximation error in dependence of the spatial resolution by comparing the profiles produced by our scheme with a reference solution at time $T = 10^{-8}$. To obtain a clear rate of convergence with respect to the spatial mesh width and to keep the number of Newton iterations per time step low, very small time increments were used: we started with a tiny $\tau = 10^{-14}$ to resolve the initial time layer up to $t \approx 10^{-12}$; afterward, we progressively increase $\tau$ in powers of ten after every one hundred iterations, until we reach $\tau = 10^{-9}$, thus reducing the running time and the accumulation of round-off errors. The reference solution could not be
calculated using the time step $\tau_{\text{ref}} = 5 \times 10^{-9}$ from before, but a much smaller step size of $\tau_{\text{ref}} = 10^{-16}$ was needed to avoid numerical instabilities during the initial phase of the evolution. Similarly as for our own scheme, the step size is progressively increased after each one thousand iterations, up to $\tau = 10^{-9}$.

For the numerical error estimation, we perform a series of experiments using $K = 25, 50, 100$ and 200 spatial grid points. To enhance spatial resolution, the spatial grid is not equidistant but finer close to the jump discontinuities. The $L^2$-differences of the densities have been evaluated at $T = 10^{-8}$, see Fig. 4/right. As before, we compare $u_{\text{ref}}(T; \cdot)$ to the piecewise linear interpolation $\tilde{u}_{\Delta}(T; \cdot)$. The observed rate of convergence is slightly worse than the rate $\delta^2 \propto K^{-2}$ predicted by the formal calculation for smooth solutions. To convey an idea of the behavior of the relative error, we also plotted the $L^2$-difference of the logarithms, which is in fair agreement with the rate $K^{-1.5}$.

Acknowledgements: The authors are indebted to Giuseppe Savaré for fruitful discussions on the subject, and especially for contributing the initial idea for the entropy preserving discretization scheme. We further thank the anonymous referees for their extremely careful reading of the initial manuscript and their suggestions for improvement.

Appendix 1: Some technical lemmas

Lemma 28 For each $p > 1$ and $\tilde{x} \in x_\delta$ with $\tilde{z} = z[\tilde{x}]$, one has that

$$\sum_{\kappa \in \mathbb{I}_{1/2}^2} \left( \frac{\delta}{z_{\kappa}} \right)^p = \sum_{\kappa \in \mathbb{I}_{1/2}^2} \left( x_{\kappa + \frac{1}{2}} - x_{\kappa - \frac{1}{2}} \right)^p \leq (b - a)^p. \quad (99)$$
Proof The first equality is simply the definition (22) of $z_\kappa$. Since trivially $x_{\kappa + \frac{1}{2}} - x_{\kappa - \frac{1}{2}} < b-a$ for each $\kappa \in \mathbb{I}_K^{1/2}$, and since $p-1 > 0$, it follows that

$$\sum_{\kappa \in \mathbb{I}_K^{1/2}} (x_{\kappa + \frac{1}{2}} - x_{\kappa - \frac{1}{2}})^p \leq (b-a)^{p-1} \sum_{\kappa \in \mathbb{I}_K^{1/2}} (x_{\kappa + \frac{1}{2}} - x_{\kappa - \frac{1}{2}}) = (b-a)^p.$$ 

\[ \Box \]

Lemma 29 For each $\vec{x} \in \mathfrak{x}_\delta$ with $\vec{z} = z[\vec{x}]$, one has that

$$\frac{\delta}{b-a} \leq z_\kappa \leq M^{1-1/q} \left( \frac{\delta}{b-a} \sum_{k} \frac{|z_{\kappa+\frac{1}{2}} - z_{\kappa-\frac{1}{2}}|}{\delta} \right)^{1/q} + \frac{M}{b-a} \quad \text{for all } \kappa \in \mathbb{I}_K^{1/2},$$

and consequently,

$$z_\kappa \leq (2MF_\delta[\vec{x}])^{1/2} + \frac{M}{b-a} \quad \text{for all } \kappa \in \mathbb{I}_K^{1/2}. \quad (100)$$

Proof The first estimate in (100) is an immediate consequence of the definition of $z_\kappa$ in (22). To prove the second estimate, let $\kappa^* \in \mathbb{I}_K^{1/2}$ be such that $z_{\kappa^*} = \max z_\kappa$. Observe that there exists a $\kappa_\ast \in \mathbb{I}_K^{-1}$ such that

$$z_{\kappa_\ast} \leq \frac{M}{b-a} \leq z_{\kappa^*}. \quad (102)$$

Writing out $z_{\kappa^*} - z_{\kappa_\ast}$ as a sum over differences of adjacent values of $z_k$ and applying the triangle and Cauchy Schwarz inequality, one obtains

$$z_{\kappa^*} - z_{\kappa_\ast} \leq \sum_{k} |z_{\kappa+\frac{1}{2}} - z_{\kappa-\frac{1}{2}}| \leq \left( \delta \sum_{k} \frac{1}{\delta} \right)^{1-1/q} \left( \delta \sum_{k} \frac{|z_{\kappa+\frac{1}{2}} - z_{\kappa-\frac{1}{2}}|}{\delta} \right)^{1/q}. \quad (101)$$

Now combine this with (102).

\[ \Box \]

Lemma 30 With $\vec{u}$ and $\vec{\bar{u}}$ being, respectively, the piecewise linear and the piecewise constant densities associated with a given vector $\vec{x}$, then

$$\mathcal{H}_{\delta}(\vec{x}) = \mathcal{H}(\vec{u}) \leq \mathcal{H}(\vec{\bar{u}}). \quad (103)$$

Proof First observe that

$$\int_0^1 \ln \left( p(1-\lambda) + q\lambda \right) d\lambda = \frac{p \ln p - q \ln q}{p-q} - 1 \geq \frac{1}{2} (\ln p + \ln q), \quad (104)$$

$$\therefore \text{ Springer}$$
which is an easy consequence of a Taylor expansion for the function \( s \mapsto (1 + s) \ln s \) around \( s = 1 \), substituting \( s = p/q \). On the one hand, we have that

\[
\int_a^b \tilde{u}(x) \ln \tilde{u}(x) \, dx = \delta \sum_{k \in 1/2} \ln z_k = \delta \frac{\ln z_{1/2} + \ln z_{K-1/2}}{2} + \delta \sum_{k \in \mathbb{Z}^+} \frac{\ln z_{k+1/2} + \ln z_{k-1/2}}{2},
\]

and on the other hand,

\[
\int_a^b \tilde{u}(x) \ln \tilde{u}(x) \, dx = \int_0^M \ln \tilde{z}(\xi) \, d\xi = \frac{\delta}{2} (\ln z_0 + \ln z_K) + \delta \sum_{k \in \mathbb{Z}^+} \int_0^1 \ln \left( z_{k-1/2} (1 - \lambda) + z_{k+1/2} \right) \, d\lambda
\geq \frac{\delta}{2} \frac{\ln z_{1/2} + \ln z_{K-1/2}}{2} + \delta \sum_{k \in \mathbb{Z}^+} \frac{\ln z_{k+1/2} + \ln z_{k-1/2}}{2},
\]

where we have used (104). This clearly implies (103).

\[\square\]

**Lemma 31** (Gagliardo–Nirenberg inequality). For each \( f \in H^1([a, b]) \), one has that

\[
\|f\|_{C^{1/6}([a, b])} \leq (9/2)^{1/3} \|f\|_{H^1([a, b])}^{2/3} \|f\|_{L^2([a, b])}^{1/3},
\]  

(105)

**Proof** Assume first that \( f \geq 0 \). Then, for arbitrary \( a < x < y < b \), the fundamental theorem of calculus and Hölder’s inequality imply that

\[
|f(x)^{3/2} - f(y)^{3/2}| \leq \frac{3}{2} \int_x^y \frac{1}{2} f'(z)^{1/2} |f'(z)| \, dz
\leq \frac{3}{2} |x - y|^{1/4} \|f\|_{L^2([a, b])}^{1/2} \|f\|_{L^2([a, b])}.
\]

Since \( f \geq 0 \), we can further estimate

\[
|f(x) - f(y)| \leq |f(x)^{3/2} - f(y)^{3/2}|^{2/3}
\leq (3/2)^{2/3} |x - y|^{1/6} \|f\|_{L^2([a, b])}^{1/3} \|f\|_{H^1([a, b])}.
\]

This shows (105) for nonnegative functions \( f \). A general \( f \) can be written in the form \( f = f_+ - f_- \), where \( f_\pm \geq 0 \). By the triangle inequality, and since \( \|f_\pm\|_{H^1([a, b])} \leq \|f\|_{H^1([a, b])} \),

\[
\|f\|_{C^{1/6}([a, b])} \leq \|f_\pm\|_{C^{1/6}([a, b])} + \|f_-\|_{C^{1/6}([a, b])} \leq 2(3/2)^{2/3} \|f\|_{L^2([a, b])}^{1/3} \|f\|_{H^1([a, b])}.
\]

This proves the claim.  

\[\square\]
Appendix 2: Proof of Lemma 23

Proof of estimate (76) First, observe that by definition of \( \tilde{z} \),

\[
\int_0^M \left[ \partial_\xi \tilde{z}_\Delta^n (\xi) \right]^2 \rho'' = \sum_{k \in \mathbb{K}^+} \left( \frac{z_{k+1/2}^n - z_{k-1/2}^n}{\delta} \right)^2 \int_{\xi_{k-1/2}^+}^{\xi_{k+1/2}^+} \rho'' \circ X^\Delta^n (\xi) \, d\xi,
\]

and therefore, by Hölder’s inequality,

\[
R_1 \leq R_{1,\alpha}^{1/2} R_{1,\beta}^{1/2},
\]

with, recalling (41),

\[
R_{1,\alpha} = \tau \sum_{n=1}^N \delta \sum_{k \in \mathbb{K}^+} \left( \frac{z_{k+1/2} - z_{k-1/2}}{\delta} \right)^4 \leq \tau \sum_{n=1}^\infty \| \tilde{z}_\Delta^n \|_{L^4([a,b])}^4 \leq 9 \mathcal{H},
\]

\[
R_{1,\beta} = \tau \sum_{n=1}^N \delta \sum_{k \in \mathbb{K}^+} \left[ \frac{z_{k+1/2} + z_{k-1/2}}{2} \frac{\rho'(x_{k+1}^n) - \rho'(x_{k-1}^n)}{\delta} \right] - \frac{2}{\delta} \int_{\xi_{k-1/2}^+}^{\xi_{k+1/2}^+} \rho'' \circ X^\Delta^n \, d\xi.
\]

To simplify \( R_{1,\beta} \), let us fix \( n \), and introduce \( \tilde{x}_k^+ \in (x_k^n, x_{k+1}^n) \) and \( \tilde{x}_k^- \in (x_{k-1}^n, x_k^n) \) such that

\[
\frac{\rho'(x_{k+1}^n) - \rho'(x_{k-1}^n)}{\delta} = \frac{\rho'(x_{k+1}^n) - \rho'(x_{k-1}^n)}{\delta} = \frac{\rho''(\tilde{x}_k^+)}{z_{k+1/2}^n} + \frac{\rho''(\tilde{x}_k^-)}{z_{k-1/2}^n}.
\]

For each \( k \in \mathbb{K}^+ \), we have that—recalling (70)—

\[
\frac{z_{k+1/2}^n + z_{k-1/2}^n}{\delta} \left( \frac{\rho''(\tilde{x}_k^+)}{z_{k+1/2}^n} + \frac{\rho''(\tilde{x}_k^-)}{z_{k-1/2}^n} \right) - \frac{2}{\delta} \int_{\xi_{k-1/2}^+}^{\xi_{k+1/2}^+} \rho'' \circ X^\Delta^n \, d\xi
\]

\[
= \frac{1}{2} \left( \frac{z_{k-1/2}^n}{z_{k+1/2}^n} + 1 \right) \rho''(\tilde{x}_k^+) + \left( \frac{z_{k+1/2}^n}{z_{k-1/2}^n} + 1 \right) \rho''(\tilde{x}_k^-) - \frac{2}{\delta} \int_{\xi_{k-1/2}^+}^{\xi_{k+1/2}^+} \rho'' \circ X^\Delta^n \, d\xi
\]
\[
= \frac{1}{2} \left[ \left( \frac{z_{k+1}^n}{z_k^{n+1}} - 1 \right) \rho''(\tilde{x}_k^+) + \left( \frac{z_{k-1}^n}{z_k^{n-1}} - 1 \right) \rho''(\tilde{x}_k^-) \right] \\
- \frac{2}{\delta} \int_{\xi_{k+1}^{1/2}}^{\xi_k^{1/2}} \left[ \rho'' \circ X^\Delta - \rho''(\tilde{x}_k^+) \right] d\xi - \frac{2}{\delta} \int_{\xi_{k-1}^{1/2}}^{\xi_k^{1/2}} \left[ \rho'' \circ X^\Delta - \rho''(\tilde{x}_k^-) \right] d\xi.
\]

Since \( X^\Delta(\xi) \in [x_k^n, x_{k+1}^n] \) for each \( \xi \in [\xi_k, \xi_{k+1}] \), and \( \tilde{x}_k^+ \in [x_k^n, x_{k+1}^n] \), it follows that \( |X^\Delta(\xi) - \tilde{x}_k^+| \leq x_{k+1}^n - x_k^n \), and therefore,

\[
\frac{2}{\delta} \int_{\xi_k^{1/2}}^{\xi_{k+1}^{1/2}} \left| \rho'' \circ X^\Delta(\xi) - \rho''(\tilde{x}_k^+) \right| d\xi \leq B(x_{k+1}^n - x_k^n). \quad (108)
\]

A similar estimate holds for the other integral. Thus,

\[
R_{1,\beta} \leq B^2 \tau \sum_{n=1}^{N} \sum_{k \in \mathbb{K}} \left[ \left( \frac{z_{k+1}^n}{z_k^{n+1}} - 1 \right)^2 + \left( \frac{z_{k-1}^n}{z_k^{n-1}} - 1 \right)^2 + 2(x_{k+1}^n - x_k^n)^2 \right].
\]

Combining the estimate (43) with inequality (99) from the “Appendix,” we further conclude that

\[
R_{1,\beta} \leq B^2 (6(b-a)^2(\overline{H} \delta)^1/2 + 4T(b-a)^2 \delta). \quad (109)
\]

In combination with (106) and (107), this proves the claim. \qed

**Proof of estimate (77)** The proof is almost identical to (and even easier than) the one for estimate (76) above. Again, we have a decomposition of the form

\[
R_2 \leq R_{2,\alpha}^{1/2} R_{2,\beta}^{1/2},
\]

where \( R_{2,\alpha} \) equals \( R_{1,\alpha} \) from (107), and

\[
R_{2,\beta} = \tau \sum_{n=1}^{N} \sum_{k \in \mathbb{K}} \left[ \frac{z_{k+1}^n + z_{k-1}^n}{2z_k^{n+1}z_k^{n-1}} \rho''(x_k^n) - \frac{1}{\delta} \int_{\xi_{k-1}^{1/2}}^{\xi_k^{1/2}} \rho'' \circ X^\Delta d\xi \right]^2.
\]

By writing

\[
\frac{(z_{k+1}^n)^2 + (z_{k-1}^n)^2}{2z_{k+1}^n z_{k-1}^n} = \frac{1}{2} \left( \frac{z_{k+1}^n}{z_k^{n+1}} - 1 \right) + \frac{1}{2} \left( \frac{z_{k-1}^n}{z_k^{n-1}} - 1 \right) + 1,
\]

\[ Springer \]

\[ \mathbb{K} \]
and observing—in analogy to (108)—that
\[
\frac{1}{\delta} \int_{\xi_{k-\frac{1}{2}}}^{\xi_{k+\frac{1}{2}}} \left| \rho'' \circ X^\alpha_\Delta (\xi) - \rho''(x^*_k) \right| \, d\xi \leq B \left( x^*_k + \frac{1}{2} - x^*_k - \frac{1}{2} \right),
\]
we obtain the same bound on \( R_{2,\beta} \) as the one on \( R_{1,\beta} \) from (109).

**Proof of estimate (78)** Arguing like in the previous proofs, we first deduce—now by means of Hölder’s inequality instead of the Cauchy–Schwarz inequality—that
\[
R_3 \leq R_{3,\alpha}^{1/4} R_{3,\beta}^{3/4},
\]
where \( R_{3,\alpha} = R_{1,\alpha} \), and
\[
R_{3,\beta} = \tau \sum_{n=1}^{N} \sum_{k \in \mathbb{I}^+} \left( \frac{\left( z^n_{k+\frac{1}{2}} \right)^2 + \left( z^n_{k-\frac{1}{2}} \right)^2}{2} \right) \rho''(x^*_k) - (x^*_k + 1 - x^*_k) \rho''(x^*_k) \right)
\]
\[
\times \frac{1}{\delta} \int_{\xi_{k-\frac{1}{2}}}^{\xi_{k+\frac{1}{2}}} \rho''' \circ X^n_\delta \, d\xi \bigg|^{4/3}.
\]

Introduce intermediate values \( \tilde{x}^*_k \) such that
\[
\rho'(x^*_k + 1) - \rho'(x^*_k) - (x^*_k + 1 - x^*_k) \rho''(x^*_k) = \frac{1}{2} \left( x^*_k + 1 - x^*_k \right)^2 \rho'''(\tilde{x}^*_k) = \frac{\delta^2}{2} \left( z^n_{k+\frac{1}{2}} \right)^2 \rho'''(\tilde{x}^*_k).
\]

Thus, we have that
\[
\left( \frac{\left( z^n_{k+\frac{1}{2}} \right)^2 + \left( z^n_{k-\frac{1}{2}} \right)^2}{2} \right) \rho''(x^*_k) - (x^*_k + 1 - x^*_k) \rho''(x^*_k) \right)
\]
\[
\times \frac{1}{\delta} \int_{\xi_{k-\frac{1}{2}}}^{\xi_{k+\frac{1}{2}}} \rho''' \circ X^n_\delta \, d\xi
\]
\[
= \frac{1}{4} \left( \frac{z^n_{k-\frac{1}{2}}}{z^n_{k+\frac{1}{2}}} + 1 \right) \rho'''(\tilde{x}^*_k) - \frac{1}{2\delta} \int_{\xi_{k-\frac{1}{2}}}^{\xi_{k+\frac{1}{2}}} \rho''' \circ X^n_\delta \, d\xi
\]
\[
= \frac{1}{4} \left( \frac{z^n_{k-\frac{1}{2}}}{z^n_{k+\frac{1}{2}}} + 1 \right) \rho'''(\tilde{x}^*_k) - \frac{1}{2\delta} \int_{\xi_{k-\frac{1}{2}}}^{\xi_{k+\frac{1}{2}}} \left[ \rho''' \circ X^n_\Delta - \rho'''(\tilde{x}^*_k) \right] \, d\xi.
\]
By the analogue of (108), it follows further that

$$R_{3,\beta} \leq 2B^{4/3} \tau \sum_{n=1}^{N} \delta \sum_{k \in I_K^+} \left( \frac{z_{n-1}^K}{z_{n+1/2}^K} + 1 \right)^{4/3} \left( \frac{z_{n-1}^K}{z_{n+1/2}^K} - 1 \right)^{4/3} \left( x_{k+1}^n - x_{k-1}^n \right)^{4/3}$$

$$\leq 2B^{4/3} \left( \tau \sum_{n=1}^{N} \delta \sum_{k \in I_K^+} \left( \frac{z_{n-1}^K}{z_{n+1/2}^K} + 1 \right)^{4} \right)^{1/3} \left( \tau \sum_{n=1}^{N} \delta \sum_{k \in I_K^+} \left( \frac{z_{n-1}^K}{z_{n+1/2}^K} - 1 \right)^{2} \right)^{2/3}$$

$$+ 2B^{4/3} T (b - a)^{4/3} \delta,$$

where we have used (99) from the “Appendix.” At this point, the estimates (42) and (43) are used to control the first and the second sum, respectively. □

**Proof of estimate (79)** Here, one proceeds in full analogy to the proof of estimate (78) above. □

**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 Math., 121 (2006), pp. 1–50.
3. J.-D. Benamou, G. Carlier, Q. Mérigot, and E. Oudet, *Discretization of functionals involving the monge-ampère operator*, arXiv preprint arXiv:1408.4536, (2014).
4. A. Blanchem, V. Calvez, and J. A. Carrillo, *Convergence of the mass-transport steepest descent scheme for the subcritical Patlak-Keller-Segel model*, SIAM J. Numer. Anal., 46 (2008), pp. 691–721.
5. P. M. Bleher, J. L. Lebowitz, and E. R. Speer, *Existence and positivity of solutions of a fourth-order nonlinear PDE describing interface fluctuations*, Comm. Pure Appl. Math., 47 (1994), pp. 923–942.
6. 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*, R. Soc. Lond. Philos. Trans. Ser. A Math. Phys. Eng. Sci., 357 (1999), pp. 1047–1077.
7. M. Bukal, E. Emmrich, and A. Jüngel, *Entropy-stable and entropy-dissipative approximations of a fourth-order quantum diffusion equation*, Numer. Math., 127 (2014), pp. 365–396.
8. M. Burger, J. A. Carrillo, and M.-T. Wolfram, *A mixed finite element method for nonlinear diffusion equations*, Kinet. Relat. Models, 3 (2010), pp. 59–83.
9. M. J. Caceres, J. A. Carrillo, and G. Toscani, *Long-time behavior for a nonlinear fourth-order parabolic equation*, Trans. Amer. Math. Soc., 357 (2005), pp. 1161–1175.
10. J. A. Carrillo, J. Dolbeault, I. Gentil, and A. Jüngel, *Entropy-energy inequalities and improved convergence rates for nonlinear parabolic equations*, Discrete Contin. Dyn. Syst. Ser. B, 6 (2006), pp. 1027–1050.
11. J. A. Carrillo, A. Jüngel, and S. Tang, *Positive entropic schemes for a nonlinear fourth-order parabolic equation*, Discrete Contin. Dyn. Syst. Ser. B, 3 (2003), pp. 1–20.
12. J. A. Carrillo and J. S. Moll, *Numerical simulation of diffusive and aggregation phenomena in nonlinear continuity equations by evolving diffeomorphisms*, SIAM J. Sci. Comput., 31 (2009/10), pp. 4305–4329.
13. J. A. Carrillo and D. Slepcev, *Example of a displacement convex functional of first order*, Calc. Var. Partial Differential Equations, 36 (2009), pp. 547–564.
14. J. A. Carrillo and G. Toscani, *Long-time asymptotics for strong solutions of the thin film equation*, Comm. Math. Phys., 225 (2002), pp. 551–571.
31. G. Grün and M. Rumpf

30. L. Gosse and G. Toscani

29. N. Gigli and J. Maas

28. L. Gosse and G. Toscani

27. A. Jüngel and G. Toscani

26. U. Gianazza, G. Savaré, and G. Toscani

25. J. Fischer

24. B. Düring, D. Matthes, and J. P. Milišić

23. B. Derrida, J. L. Lebowitz, E. R. Speer, and H. Spohn

22. D. Kinderlehrer and N. J. Walkington

21. R. C. MacCamy and E. Socolovsky

20. B. Derrida, J. L. Lebowitz, E. R. Speer, and H. Spohn

19. P. Degond, F. Méhats, and C. Ringhofer

18. J. Denzler and R. J. McCann

17. M. P. Gualdani, A. Jüngel, and G. Toscani

16. A. Jüngel and D. Matthes

15. A. Jüngel and R. Pinnau

14. A. Jüngel and D. Matthes

13. J. A. Carrillo and M.-T. Wolfram, A finite element method for nonlinear continuity equations in lagrangian coordinates. Working paper.

12. F. Cavalli and G. Naldi, A Wasserstein approach to the numerical solution of the one-dimensional Cahn-Hilliard equation, Kinet. Relat. Models, 3 (2010), pp. 123–142.

11. P. Degond, F. Mehats, and C. Ringhofer, Quantum energy-transport and drift-diffusion models, J. Stat. Phys., 118 (2005), pp. 625–667.

10. J. Denzler and R. J. McCann, Nonlinear diffusion from a delocalized source: affine self-similarity, time reversal, & nonradial focusing geometries, Ann. Inst. H. Poincaré Anal. Non Linéaire, 25 (2008), pp. 865–888.

9. B. Derrida, J. L. Lebowitz, E. R. Speer, and H. Spohn, Dynamics of an anchored Toom interface, J. Phys. A, 24 (1991), pp. 4805–4834.

8. B. Derrida, J. L. Lebowitz, E. R. Speer, and H. Spohn, Fluctuations of a stationary non-equilibrium interface, Phys. Rev. Lett., 67 (1991), pp. 165–168.

7. B. Düring, D. Matthes, and J. P. Milišić, A gradient flow scheme for nonlinear fourth order equations, Discrete Contin. Dyn. Syst. Ser. B, 14 (2010), pp. 935–959.

6. L. C. Evans, O. Savin, and W. Gangbo, Diffeomorphisms and nonlinear heat flows, SIAM J. Math. Anal., 37 (2005), pp. 737–751.

5. J. Fischer, Uniqueness of solutions of the Derrida-Lebowitz-Speer-Spohn equation and quantum drift-diffusion models, Comm. Partial Differential Equations, 38 (2013), pp. 2004–2047.

4. J. Fischer, Infinite speed of support propagation for the Derrida-Lebowitz-Speer-Spohn equation and quantum drift-diffusion models, NoDEA Nonlinear Differential Equations Appl., 21 (2014), pp. 27–50.

3. W. Gang, Z. Martin, and S. Waruhiu, The sharp log-Sobolev inequality on a compact interval, Involue, 7 (2014), pp. 181–186.

2. U. Gianazza, G. Savaré, and G. Toscani, The Wasserstein gradient flow of the Fisher information and the quantum drift-diffusion equation, Arch. Ration. Mech. Anal., 194 (2009), pp. 133–220.

1. N. Gigli and J. Maas, Gromov-Hausdorff convergence of discrete transportation metrics, SIAM J. Math. Anal., 43 (2011), pp. 879–899.

Gigli and J. Maas, Gromov-Hausdorff convergence of discrete transportation metrics, SIAM J. Math. Anal., 43 (2011), pp. 879–899.

L. Gosse and G. Toscani, Identification of asymptotic decay to self-similarity for one-dimensional filtration equations, SIAM J. Numer. Anal., 43 (2006), pp. 2590–2606 (electronic).

L. Gosse and G. Toscani, Lagrangian numerical approximations to one-dimensional convolution-diffusion equations, SIAM J. Sci. Comput., 28 (2006), pp. 1203–1227 (electronic).

G. Grün, On the convergence of entropy consistent schemes for lubrication type equations in multiple space dimensions, Math. Comp., 72 (2003), pp. 1251–1279 (electronic).

G. Grün and M. Rumpf, Nonnegativity preserving convergent schemes for the thin film equation, Numer. Math., 87 (2000), pp. 113–152.

M. P. Gualdani, A. Jüngel, and G. Toscani, A nonlinear fourth-order parabolic equation with nonhomogeneous boundary conditions, SIAM J. Math. Anal., 37 (2006), pp. 1761–1779 (electronic).

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

A. Jüngel, Transport equations for semiconductors, vol. 773 of Lecture Notes in Physics, Springer-Verlag, Berlin, 2009.

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

A. Jüngel and D. Matthes, The Derrida-Lebowitz-Speer-Spohn equation: existence, nonuniqueness, and decay rates of the solutions, SIAM J. Math. Anal., 39 (2008), pp. 1996–2015.

A. Jüngel and R. Pinnau, Global nonnegative solutions of a nonlinear fourth-order parabolic equation for quantum systems, SIAM J. Math. Anal., 32 (2000), pp. 760–777 (electronic).

A. Jüngel and R. Pinnau, A positivity-preserving numerical scheme for a nonlinear fourth order parabolic system, SIAM J. Numer. Anal., 39 (2001), pp. 385–406 (electronic).

A. Jüngel and G. Toscani, Exponential time decay of solutions to a nonlinear fourth-order parabolic equation, Z. Angew. Math. Phys., 54 (2003), pp. 377–386.

A. Jüngel and I. Violet, First-order entropies for the Derrida-Lebowitz-Speer-Spohn equation, Discrete Contin. Dyn. Syst. Ser. B, 8 (2007), pp. 861–877.

D. Kinderlehrer and N. J. Walkington, Approximation of parabolic equations using the Wasserstein metric, M2AN Math. Model. Numer. Anal., 33 (1999), pp. 837–852.

R. C. MacCamy and E. Socolovsky, A numerical procedure for the porous media equation, Comput. Math. Appl., 11 (1985), pp. 315–319. Hyperbolic partial differential equations, II.
43. D. Matthes, R. J. McCann, and G. Savaré, *A family of nonlinear fourth order equations of gradient flow type*, Comm. Partial Differential Equations, 34 (2009), pp. 1352–1397.
44. D. Matthes and H. Osberger, *Convergence of a variational Lagrangian scheme for a nonlinear drift diffusion equation*, ESAIM Math. Model. Numer. Anal., 48 (2014), pp. 697–726.
45. G. Russo, *Deterministic diffusion of particles*, Comm. Pure Appl. Math., 43 (1990), pp. 697–733.
46. S. Serfaty, *Gamma-convergence of gradient flows on Hilbert and metric spaces and applications*, Discrete Contin. Dyn. Syst., 31 (2011), pp. 1427–1451.
47. C. Villani, *Topics in optimal transportation*, vol. 58 of Graduate Studies in Mathematics, American Mathematical Society, Providence, RI, 2003.
48. L. Zhornitskaya and A. L. Bertozzi, *Positivity-preserving numerical schemes for lubrication-type equations*, SIAM J. Numer. Anal., 37 (2000), pp. 523–555.