

# A CONVERGENT LAGRANGIAN DISCRETIZATION FOR A NONLINEAR FOURTH ORDER EQUATION

DANIEL MATTHES AND 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. We prove that the discrete solutions are strictly positive and mass conserving. Further, they dissipate *both* the Fisher information and the logarithmic entropy. Numerical experiments illustrate the practicability of the scheme.

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 non-negative initial data with finite entropy, without any CFL type condition. The key ingredient in the proof is a discretized version of the classical entropy dissipation estimate.

## 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  $\Omega = [a, b]$ :

$$\partial_t u + \partial_x \left( u \partial_x \left( \frac{\partial_{xx} \sqrt{u}}{\sqrt{u}} \right) \right) = 0 \quad \text{for } t > 0 \text{ and } x \in \Omega, \quad (1)$$

$$\partial_x u = 0, \quad u \partial_x \left( \frac{\partial_{xx} \sqrt{u}}{\sqrt{u}} \right) = 0 \quad \text{for } t > 0 \text{ and } x \in \partial\Omega, \quad (2)$$

$$u = u^0 \quad \text{at } t = 0. \quad (3)$$

Equation (1) is known as the *DLSS* equation, where the acronym refers to Derrida, Lebowitz, Speer and Spohn, who introduced (1) in [18, 19] for studying interface fluctuations in the anchored Toom model. In the context of semi-conductor physics, (1) appears as a simplified *quantum drift diffusion equation* [16, 30].

The analytical treatment of (1) is far from trivial: see e.g. [4, 22, 28, 24, 32, 33] for results on existence and uniqueness of solutions in various different settings, and [8, 13, 9, 24, 32, 35, 39] 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 maximum 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 [4], 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 [23]. In order to deal with the general case — allowing arbitrary non-negative initial data  $u^0$  of finite entropy — a theory for non-negative weak solutions has been developed [24, 32] on grounds of the a priori regularity estimate

$$\sqrt{u} \in L_{\text{loc}}^2(\mathbb{R}_+; H^2(\Omega)), \quad (4)$$

---

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

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

$$\partial_t u + \partial_{xxxx} u - \partial_{xx} (\partial_x \sqrt{u})^2 = 0. \quad (5)$$

On the other hand, the problem (1)–(3) has several remarkable structural properties, and these eventually paved the way to a rigorous analysis. We list some of those properties:

- 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 are infinitely many (formal) Lyapunov functionals [4, 8, 31]. The two most important ones are the (logarithmic) *entropy*,

$$\mathcal{H}(u) = \int_a^b u \ln u dx - \mathcal{H}_0 \quad \text{with} \quad \mathcal{H}_0 = M \ln \left( \frac{M}{b-a} \right), \quad (6)$$

and the *Fisher information*,

$$\mathcal{F}(u) = \int_a^b (\partial_x \sqrt{u})^2 dx. \quad (7)$$

- The Fisher information is more than just a Lyapunov functional: in [24], 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  $\mathbb{W}$ . That is, formally one can write (1)&(2) as

$$\partial_t u = -\text{grad}_{\mathbb{W}} \mathcal{F}(u). \quad (8)$$

- Also  $\mathcal{H}$  is not “an arbitrary” Lyapunov functional: the  $L^2$ -Wasserstein gradient flow of  $\mathcal{H}$  is the *heat equation* [29],

$$\partial_s v = -\text{grad}_{\mathbb{W}} \mathcal{H}(v) = \partial_{xx} v, \quad (9)$$

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

$$\mathcal{F}(v(s)) = -\frac{d}{ds} \mathcal{H}(v(s)). \quad (10)$$

In view of (8), this relation makes the DLSS equation the “big brother” of the heat equation, see [17, 39] 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 non-negative (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 [6, 10, 34, 36]. 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 [6, 10, 34], or a variant of the Fisher information [6, 20, 36]. An exception is the discretization from [20], 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 [20], which lead to a discretization with a very rich structure: the scheme 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). By a small change of that discretization, we obtain a new scheme which still has all of these properties, but in addition also dissipates the logarithmic entropy. 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.

We emphasize that our scheme is the first one to preserve more than one Lyapunov functional for (1), and it is the only fully discrete scheme for which a rigorous convergence analysis is available.

Below, we give the “pragmatic” definition of our full discretization, which is actually very simple. In Section 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  $X(t; \cdot) : [0, M] \rightarrow \Omega$  — the so-called *pseudo-inverse distribution function* of  $u(t; \cdot)$  — such that

$$\xi = \int_0^{X(t; \xi)} u(t; x) dx, \quad \text{for each } \xi \in [0, M]. \quad (11)$$

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

$$\mathbf{F}(X) = \int_0^M \left[ \partial_\xi \left( \frac{1}{\partial_\xi X} \right) \right]^2 d\xi,$$

that is,

$$\partial_t X = \partial_\xi (Z^2 \partial_{\xi\xi} Z), \quad \text{where } Z(t; \xi) := \frac{1}{\partial_\xi X(t; \xi)} = u(t; X(t; \xi)). \quad (12)$$

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  $\bar{x}_\Delta^n = (x_k^n)$  a fully discrete solution on the  $\Delta$ -mesh, so that  $x_k^n$  approximates  $X(n\tau; k\delta)$ , then the  $x_k^n$  satisfy

$$\frac{x_k^n - x_k^{n-1}}{\tau} = \frac{1}{\delta} \left[ (z_{k+\frac{1}{2}}^n)^2 \left( \frac{z_{k+\frac{3}{2}}^n - 2z_{k+\frac{1}{2}}^n + z_{k-\frac{1}{2}}^n}{\delta^2} \right) - (z_{k-\frac{1}{2}}^n)^2 \left( \frac{z_{k+\frac{1}{2}}^n - 2z_{k-\frac{1}{2}}^n + z_{k-\frac{3}{2}}^n}{\delta^2} \right) \right], \quad (13)$$

where the values  $z_{\ell-\frac{1}{2}}^n = \delta/(x_\ell^n - x_{\ell-1}^n)$  are associated to the mid-points of the spatial grid. At each time step  $n \in \mathbb{N}$ ,  $\bar{x}_\Delta^n = (x_1^n, \dots, x_{K-1}^n)$  approximates a Lagrangian map, so we assume that  $\bar{x}_\Delta^n$  is monotone, i.e.,  $x_k^n > x_{k-1}^n$ , and in accordance with (11), we associate to  $\bar{x}_\Delta^n$  a piecewise constant function  $\bar{u}_\Delta^n : \Omega \rightarrow \mathbb{R}_+$  with

$$\bar{u}_\Delta^n(x) = \frac{\delta}{x_k^n - x_{k-1}^n} \quad \text{for } x_{k-1}^n < x < x_k^n.$$

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

$$\mathbf{H}_\delta(\bar{x}^n) = \delta \sum_{k=1}^K \log z_{k-\frac{1}{2}}^n, \quad \mathbf{F}_\delta(\bar{x}^n) = \sum_{k=1}^{K-1} \left( \frac{z_{k+\frac{1}{2}}^n - z_{k-\frac{1}{2}}^n}{\delta} \right)^2.$$

These choices are made such that  $\mathbf{H}_\delta$  is the restriction of  $\mathcal{H}$ , i.e.,  $\mathbf{H}_\delta(\bar{x}_\Delta^n) = \mathcal{H}(\bar{u}_\Delta^n)$ , and such that  $\mathbf{F}_\delta$  is related to  $\mathbf{H}_\delta$  in the same way (10) as  $\mathcal{F}$  is related to  $\mathcal{H}$ ; see Section 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  $\bar{x}_\Delta^0$ , a sequence of monotone  $\bar{x}_\Delta^n$  satisfying (13) can be constructed by inductively defining  $\bar{x}_\Delta^n$  as a global minimizer of*

$$\bar{x} \mapsto \frac{\delta}{2\tau} \sum_k (x_k - x_k^{n-1})^2 + \mathbf{F}_\delta(\bar{x}).$$

*This sequence of vectors  $\bar{x}_\Delta^n$  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,

$$\mathbf{H}_\delta(\bar{x}_\Delta^n) \leq \mathbf{H}_\delta(\bar{x}_\Delta^{n-1}) \quad \text{and} \quad \mathbf{F}_\delta(\bar{x}_\Delta^n) \leq \mathbf{F}_\delta(\bar{x}_\Delta^{n-1}).$$

- Equilibration: There is a constant  $r > 0$  only depending on  $b - a$  such that

$$\mathbf{H}_\delta(\bar{x}_\Delta^n) \leq \mathbf{H}_\delta(\bar{x}_\Delta^0) e^{-rn\tau}. \quad (14)$$

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  $\{\bar{u}_\Delta\}_\tau : \mathbb{R}_+ \times \Omega \rightarrow \mathbb{R}_+$ , which is given by

$$\{\bar{u}_\Delta\}_\tau(t; x) = \bar{u}_\Delta^n(x) \quad \text{for } (n-1)\tau < t \leq n\tau.$$

Further,  $\Delta$  symbolizes a whole sequence of mesh parameters from now on, and we write  $\Delta \rightarrow 0$  to indicate that  $\tau \rightarrow 0$  and  $\delta \rightarrow 0$  simultaneously.

**Theorem 2.** *Let a non-negative initial condition  $u^0$  with  $\mathcal{H}(u^0) < \infty$  be given. Choose initial conditions  $\bar{x}_\Delta^0$  such that  $\bar{u}_\Delta^0$  converges to  $u^0$  weakly as  $\Delta \rightarrow 0$ , and*

$$\bar{\mathcal{H}} := \sup_\Delta \mathbf{H}_\delta(\bar{x}_\Delta^0) < \infty \quad \text{and} \quad \lim_{\Delta \rightarrow 0} (\tau + \delta) \mathbf{F}_\delta(\bar{x}_\Delta^0) = 0. \quad (15)$$

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

- $\{\bar{u}_\Delta\}_\tau$  converges to  $u_*$  locally uniformly on  $\mathbb{R}_+ \times \Omega$ ,
- $\sqrt{u_*} \in L_{loc}^2(\mathbb{R}_{\geq 0}; H^1(\Omega))$ ,
- there are non-increasing functions  $f, h : \mathbb{R}_+ \rightarrow \mathbb{R}$  such that  $\mathcal{F}(u_*(t)) = f(t)$  and  $\mathcal{H}(u_*(t)) = h(t)$  for a.e.  $t > 0$ , and additionally  $h(t) \leq \bar{\mathcal{H}} e^{-rt}$  with the constant  $r > 0$  from (14),
- $u_*$  satisfies the following weak formulation of (1)<sup>ℓ</sup>(2), see (5):

$$\int_0^\infty \int_\Omega [\partial_t \varphi u_* + \partial_{xxx} \varphi \partial_x u_* + 4\partial_{xx} \varphi (\partial_x \sqrt{u_*})^2] dx dt + \int_\Omega \varphi(0, x) u^0(x) dx = 0 \quad (16)$$

for every test function  $\varphi \in C_c^\infty(\mathbb{R}_{\geq 0} \times \Omega)$  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 27, and this is also the observed rate of convergence in numerical experiments with smooth initial data  $u^0$ , see Section 6.2.4.
  - (3) Initial condition: We emphasize that our only hypothesis on  $u^0$  is  $\mathcal{H}(u^0) < \infty$ , which allows the same general initial conditions as in [24, 32]. If  $\mathcal{F}(u^0)$  happens to be finite, and also  $\sup_\Delta \mathbf{F}_\delta(\bar{x}_\Delta^0) < \infty$ , then the uniform convergence of  $\{\bar{u}_\Delta\}_\tau$  holds up to  $t = 0$ .

- (4) Long time behavior: *By means of the Csiszar-Kullback inequality, the exponential decay of  $\mathcal{H}(u_*(t))$  to zero implies exponential convergence of  $u_*$  to the constant function  $u_\infty \equiv M/(b-a)$  in  $L^1(\Omega)$ .*
- (5) No uniqueness: *Since our notion of solution is too weak to apply the uniqueness result from [22], we cannot exclude that different subsequences of  $\{u_\Delta\}_\tau$  converge to different limits.*

The idea to derive numerical discretizations for solution of Wasserstein gradient flows from the Lagrangian representation is not new in the literature, see e.g. [37] for a general treatise. Several practical schemes have been developed on grounds of the Lagrangian representation for this class of evolution problems, mainly for second-order diffusion equations [5, 7, 38, 41], but also for chemotaxis systems [3], for non-local aggregation equations [11, 14], and for variants of the Boltzmann equation [27]. For certain nonlinear fourth order equations, Lagrangian numerical schemes have been developed as well, e.g., for the Hele-Shaw flow [15] and for a class of thin film equations [27]. On the other hand, a rigorous analysis of stability and convergence of the *fully discrete* schemes is rare and apparently limited to the case of nonlinear diffusion in one space dimension, see [26, 40]. There are, however, results available for *semi-discrete* Lagrangian approximations, see e.g. [2, 21].

The primary challenge in our convergence analysis is to carry out all estimates under *no additional assumptions on the regularity* of the limit solution  $u_*$ . In particular, we do not exclude a priori the formation of zeros — and the induced loss of regularity — in the limit  $u_*$ , 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  $u_*$ . 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 derivation of the relevant estimates turns out to be much harder than for the original problem (1)–(3). For instance, we are able to prove a compactness estimate for  $\bar{u}_\Delta$ , but not for its inverse distribution function, although both estimates would be equivalent in a smooth setting. This forces us to switch back and forth between the original (1) and the Lagrangian (12) formulation of the DLSS equation.

We further remark that the convergence of a family of gradient flows to a limiting gradient flow has been thoroughly investigated on a very abstract level, see e.g. in [1, 42], using methods of  $\Gamma$ -convergence. Unfortunately, these appealing abstract results would not help to simplify our proof significantly, 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. Therefore, we decided to give a “hands-on proof”, which requires only very few elements from the general theory of metric gradient flows.

**1.4. Structure of the paper.** We start with a description of our Lagrangian discretization in Section 2; the fully discrete scheme is defined in Subsection 2.5. In Section 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  $\Delta \rightarrow 0$ . In Section 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 Section 5. Finally, Section 6 provides a consistency analysis and results from numerical simulations of (13).

**Acknowledgement.** 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.

## 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

$$\mathcal{P}(\Omega) = \left\{ u : \Omega \rightarrow \mathbb{R}_+ : \int_{\Omega} u(x) \, dx = M \right\}$$

the space of densities of total mass  $M$  on  $\Omega$ , and we endow  $\mathcal{P}(\Omega)$  with the  $L^2$ -Wasserstein metric  $\mathbb{W}$ . We refer to [43] 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  $L^1(\Omega)$ , and that the  $L^2$ -Wasserstein distance on  $\mathcal{P}(\Omega)$  is isometrically equivalent to the usual  $L^2$ -distance on the space

$$\mathfrak{X} = \{X : [0, M] \rightarrow \Omega : X \text{ continuous and strictly increasing, with } X(0) = a, X(M) = b\}$$

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

**Lemma 4.** *Given  $u^0, u^1 \in \mathcal{P}(\Omega)$ , introduce their Lagrangian maps  $X^0, X^1 \in \mathfrak{X}$  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,

$$\int_{\Omega} \varphi(x) u(x) \, dx = \int_0^M \varphi(X(\xi)) \, d\xi, \quad (17)$$

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

$$\mathbb{I}_K^+ = \{1, 2, \dots, K-1\}, \quad \mathbb{I}_K^0 = \mathbb{I}_K^+ \cup \{0, K\}, \quad \text{and} \quad \mathbb{I}_K^{1/2} = \left\{ \frac{1}{2}, \frac{3}{2}, \dots, K - \frac{1}{2} \right\}.$$

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

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

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

$$\mathfrak{r}_{\delta} = \{ \vec{x} = (x_1, \dots, x_{K-1}) \mid a < x_1 < \dots < x_{K-1} < b \} \subseteq (a, b)^{K-1}.$$

By definition,  $\vec{x} \in \mathfrak{r}_{\delta}$  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  $\mathfrak{X}$  of inverse distribution functions, we single out the  $(K-1)$ -dimensional open and convex subset

$$\mathfrak{X}_{\delta} = \{X \in \mathfrak{X} \mid X \text{ is affine on each } [\xi_{k-1}, \xi_k]\}.$$

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

$$X = \mathbf{X}_{\delta}[\vec{x}] = \sum_{k \in \mathbb{I}_K^0} x_k \theta_k, \quad (18)$$

where the  $\theta_k : [0, M] \rightarrow \mathbb{R}$  are the usual affine hat functions, with  $\theta_k(\xi_\ell) = \delta_{k,\ell}$ . Further, the density function  $\mathbf{u}_\delta[\vec{x}] \in \mathcal{P}(\Omega)$  associated to  $\mathbf{X}_\delta[\vec{x}]$  is

$$\mathbf{u}_\delta[\vec{x}](x) = \sum_{\kappa \in \mathbb{1}_K^{1/2}} z_\kappa \mathbf{1}_{(x_{\kappa-\frac{1}{2}}, x_{\kappa+\frac{1}{2}}]}(x), \quad (19)$$

where the vector

$$\vec{z} = \mathbf{z}_\delta[\vec{x}] = (z_{1/2}, \dots, z_{K-1/2}) \quad \text{of weights} \quad z_\kappa = \frac{\delta}{x_{\kappa+\frac{1}{2}} - x_{\kappa-\frac{1}{2}}} \quad (20)$$

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:

$$z_{-\frac{1}{2}} = z_{\frac{1}{2}}, \quad z_{K+\frac{1}{2}} = z_{K-\frac{1}{2}}. \quad (21)$$

We finally introduce the associated  $(K-1)$ -dimensional submanifold  $\mathcal{P}_\delta(\Omega) := \mathbf{u}_\delta[\mathfrak{r}_\delta] \subset \mathcal{P}(\Omega)$  as the image of the injective map  $\mathbf{u}_\delta : \mathfrak{r}_\delta \rightarrow \mathcal{P}_\delta(\Omega)$ .

**2.3. A metric on the ansatz space.** Below, we define a ‘‘Wasserstein-like’’ metric  $\mathbb{W}_\delta$  on the ansatz space  $\mathcal{P}_\delta(\Omega)$ . For motivation of that definition, observe that  $\mathcal{P}_\delta(\Omega)$  is a geodesic submanifold of  $\mathcal{P}(\Omega)$ , hence the restriction  $\widetilde{\mathbb{W}}_\delta$  of the genuine  $L^2$ -Wasserstein distance  $\mathbb{W}$  to  $\mathcal{P}_\delta(\Omega)$  appears as a natural candidate for  $\mathbb{W}_\delta$ . Thanks to the flatness of  $\mathbb{W}$  in one space dimension, see Lemma 4, the pull-back metric of  $\mathbb{W}$  on  $\mathfrak{r}_\delta$  induced by  $\mathbf{u}_\delta$  is a homogeneous quadratic form. More precisely,

$$\mathbb{W}(\mathbf{u}_\delta[\vec{x}^0], \mathbf{u}_\delta[\vec{x}^1])^2 = \sum_{k=1}^{K-1} (\vec{x}_k^1 - \vec{x}_k^0) \widetilde{\mathbb{W}}_{k\ell} (\vec{x}_\ell^1 - \vec{x}_\ell^0) \quad \text{for all } \vec{x}^0, \vec{x}^1 \in \mathfrak{r}_\delta, \quad (22)$$

where the positive matrix  $\widetilde{\mathbb{W}} \in \mathbb{R}^{(K-1) \times (K-1)}$  is tridiagonal. This approach has been followed in our previous work [40].

Here, we take a modified approach and use (22) to *define* a metric  $\mathbb{W}_\delta$  on  $\mathcal{P}_\delta(\Omega)$ , but with the simpler matrix  $\delta \mathbb{1}_{K-1}$  in place of  $\widetilde{\mathbb{W}}$  above. In other words: up to a factor  $\delta^{1/2}$ , the pull-back metric of  $\mathbb{W}_\delta$  via  $\mathbf{u}_\delta$  is the usual Euclidean distance on  $\mathfrak{r}_\delta$ .

**Remark 5.** *Our proof of convergence heavily relies on several explicit estimates of quantities with respect to the metric  $\mathbb{W}_\delta$ .*

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

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

the distance  $\mathbb{W}_\delta$  is conveniently written as

$$\mathbb{W}_\delta(\mathbf{u}_\delta[\vec{x}^0], \mathbf{u}_\delta[\vec{x}^1]) = \|\vec{x}^1 - \vec{x}^0\|_\delta.$$

In [40, Lemma 3.2], we have shown the following.

**Lemma 6.**  *$\mathbb{W}_\delta$  is equivalent to the Wasserstein metric restricted to  $\mathcal{P}_\delta(\Omega)$ , uniformly in  $K$ :*

$$\frac{1}{6} \mathbb{W}_\delta(u_0, u_1)^2 \leq \mathbb{W}(u_0, u_1)^2 \leq \mathbb{W}_\delta(u_0, u_1)^2 \quad \text{for all } u_0, u_1 \in \mathcal{P}_\delta(\Omega). \quad (23)$$

Note that, as a direct consequence of (23), we obtain that

$$\|\mathbf{X}_\delta[\bar{x}^0] - \mathbf{X}_\delta[\bar{x}^1]\|_{L^2([0,M])} \leq \|\bar{x}^0 - \bar{x}^1\|_\delta.$$

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  $\mathcal{P}_\delta(\Omega)$ . However, Theorem 2 validates our choice a posteriori. For results concerning the  $\Gamma$ -convergence of discretized transport metrics to the Wasserstein distance see [25].

**2.4. Functions on  $\mathcal{P}_\delta(\Omega)$ .** When discussing functions on  $\mathcal{P}_\delta(\Omega)$  in the following, we always assume that these are given in the form  $f : \mathfrak{r}_\delta \rightarrow \mathbb{R}$ . We denote the first and second derivatives of  $f$  by  $\partial_{\bar{x}} f : \mathfrak{r}_\delta \rightarrow \mathbb{R}^{K-1}$  and by  $\partial_{\bar{x}}^2 f : \mathfrak{r}_\delta \rightarrow \mathbb{R}^{(K-1) \times (K-1)}$ , respectively, with components

$$[\partial_{\bar{x}} f(\bar{x})]_k = \partial_{x_k} f(\bar{x}) \quad \text{and} \quad [\partial_{\bar{x}}^2 f(\bar{x})]_{k,l} = \partial_{x_k} \partial_{x_l} f(\bar{x}). \quad (24)$$

**Example 7.** Each component  $z_\kappa$  of  $\bar{z} = \mathbf{z}_\delta[\bar{x}]$  is a function on  $\mathfrak{r}_\delta$ , and

$$\partial_{\bar{x}} z_\kappa = -z_\kappa^2 \frac{\mathbf{e}_{\kappa+\frac{1}{2}} - \mathbf{e}_{\kappa-\frac{1}{2}}}{\delta}, \quad (25)$$

where  $\mathbf{e}_k \in \mathbb{R}^{K-1}$  is the  $k$ th canonical unit vector, with the convention  $\mathbf{e}_0 = \mathbf{e}_K = 0$ .

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  $\vec{v} \in \mathbb{R}^{K-1}$ ,

$$\langle \vec{v}, \nabla_\delta f(\bar{x}) \rangle_\delta = \sum_{k=1}^{K-1} v_k \partial_{x_k} f(\bar{x}).$$

The *gradient flow* of a function  $f$  on  $\mathcal{P}_\delta(\Omega)$  with respect to  $\mathbb{W}_\delta$  is then defined as the solution  $\bar{x} : [0; \infty) \rightarrow \mathfrak{r}_\delta$  for the system of ordinary differential equations

$$\dot{\bar{x}} = -\nabla_\delta f(\bar{x}), \quad \text{or, more explicitly,} \quad \dot{x}_k = -\delta^{-1} \partial_{x_k} f(\bar{x}), \quad \text{for each } k \in \mathbb{I}_K^+. \quad (26)$$

**2.4.1. The discretized Boltzmann entropy.** The Boltzmann entropy  $\mathcal{H}$  as defined in (6) is a non-negative functional on  $\mathcal{P}(\Omega)$ , which vanishes precisely on the constant function  $u \equiv M/(b-a)$ . In analogy to [40], we introduce a discretization  $\mathbf{H}_\delta : \mathfrak{r}_\delta \rightarrow \mathbb{R}$  of the Boltzmann entropy  $\mathcal{H}$  by restriction to  $\mathcal{P}_\delta(\Omega)$ :

$$\mathbf{H}_\delta(\bar{x}) := \mathcal{H}(\mathbf{u}_\delta[\bar{x}]) = \int_\Omega \mathbf{u}_\delta[\bar{x}] \ln \mathbf{u}_\delta[\bar{x}] dx - \mathcal{H}_0 = \delta \sum_{\kappa \in \mathbb{I}_K^{1/2}} \ln z_\kappa - \mathcal{H}_0,$$

where  $\mathcal{H}_0$  was defined in (6), and  $\bar{z} = \mathbf{z}_\delta[\bar{x}]$ . Naturally,  $\mathbf{H}_\delta$  inherits non-negativity, and vanishes only for  $\bar{x}$  with  $x_k = a + (b-a)k/K$ . For the derivatives, we obtain — using the rule (25) —

$$\partial_{\bar{x}} \mathbf{H}_\delta(\bar{x}) = -\delta \sum_{\kappa \in \mathbb{I}_K^{1/2}} z_\kappa \frac{\mathbf{e}_{\kappa-\frac{1}{2}} - \mathbf{e}_{\kappa+\frac{1}{2}}}{\delta} = \delta \sum_{k \in \mathbb{I}_K^+} \frac{z_{k+\frac{1}{2}} - z_{k-\frac{1}{2}}}{\delta} \mathbf{e}_k, \quad (27)$$

$$\partial_{\bar{x}}^2 \mathbf{H}_\delta(\bar{x}) = \delta \sum_{\kappa \in \mathbb{I}_K^{1/2}} z_\kappa^2 \left( \frac{\mathbf{e}_{\kappa-\frac{1}{2}} - \mathbf{e}_{\kappa+\frac{1}{2}}}{\delta} \right) \left( \frac{\mathbf{e}_{\kappa-\frac{1}{2}} - \mathbf{e}_{\kappa+\frac{1}{2}}}{\delta} \right)^T. \quad (28)$$

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

2.4.2. *The discretized Fisher information.* The discrete Fisher information  $\mathbf{F}_\delta : \mathfrak{r}_\delta \rightarrow \mathbb{R}$  is *not* defined by restriction of  $\mathcal{F}$  from (7). Instead, we mimic (10) and define accordingly

$$\mathbf{F}_\delta(\vec{x}) = \frac{1}{2} \|\nabla_\delta \mathbf{H}_\delta(\vec{x})\|_\delta^2 = \frac{\delta}{2} \sum_{k \in \mathbb{I}_K^+} \left( \frac{z_{k+\frac{1}{2}} - z_{k-\frac{1}{2}}}{\delta} \right)^2,$$

using (27). Thanks to this simple structure, the gradient flow equation for  $\mathbf{F}_\delta$  has an explicit and compact representation. Using the rule (25), the representation (28) and the convention (21), we obtain with  $\vec{z} = \mathbf{z}_\delta[\vec{x}]$ :

$$\begin{aligned} \nabla_\delta \mathbf{F}_\delta(\vec{x}) &= \delta^{-2} \partial_{\vec{x}}^2 \mathbf{H}_\delta(\vec{x}) \partial_{\vec{x}} \mathbf{H}_\delta(\vec{x}) \\ &= \sum_{\kappa \in \mathbb{I}_K^{1/2}, k \in \mathbb{I}_K^+} z_\kappa^2 \left( \frac{z_{k+\frac{1}{2}} - z_{k-\frac{1}{2}}}{\delta} \right) \left( \frac{\mathbf{e}_{\kappa+\frac{1}{2}} - \mathbf{e}_{\kappa-\frac{1}{2}}}{\delta} \right) \left( \frac{\mathbf{e}_{\kappa+\frac{1}{2}} - \mathbf{e}_{\kappa-\frac{1}{2}}}{\delta} \right)^T \mathbf{e}_k \end{aligned} \quad (29)$$

$$= \sum_{\kappa \in \mathbb{I}_K^{1/2}} z_\kappa^2 \left( \frac{z_{\kappa+1} - 2z_\kappa + z_{\kappa-1}}{\delta^2} \right) \left( \frac{\mathbf{e}_{\kappa+\frac{1}{2}} - \mathbf{e}_{\kappa-\frac{1}{2}}}{\delta} \right). \quad (30)$$

This should be understood as a discretization of the differential operator  $(Z^2 Z_{\xi\xi})_\xi$  appearing on the right hand side of (12).

**Remark 8.** *Without calculating the second derivative  $\partial_{\vec{x}}^2 \mathbf{F}_\delta$  explicitly, we remark that it is unbounded from below on  $\mathfrak{X}_\delta$ , hence  $\mathbf{F}_\delta$  is not  $\lambda$ -convex for any  $\lambda \in \mathbb{R}$ . This is in agreement with the fact that already the original Fisher information  $\mathcal{F}$  is not geodesically  $\lambda$ -convex in the Wasserstein metric, see [12].*

2.5. **Time stepping.** For the definition of the fully discrete scheme for solution of (8), we discretize the spatially discrete gradient flow equation

$$\dot{\vec{x}} = -\nabla_\delta \mathbf{F}_\delta(\vec{x}) \quad (31)$$

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  $\vec{y} \in \mathfrak{r}_\delta$ , introduce the *Yosida-regularized Fisher information*  $\mathbf{F}_\Delta(\cdot; \vec{y}) : \mathfrak{r}_\delta \rightarrow \mathbb{R}$  by

$$\mathbf{F}_\Delta(\vec{x}; \vec{y}) = \frac{1}{2\tau} \|\vec{x} - \vec{y}\|_\delta^2 + \mathbf{F}_\delta(\vec{x}).$$

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

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

$$\frac{\vec{x} - \vec{x}_\Delta^{n-1}}{\tau} = -\nabla_\delta \mathbf{F}_\delta(\vec{x}). \quad (32)$$

Using the explicit representation (30) of  $\partial_{\vec{x}} \mathbf{F}_\delta$ , 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  $\vec{x}^0 \in \mathfrak{r}_\delta$ , the sequence of equations (32) can be solved inductively. Moreover, if  $\tau > 0$  is sufficiently small with respect to  $\delta$  and  $\mathbf{F}_\delta(\vec{x}^0)$ , then each equation (32) possesses a unique solution with  $\mathbf{F}_\delta(\vec{x}) \leq \mathbf{F}_\delta(\vec{x}^0)$ , and that solution is the unique global minimizer of  $\mathbf{F}_\Delta(\cdot; \vec{x}_\Delta^{n-1})$ .*

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  $\vec{y} \in \mathfrak{r}_\delta$  with  $\mathbf{F}_\delta(\vec{y}) \leq C$ , the following are true:*

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

*Proof.* First, observe that the sublevel  $A_C := \mathbf{F}_\delta^{-1}([0, C + 1]) \subset \mathfrak{r}_\delta$  is a compact subset of  $\mathbb{R}^{K-1}$ . Indeed,  $A_C$  is a relatively closed subset of  $\mathfrak{r}_\delta$  by continuity of  $\mathbf{F}_\delta$ . Moreover, thanks to (103), every  $\vec{x} \in A_C$  satisfies  $x_{\kappa+\frac{1}{2}} - x_{\kappa-\frac{1}{2}} \geq \underline{x}$  for all  $\kappa \in \mathbb{I}_K^{1/2}$  with a positive constant  $\underline{x}$  that depends on  $C$  only. Thus  $A_C$  does not touch the boundary (in the ambient  $\mathbb{R}^{K-1}$ ) of  $\mathfrak{r}_\delta$ . Consequently,  $A_C$  is closed and bounded in  $\mathbb{R}^{K-1}$ .

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

Since  $\mathbf{F}_\delta : \mathfrak{r}_\delta \rightarrow \mathbb{R}$  is smooth, its restriction to  $A_C$  is  $\lambda_C$ -convex with some  $\lambda_C \leq 0$ , i.e.,  $\partial_{\vec{x}}^2 \mathbf{F}_\delta(\vec{x}) \geq \lambda_C \mathbf{1}_{K-1}$  for all  $\vec{x} \in A_C$ . Independently of  $\vec{y}$ , we have that

$$\partial_{\vec{x}}^2 \mathbf{F}_\Delta(\vec{x}; \vec{y}) = \partial_{\vec{x}}^2 \mathbf{F}_\delta(\vec{x}) + \frac{\delta}{\tau} \mathbf{1}_{K-1},$$

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

$$0 < \tau < \tau_C := \frac{\delta}{(-\lambda_C)}.$$

Consequently, each such  $\mathbf{F}_\Delta(\cdot; \vec{y})$  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  $\vec{z}_\Delta^n = \mathbf{z}_\delta[\vec{x}_\Delta^n]$  as  $x_k$  and  $z_k$ , respectively, whenever there is no ambiguity in the choice of  $\Delta$  and the time step  $n$ .

Recall that  $u_\Delta^n = \mathbf{u}_\delta[\vec{x}_\Delta^n] \in \mathcal{P}_\delta(\Omega)$  defines a sequence of densities on  $[a, b]$  which are piecewise constant with respect to the (non-uniform) grid  $(a, x_1, \dots, x_{K-1}, b)$ . To facilitate the study of convergence of weak derivatives, we introduce also *piecewise affine* interpolations  $\widehat{z}_\Delta^n : [0, M] \rightarrow \mathbb{R}_+$  and  $\widehat{u}_\Delta^n : [a, b] \rightarrow \mathbb{R}_+$ .

In addition to  $\xi_k = k\delta$  for  $k \in \mathbb{I}_K^0$ , introduce the intermediate points  $\xi_\kappa = \kappa\delta$  for  $\kappa \in \mathbb{I}_K^{1/2}$ . Accordingly, introduce the intermediate values for the vectors  $\vec{x}_\Delta^n$  and  $\vec{z}_\Delta^n$ :

$$\begin{aligned} x_\kappa &= \frac{1}{2}(x_{\kappa+\frac{1}{2}} + x_{\kappa-\frac{1}{2}}) \quad \text{for } \kappa \in \mathbb{I}_K^{1/2}, \\ z_k &= \frac{1}{2}(z_{k+\frac{1}{2}} + z_{k-\frac{1}{2}}) \quad \text{for } k \in \mathbb{I}_K^+. \end{aligned}$$

Now define

- $\widehat{z}_\Delta^n : [0, M] \rightarrow \mathbb{R}$  as the piecewise affine interpolation of the values  $(z_{\frac{1}{2}}, z_{\frac{3}{2}}, \dots, z_{K-\frac{1}{2}})$  with respect to the equidistant grid  $(\xi_{\frac{1}{2}}, \xi_{\frac{3}{2}}, \dots, \xi_{K-\frac{1}{2}})$ , and
- $\widehat{u}_\Delta^n : [a, b] \rightarrow \mathbb{R}$  as the piecewise affine function with

$$\widehat{u}_\Delta^n \circ X_\Delta^n = \widehat{z}_\Delta^n. \tag{33}$$

Our convention is that  $\widehat{z}_\Delta^n(\xi) = z_{\frac{1}{2}}$  for  $0 \leq \xi \leq \delta/2$  and  $\widehat{z}_\Delta^n(\xi) = z_{K-\frac{1}{2}}$  for  $M - \delta/2 \leq \xi \leq M$ , and accordingly  $\widehat{u}_\Delta^n(x) = z_{\frac{1}{2}}$  for  $x \in [a, x_{\frac{1}{2}}]$  and  $\widehat{u}_\Delta^n(x) = z_{K-\frac{1}{2}}$  for  $x \in [x_{K-\frac{1}{2}}, b]$ . The definitions have been made such that

$$x_k = X_\Delta^n(\xi_k), \quad z_k = \widehat{z}(\xi_k) = \widehat{u}(x_k) \quad \text{for all } k \in \mathbb{I}_K^0 \cup \mathbb{I}_K^{1/2}. \quad (34)$$

Notice that  $\widehat{u}_\Delta^n$  is piecewise affine with respect to the ‘‘double grid’’  $(x_0, x_{\frac{1}{2}}, x_1, \dots, x_{K-\frac{1}{2}}, x_K)$ , but in general not with respect to the subgrid  $(x_0, x_1, \dots, x_K)$ . By direct calculation, we obtain for each  $k \in \mathbb{I}_K^+$  that

$$\begin{aligned} \partial_x \widehat{u} \Big|_{(x_{k-\frac{1}{2}}, x_k)} &= \frac{z_k - z_{k-\frac{1}{2}}}{x_k - x_{k-\frac{1}{2}}} = \frac{z_{k+\frac{1}{2}} - z_{k-\frac{1}{2}}}{x_k - x_{k-1}} = z_{k-\frac{1}{2}} \frac{z_{k+\frac{1}{2}} - z_{k-\frac{1}{2}}}{\delta}, \\ \partial_x \widehat{u} \Big|_{(x_k, x_{k+\frac{1}{2}})} &= \frac{z_{k+\frac{1}{2}} - z_k}{x_{k+\frac{1}{2}} - x_k} = \frac{z_{k+\frac{1}{2}} - z_{k-\frac{1}{2}}}{x_{k+1} - x_k} = z_{k+\frac{1}{2}} \frac{z_{k+\frac{1}{2}} - z_{k-\frac{1}{2}}}{\delta}. \end{aligned} \quad (35)$$

Trivially, we also have that  $\partial_x \widehat{u}$  vanishes identically on the intervals  $(a, x_{\frac{1}{2}})$  and  $(x_{K-\frac{1}{2}}, b)$ .

### 3. A PRIORI ESTIMATES AND COMPACTNESS

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

$$\mathbf{H}_\delta(\bar{x}_\Delta^0) \leq \bar{\mathcal{H}} \quad \text{for all } \Delta, \quad (36)$$

and that

$$(\tau + \delta) \mathbf{F}_\delta(\bar{x}_\Delta^0) \rightarrow 0 \quad \text{as } \Delta \rightarrow 0. \quad (37)$$

Further, we use  $\{q\}_\tau$  to denote the constant in time interpolations of sequences  $(q^n)_{n=0}^\infty$  with step size  $\tau > 0$ , that is

$$\{q\}_\tau(t) := q^n \quad \text{for } t \in ((n-1)\tau, n\tau], \quad \{q\}_\tau(0) := q^0.$$

**3.1. Energy inequality.** The following basic energy estimates are classical for gradient flows.

**Lemma 11.** *One has that  $\mathbf{F}_\delta$  is monotone, i.e.,  $\mathbf{F}_\delta(\bar{x}_\Delta^n) \leq \mathbf{F}_\delta(\bar{x}_\Delta^{n-1})$ , and further:*

$$\mathbf{F}_\delta(\bar{x}_\Delta^n) \leq \mathbf{F}_\delta(\bar{x}_\Delta^0) \quad \text{for all } n \geq 0, \quad (38)$$

$$\|\bar{x}_\Delta^{\bar{n}} - \bar{x}_\Delta^{\underline{n}}\|_\delta^2 \leq 2\mathbf{F}_\delta(\bar{x}_\Delta^0) (\bar{n} - \underline{n})\tau \quad \text{for all } \bar{n} \geq \underline{n} \geq 0, \quad (39)$$

$$\tau \sum_{n=1}^\infty \left\| \frac{\bar{x}_\Delta^n - \bar{x}_\Delta^{n-1}}{\tau} \right\|_\delta^2 = \tau \sum_{n=1}^\infty \|\nabla_\delta \mathbf{F}_\delta(\bar{x}_\Delta^n)\|_\delta^2 \leq 2\mathbf{F}_\delta(\bar{x}_\Delta^0). \quad (40)$$

*Proof.* The monotonicity (38) follows (by induction on  $n$ ) from the definition of  $\bar{x}_\Delta^n$  as minimizer of  $\mathbf{F}_\Delta(\cdot; \bar{x}_\Delta^{n-1})$ :

$$\mathbf{F}_\delta(\bar{x}_\Delta^n) \leq \frac{1}{2\tau} \|\bar{x}_\Delta^n - \bar{x}_\Delta^{n-1}\|_\delta^2 + \mathbf{F}_\delta(\bar{x}_\Delta^n) = \mathbf{F}_\Delta(\bar{x}_\Delta^n; \bar{x}_\Delta^{n-1}) \leq \mathbf{F}_\Delta(\bar{x}_\Delta^{n-1}; \bar{x}_\Delta^{n-1}) = \mathbf{F}_\delta(\bar{x}_\Delta^{n-1}).$$

Moreover, summation of these inequalities from  $n = \underline{n} + 1$  to  $n = \bar{n}$  yields

$$\frac{\tau}{2} \sum_{n=\underline{n}+1}^{\bar{n}} \left[ \frac{\|\bar{x}_\Delta^n - \bar{x}_\Delta^{n-1}\|_\delta}{\tau} \right]^2 \leq \mathbf{F}_\delta(\bar{x}_\Delta^{\bar{n}}) - \mathbf{F}_\delta(\bar{x}_\Delta^{\underline{n}}) \leq \mathbf{F}_\delta(\bar{x}_\Delta^0).$$

For  $\underline{n} = 0$  and  $\bar{n} \rightarrow \infty$ , we obtain the first part of (40). The second part follows by (32). If instead we combine the estimate with Jensen's inequality, we obtain

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

which leads to (39).  $\square$

**3.2. Entropy dissipation.** The key to our convergence analysis is a refined a priori estimate, which follows from the dissipation of the entropy  $\mathbf{H}_\delta$  along the fully discrete solution.

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

$$\tau \sum_{n=1}^{\infty} \delta \sum_{\kappa \in \mathbb{I}_K^{1/2}} z_\kappa^2 \left( \frac{z_{\kappa+1} - 2z_\kappa + z_{\kappa-1}}{\delta^2} \right)^2 \leq \mathbf{H}_\delta(\bar{x}_\Delta^0). \quad (41)$$

*Proof.* By convexity of  $\mathbf{H}_\delta$  and the discrete evolution (32), we have

$$\mathbf{H}_\delta(\bar{x}_\Delta^{n-1}) - \mathbf{H}_\delta(\bar{x}_\Delta^n) \geq \langle \nabla_\delta \mathbf{H}_\delta(\bar{x}_\Delta^n), \bar{x}_\Delta^{n-1} - \bar{x}_\Delta^n \rangle_\delta = \tau \langle \nabla_\delta \mathbf{H}_\delta(\bar{x}_\Delta^n), \nabla_\delta \mathbf{F}_\delta(\bar{x}_\Delta^n) \rangle_\delta$$

for each  $n = 1, 2, \dots$ . Evaluate the (telescopic) sum with respect to  $n$  and use that  $\mathbf{H}_\delta \geq 0$  to obtain

$$\tau \sum_{n=1}^{\infty} \langle \nabla_\delta \mathbf{H}_\delta(\bar{x}_\Delta^n), \nabla_\delta \mathbf{F}_\delta(\bar{x}_\Delta^n) \rangle_\delta \leq \mathbf{H}_\delta(\bar{x}_\Delta^0).$$

It remains to make the scalar product explicit, using (27) and (30):

$$\begin{aligned} \langle \nabla_\delta \mathbf{H}_\delta, \nabla_\delta \mathbf{F}_\delta \rangle_\delta &= \delta \sum_{\kappa \in \mathbb{I}_K^{1/2}, k \in \mathbb{I}_K^+} z_\kappa^2 \left( \frac{z_{\kappa+1} - 2z_\kappa + z_{\kappa-1}}{\delta^2} \right) \left( \frac{z_{k+\frac{1}{2}} - z_{k-\frac{1}{2}}}{\delta} \right) \left( \frac{\mathbf{e}_{\kappa+\frac{1}{2}} - \mathbf{e}_{\kappa-\frac{1}{2}}}{\delta} \right)^T \mathbf{e}_k \\ &= \delta \sum_{\kappa \in \mathbb{I}_K^{1/2}} z_\kappa^2 \left( \frac{z_{\kappa+1} - 2z_\kappa + z_{\kappa-1}}{\delta^2} \right)^2, \end{aligned}$$

using that  $z_{-\frac{1}{2}} = z_{\frac{1}{2}}$  and  $z_{K+\frac{1}{2}} = z_{K-\frac{1}{2}}$ , according to our convention (21).  $\square$

We draw several conclusions from (41). The first is an a priori estimate on the the  $\xi$ -derivative of the affine functions  $\widehat{z}_\Delta^n$ .

**Lemma 13.** *One has that*

$$\tau \sum_{n=1}^{\infty} \|\partial_\xi \widehat{z}_\Delta^n\|_{L^4([0, M])}^4 = \tau \sum_{n=1}^{\infty} \delta \sum_{k \in \mathbb{I}_K^+} \left( \frac{z_{k+\frac{1}{2}}^n - z_{k-\frac{1}{2}}^n}{\delta} \right)^4 \leq 9\overline{\mathcal{H}}. \quad (42)$$

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

*Proof.* Fix  $n \in \mathbb{N}$ . Invoking our convention (21), one obtains

$$\begin{aligned} \|\partial_\xi \widehat{z}_\Delta^n\|_{L^4(\Omega)}^4 &= \sum_{k \in \mathbb{I}_K^0} (z_{k+\frac{1}{2}} - z_{k-\frac{1}{2}}) \left( \frac{z_{k+\frac{1}{2}} - z_{k-\frac{1}{2}}}{\delta} \right)^3 \\ &= - \sum_{k \in \mathbb{I}_K^{1/2}} z_\kappa \left[ \left( \frac{z_{\kappa+1} - z_\kappa}{\delta} \right)^3 - \left( \frac{z_\kappa - z_{\kappa-1}}{\delta} \right)^3 \right] = (A) \end{aligned}$$

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

$$\begin{aligned}
(A) &= -\delta \sum_{\kappa \in \mathbb{I}_K^{1/2}} z_\kappa \frac{z_{\kappa+1} - 2z_\kappa + z_{\kappa-1}}{\delta^2} \times \\
&\quad \times \left[ \left( \frac{z_{\kappa+1} - z_\kappa}{\delta} \right)^2 + \left( \frac{z_\kappa - z_{\kappa-1}}{\delta} \right)^2 + \left( \frac{z_{\kappa+1} - z_\kappa}{\delta} \right) \left( \frac{z_\kappa - z_{\kappa-1}}{\delta} \right) \right] \\
&\leq \frac{3\delta}{2} \sum_{\kappa \in \mathbb{I}_K^{1/2}} \left| z_\kappa \frac{z_{\kappa+1} - 2z_\kappa + z_{\kappa-1}}{\delta^2} \right| \left[ \left( \frac{z_{\kappa+1} - z_\kappa}{\delta} \right)^2 + \left( \frac{z_\kappa - z_{\kappa-1}}{\delta} \right)^2 \right] \\
&\leq \frac{3}{2} \left( \delta \sum_{\kappa \in \mathbb{I}_K^{1/2}} z_\kappa^2 \left[ \frac{z_{\kappa+1} - 2z_\kappa + z_{\kappa-1}}{\delta^2} \right]^2 \right)^{1/2} \left( 4\delta \sum_{k \in \mathbb{I}_K^0} \left( \frac{z_{k+\frac{1}{2}} - z_{k-\frac{1}{2}}}{\delta} \right)^4 \right)^{1/2}.
\end{aligned}$$

Note that the last sum above is again the  $L^4$ -norm of  $\partial_\xi \widehat{z}^n$ . Taking the square on both sides, dividing by the  $L^4$ -norm, summing over  $n = 1, 2, \dots$ , and finally applying the entropy dissipation estimate (41), one arrives at (42).  $\square$

The a priori estimate (42) 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 (42).

**Lemma 15.** *One has*

$$\tau \sum_{n=1}^{\infty} \delta \sum_{k \in \mathbb{I}_K^+} \left[ \left( \frac{z_{k+\frac{1}{2}}^n}{z_{k-\frac{1}{2}}^n} - 1 \right)^4 + \left( \frac{z_{k-\frac{1}{2}}^n}{z_{k+\frac{1}{2}}^n} - 1 \right)^4 \right] \leq 18(b-a)^4 \mathbf{H}_\delta(\bar{\mathbf{x}}_\Delta^0). \quad (43)$$

Moreover, given  $T > 0$ , then for each  $N \in \mathbb{N}$  with  $N\tau \leq T$ , one has

$$\tau \sum_{n=1}^N \delta \sum_{k \in \mathbb{I}_K^+} \left[ \left( \frac{z_{k+\frac{1}{2}}^n}{z_{k-\frac{1}{2}}^n} - 1 \right)^2 + \left( \frac{z_{k-\frac{1}{2}}^n}{z_{k+\frac{1}{2}}^n} - 1 \right)^2 \right] \leq 6(b-a)^2 T^{1/2} \mathbf{H}_\delta(\bar{\mathbf{x}}_\Delta^0)^{1/2} \delta^{1/2}. \quad (44)$$

*Proof.* Recall that  $z_\kappa \geq \delta/(b-a)$  for all  $\kappa$ , see (102). Consider the first term in the inner summation in (43):

$$\delta \sum_{k \in \mathbb{I}_K^+} \left( \frac{z_{k+\frac{1}{2}}^n}{z_{k-\frac{1}{2}}^n} - 1 \right)^4 = \delta \sum_{k \in \mathbb{I}_K^+} \left( \frac{\delta}{z_{k-\frac{1}{2}}^n} \right)^4 \left( \frac{z_{k+\frac{1}{2}}^n - z_{k-\frac{1}{2}}^n}{\delta} \right)^4 \leq (b-a)^4 \|\widehat{z}_\Delta^n\|_{L^4(\Omega)}^4.$$

The same estimate holds for the second term. The claim (43) is now directly deduced from (42) above. The proof of the second claim (44) is similar, using the Cauchy-Schwarz inequality instead of the modulus estimate:

$$\delta \sum_{k \in \mathbb{I}_K^+} \left( \frac{z_{k+\frac{1}{2}}^n}{z_{k-\frac{1}{2}}^n} - 1 \right)^2 = \delta \sum_{k \in \mathbb{I}_K^+} \left( \frac{\delta}{z_{k-\frac{1}{2}}^n} \right)^2 \left( \frac{z_{k+\frac{1}{2}}^n - z_{k-\frac{1}{2}}^n}{\delta} \right)^2 \leq \left( \delta \sum_{k \in \mathbb{I}_K^+} \left( \frac{\delta}{z_{k-\frac{1}{2}}^n} \right)^4 \right)^{1/2} \|\widehat{z}_\Delta^n\|_{L^4(\Omega)}^2.$$

Use estimate (101), sum over  $n = 1, \dots, N$ , and apply the Cauchy-Schwarz inequality to this second summation. This yields

$$\tau \sum_{n=1}^N \delta \sum_{k \in \mathbb{I}_K^+} \left( \frac{z_{k+\frac{1}{2}}^n}{z_{k-\frac{1}{2}}^n} - 1 \right)^2 \leq \delta^{1/2} (b-a)^2 \left( \tau \sum_{n=1}^N 1 \right)^{1/2} \left( \tau \sum_{n=1}^{\infty} \|\widehat{z}_\Delta^n\|_{L^4(\Omega)}^4 \right)^{1/2}.$$

Invoking again (42), and recalling that  $N\tau \leq T$ , we arrive at (44).  $\square$

We are now going to prove the main consequence from the entropy dissipation (38), namely a control on the total variation of  $\sqrt{\widehat{u}_\Delta^n}$ . This estimate is the key ingredient for obtaining strong compactness in Proposition 20. Recall that several equivalent definitions of the *total variation* of  $f \in L^1(\Omega)$  exist. Most generally,

$$\mathrm{TV}[f] = \sup \left\{ \int_{\Omega} f(x) \partial_x \phi(x) \, dx; \phi \in C^{0,1}(\Omega), \sup_x |\phi(x)| \leq 1 \right\}. \quad (45)$$

Since we are dealing with functions  $f : \Omega \rightarrow \mathbb{R}$  that are piecewise smooth on intervals and only have jump discontinuities, the following definition is most appropriate:

$$\mathrm{TV}[f] = \sup \left\{ \sum_{j=1}^{J-1} |f(r_{j+1}) - f(r_j)| : J \in \mathbb{N}, a < r_1 < r_2 < \dots < r_N < b \right\}. \quad (46)$$

Further recall the notation

$$\llbracket f \rrbracket_{\bar{x}} = \lim_{x \downarrow \bar{x}} f(x) - \lim_{x \uparrow \bar{x}} f(x).$$

for the height of the jump in  $f(x)$ 's value at  $x = \bar{x}$ .

**Lemma 16.** *One has*

$$\tau \sum_{n=1}^{\infty} \mathrm{TV} \left[ \partial_x \sqrt{\widehat{u}_\Delta^n} \right]^2 \leq 10(b-a)\overline{\mathcal{H}}. \quad (47)$$

*Proof.* Fix  $n$ . Observe that  $\sqrt{\widehat{u}_\Delta^n}$  is smooth on  $\Omega$  except for the points  $x_{\frac{1}{2}}, x_1, \dots, x_{K-\frac{1}{2}}$ , with derivatives given by

$$\partial_x \sqrt{\widehat{u}_\Delta^n} = \frac{1}{2\sqrt{\widehat{u}_\Delta^n}} \partial_x \widehat{u}_\Delta^n, \quad \partial_{xx} \sqrt{\widehat{u}_\Delta^n} = -\frac{1}{4\sqrt{\widehat{u}_\Delta^n}^3} (\partial_x \widehat{u}_\Delta^n)^2 \leq 0.$$

Therefore,  $\partial_x \sqrt{\widehat{u}_\Delta^n}$  is monotonically decreasing in between the (potential) jump discontinuities at the points  $x_{\frac{1}{2}}, x_1, \dots, x_{K-\frac{1}{2}}$ . Further, recall that

$$\partial_x \sqrt{\widehat{u}_\Delta^n}(x) = 0 \quad \text{for all } x \in (a, a + \delta/2) \text{ and all } x \in (b - \delta/2, b), \quad (48)$$

It follows that the supremum in (46) can be realized (in the limit  $\varepsilon \downarrow 0$ ) for a sequence of just  $J = 2(2K - 1)$  many points  $r_j^\varepsilon$ , chosen as follows:

$$r_{2i-1}^\varepsilon = x_{i/2} - \varepsilon \quad \text{and} \quad r_{2i}^\varepsilon = x_{i/2} + \varepsilon, \quad \text{for } i = 1, \dots, 2K - 1.$$

On the one hand,

$$\lim_{\varepsilon \downarrow 0} \left| \partial_x \sqrt{\widehat{u}_\Delta^n}(r_{2i-1}^\varepsilon) - \partial_x \sqrt{\widehat{u}_\Delta^n}(r_{2i}^\varepsilon) \right| = \left| \llbracket \partial_x \sqrt{\widehat{u}_\Delta^n} \rrbracket_{x_{i/2}} \right|. \quad (49)$$

On the other hand, since  $\partial_x \sqrt{\widehat{u}_\Delta^n}$  is monotone decreasing in between  $r_{2i}^\varepsilon$  and  $r_{2i+1}^\varepsilon$ , and vanishes near the boundary by (48), we have that

$$\begin{aligned} \lim_{\varepsilon \downarrow 0} \sum_{i=1}^{2K-2} \left( \partial_x \sqrt{\widehat{u}_\Delta^n}(r_{2i}^\varepsilon) - \partial_x \sqrt{\widehat{u}_\Delta^n}(r_{2i+1}^\varepsilon) \right) &= \lim_{\varepsilon \downarrow 0} \sum_{i=1}^{2K-1} \left( \partial_x \sqrt{\widehat{u}_\Delta^n}(r_{2i}^\varepsilon) - \partial_x \sqrt{\widehat{u}_\Delta^n}(r_{2i-1}^\varepsilon) \right) \\ &\leq \sum_{k \in \mathbb{I}_K^+} \left| \llbracket \partial_x \sqrt{\widehat{u}_\Delta^n} \rrbracket_{x_k} \right| + \sum_{\kappa \in \mathbb{I}_K^{1/2}} \left| \llbracket \partial_x \sqrt{\widehat{u}_\Delta^n} \rrbracket_{x_\kappa} \right|. \end{aligned} \quad (50)$$

Summarizing (49) and (50), we obtain the estimate

$$\text{TV} \left[ \partial_x \sqrt{\widehat{u}_\Delta^n} \right] \leq 2 \sum_{k \in \mathbb{I}_K^+} \left| \left[ \partial_x \sqrt{\widehat{u}_\Delta^n} \right]_{x_k} \right| + 2 \sum_{\kappa \in \mathbb{I}_K^{1/2}} \left| \left[ \partial_x \sqrt{\widehat{u}_\Delta^n} \right]_{x_\kappa} \right|, \quad (51)$$

In view of (35), we have that

$$\begin{aligned} \left[ \partial_x \sqrt{\widehat{u}_\Delta^n} \right]_{x_k} &= \frac{1}{2\sqrt{z_k}} \frac{(z_{k-\frac{1}{2}} - z_{k+\frac{1}{2}})^2}{\delta} \quad \text{for } k \in \mathbb{I}_K^+, \\ \left[ \partial_x \sqrt{\widehat{u}_\Delta^n} \right]_{x_\kappa} &= \frac{1}{2} \sqrt{z_\kappa} \frac{z_{\kappa+1} - 2z_\kappa + z_{\kappa-1}}{\delta} \quad \text{for } \kappa \in \mathbb{I}_K^{1/2}. \end{aligned}$$

Accordingly, using that  $1/z_k \leq (1/z_{k+\frac{1}{2}} + 1/z_{k-\frac{1}{2}})/2$  by the arithmetic-harmonic mean inequality,

$$\begin{aligned} \sum_{k \in \mathbb{I}_K^+} \left| \left[ \partial_x \sqrt{\widehat{u}_\Delta^n} \right]_{x_k} \right| &= \frac{\delta}{2} \sum_{k \in \mathbb{I}_K^+} \frac{(z_{k-\frac{1}{2}} - z_{k+\frac{1}{2}})^2}{\delta^2} \cdot \frac{1}{\sqrt{z_k}} \\ &\leq \frac{1}{2} \left( \delta \sum_{k \in \mathbb{I}_K^+} \left[ \frac{z_{k-\frac{1}{2}} - z_{k+\frac{1}{2}}}{\delta} \right]^4 \right)^{1/2} \left( \sum_{k \in \mathbb{I}_K^+} \frac{\delta}{z_k} \right)^{1/2} \\ &= \frac{1}{2} \|\partial_\xi \widehat{z}_\Delta^n\|_{L^4(\Omega)}^2 (b-a)^{1/2}, \end{aligned} \quad (52)$$

and also

$$\begin{aligned} \sum_{\kappa \in \mathbb{I}_K^{1/2}} \left| \left[ \partial_x \sqrt{\widehat{u}_\Delta^n} \right]_{x_\kappa} \right| &= \frac{\delta}{2} \sum_{\kappa \in \mathbb{I}_K^{1/2}} z_\kappa \left| \frac{z_{\kappa+1} - 2z_\kappa + z_{\kappa-1}}{\delta^2} \right| \cdot \frac{1}{\sqrt{z_\kappa}} \\ &\leq \frac{1}{2} \left( \delta \sum_{\kappa \in \mathbb{I}_K^{1/2}} z_\kappa^2 \left[ \frac{z_{\kappa+1} - 2z_\kappa + z_{\kappa-1}}{\delta^2} \right]^2 \right)^{1/2} (b-a)^{1/2}. \end{aligned} \quad (53)$$

Combine (52) with the  $L^4$  bound from (42), and (53) with the entropy dissipation inequality (41). Inserting this into (51) to obtain the claim (47).  $\square$

**3.3. Convergence of time interpolants.** Recall that we require the a priori bound (36) on the initial entropy, but only (37) on the initial Fisher information. This estimate improves over time.

**Lemma 17.** *One has, for every  $N \geq 1$ ,*

$$\mathbf{F}_\delta(\bar{x}_\Delta^N) \leq \frac{3}{2} (M\overline{\mathcal{H}})^{1/2} (N\tau)^{-1/2}. \quad (54)$$

Consequently,  $\{\mathbf{F}_\delta(\bar{x}_\Delta)\}_\tau(t)$  is bounded for each  $t > 0$ , uniformly in  $\Delta$ .

*Proof.* Since  $\mathbf{F}_\delta(\vec{x}_\Delta^n)$  is monotonically decreasing in  $n$  (for fixed  $\Delta$ ), it follows that

$$\begin{aligned} \mathbf{F}_\delta(\vec{x}_\Delta^N) &\leq \frac{1}{N} \sum_{n=1}^N \mathbf{F}_\delta(\vec{x}_\Delta^n) = \frac{1}{2N} \sum_{n=1}^N \delta \sum_{k \in \mathbb{I}_K^+} \left( \frac{z_{k+\frac{1}{2}}^n - z_{k-\frac{1}{2}}^n}{\delta} \right)^2 \\ &\leq \frac{1}{2N\tau} \left( \tau \sum_{n=1}^N \delta \sum_{k \in \mathbb{I}_K^+} 1 \right)^{1/2} \left( \tau \sum_{n=1}^{\infty} \delta \sum_{k \in \mathbb{I}_K^+} \left( \frac{z_{k+\frac{1}{2}}^n - z_{k-\frac{1}{2}}^n}{\delta} \right)^4 \right)^{1/2} \\ &\leq \frac{1}{2N\tau} (N\tau M)^{1/2} (9\overline{\mathcal{H}})^{1/2} = \frac{3}{2} (M\overline{\mathcal{H}})^{1/2} (N\tau)^{-1/2}, \end{aligned}$$

as desired.  $\square$

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

**Lemma 18.** *We have that, for each  $[t, \bar{t}] \in \mathbb{R}_+$ ,*

$$\sup_{\Delta} \sup_{t \in [t, \bar{t}]} \|\{\widehat{u}_\Delta\}_\tau(t)\|_{H^1(\Omega)} < \infty, \quad (55)$$

and that, as  $\Delta \rightarrow 0$ ,

$$\sup_{t \in \mathbb{R}_+} \|\{\widehat{u}_\Delta\}_\tau(t) - \{\bar{u}_\Delta\}_\tau(t)\|_{L^\infty(\Omega)} \rightarrow 0. \quad (56)$$

*Proof.* For each  $n \in \mathbb{N}$ ,

$$\begin{aligned} \|\partial_x \widehat{u}_\Delta^n\|_{L^2(\Omega)}^2 &= \sum_{k \in \mathbb{I}_K^+} \left[ (x_{k+\frac{1}{2}}^n - x_k^n) \left( \frac{z_{k+\frac{1}{2}}^n - z_k^n}{x_{k+\frac{1}{2}}^n - x_k^n} \right)^2 + (x_k^n - x_{k-\frac{1}{2}}^n) \left( \frac{z_k^n - z_{k-\frac{1}{2}}^n}{x_k^n - x_{k-\frac{1}{2}}^n} \right)^2 \right] \\ &\leq \delta \sum_{k \in \mathbb{I}_K^+} \frac{z_{k+\frac{1}{2}}^n + z_{k-\frac{1}{2}}^n}{2} \left( \frac{z_{k+\frac{1}{2}}^n - z_{k-\frac{1}{2}}^n}{\delta} \right)^2 \leq \mathbf{F}_\delta(\vec{x}_\Delta^n) \max_{\kappa \in \mathbb{I}_K^{1/2}} z_\kappa^n. \end{aligned}$$

Now combine this with the estimates (54) from above and (103) from the appendix to obtain (55). Estimate (56) follows directly from the elementary observation that

$$\sup_{x \in \Omega} |\bar{u}_\Delta^n(x) - \widehat{u}_\Delta^n(x)|^2 \leq \max_{k \in \mathbb{I}_K^+} |z_{k+\frac{1}{2}}^n - z_{k-\frac{1}{2}}^n|^2 \leq \delta \mathbf{F}_\delta(\vec{x}_\Delta^n) \leq \delta \mathbf{F}_\delta(\vec{x}_\Delta^0),$$

and an application of (37).  $\square$

**Proposition 19.** *There exists a function  $u_* : \mathbb{R}_{\geq 0} \times \Omega \rightarrow \mathbb{R}_{\geq 0}$  with*

$$u_* \in C_{loc}^{1/2}(\mathbb{R}_+; \mathcal{P}(\Omega)) \cap L_{loc}^\infty(\mathbb{R}_+; H^1(\Omega)), \quad (57)$$

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

$$\{u_\Delta\}_\tau(t) \longrightarrow u_*(t) \quad \text{in } \mathcal{P}(\Omega), \text{ uniformly with respect to } t \in [t, \bar{t}], \quad (58)$$

$$\{u_\Delta\}_\tau, \{\widehat{u}_\Delta\}_\tau \longrightarrow u_* \quad \text{uniformly on } [t, \bar{t}] \times \Omega, \quad (59)$$

$$\{X_\Delta\}_\tau(t) \longrightarrow X^*(t) \quad \text{in } L^2([0, M]), \text{ uniformly with respect to } t \in [t, \bar{t}], \quad (60)$$

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

*Proof.* Fix  $\underline{t} > 0$ . From the discrete energy inequality (39), the bound on the Fisher information in Lemma 17, and the equivalence (23) of  $\mathbb{W}_\delta$  with the usual  $L^2$ -Wasserstein metric  $\mathbb{W}$ , it follows by elementary considerations that

$$\mathbb{W}(\{\bar{u}_\Delta\}_\tau(t), \{\bar{u}_\Delta\}_\tau(s))^2 \leq C(\underline{t})(|t - s| + \tau), \quad (61)$$

for all  $t, s \geq \underline{t}$ . Moreover, since  $\Omega = [a, b]$  is compact, also  $\mathcal{P}(\Omega)$  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  $u_{\underline{t}}$  in  $\mathcal{P}(\Omega)$ , locally uniformly with respect to  $t \in [\underline{t}, \infty)$ . The Hölder-type estimate (61) implies  $u_{\underline{t}} \in C^{1/2}([\underline{t}, \infty); \mathcal{P}(\Omega))$ . The claim (60) is a consequence of the equivalence between the Wasserstein metric on  $\mathcal{P}(\Omega)$  and the  $L^2$ -metric on  $\mathfrak{X}$ , see Lemma 4.

Clearly, the previous argument applies to every choice of  $\underline{t} > 0$ . Using a diagonal argument, one constructs a limit  $u_*$  defined on all  $\mathbb{R}_+$ , such that  $u_{\underline{t}}$  is the restriction of  $u_*$  to  $[\underline{t}, \infty)$ .

For the rest of the proof, let some  $[\underline{t}, \bar{t}] \in \mathbb{R}_+$  be fixed.

For proving (59), it suffices to show that  $\{\hat{u}_\Delta\}_\tau \rightarrow u_*$  uniformly on  $[\underline{t}, \bar{t}] \times \Omega$ : indeed, (56) implies that if  $\{\hat{u}_\Delta\}_\tau$  converges uniformly to some limit, so does  $\{\bar{u}_\Delta\}_\tau$ . As an intermediate step towards proving uniform convergence of  $\{\hat{u}_\Delta\}_\tau$ , we show that

$$\hat{u}_\Delta(t) \longrightarrow u_*(t) \quad \text{in } L^2(\Omega), \text{ uniformly in } t \in [\underline{t}, \bar{t}]. \quad (62)$$

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

$$\begin{aligned} \|\{\hat{u}_\Delta\}_\tau(t) - u_*(t)\|_{L^2(\Omega)}^2 &= \int_\Omega [(\{\hat{u}_\Delta\}_\tau - u_*)\{\bar{u}_\Delta\}_\tau](t; x) \, dx \\ &\quad + \int_\Omega [(\{\hat{u}_\Delta\}_\tau - u_*)(\{\hat{u}_\Delta\}_\tau - \{\bar{u}_\Delta\}_\tau)](t; x) \, dx \\ &\quad - \int_\Omega [(\{\hat{u}_\Delta\}_\tau - u_*)u_*](t; x) \, dx. \end{aligned}$$

On the one hand, observe that

$$\begin{aligned} &\sup_{t \in [\underline{t}, \bar{t}]} \int_\Omega [(\{\hat{u}_\Delta\}_\tau - u_*)(\{\hat{u}_\Delta\}_\tau - \{\bar{u}_\Delta\}_\tau)](t; x) \, dx \\ &\leq \sup_{t \in [\underline{t}, \bar{t}]} ((\|\{\hat{u}_\Delta\}_\tau(t)\|_{L^1(\Omega)} + \|u_*(t)\|_{L^1(\Omega)}) \|\{\hat{u}_\Delta\}_\tau(t) - \{\bar{u}_\Delta\}_\tau(t)\|_{L^\infty} \\ &\leq \sup_{t \in [\underline{t}, \bar{t}]} ((2M + (b - a)) \|\{\hat{u}_\Delta\}_\tau(t) - \{\bar{u}_\Delta\}_\tau(t)\|_{L^\infty}) \|\{\hat{u}_\Delta\}_\tau(t) - \{\bar{u}_\Delta\}_\tau(t)\|_{L^\infty}, \end{aligned}$$

which converges to zero as  $\Delta \rightarrow 0$ , using both conclusions from Lemma 18. On the other hand, we can use property (17) to write

$$\begin{aligned} &\int_\Omega [(\{\hat{u}_\Delta\}_\tau - u_*)\{\bar{u}_\Delta\}_\tau](t; x) \, dx - \int_\Omega [(\{\hat{u}_\Delta\}_\tau - u_*)u_*](t; x) \, dx \\ &= \int_0^M [(\{\hat{u}_\Delta\}_\tau - u_*)](t; \{X_\Delta\}_\tau(t; x)) \, d\xi - \int_0^M [(\{\hat{u}_\Delta\}_\tau - u_*)](t; X^*(t; \xi)) \, d\xi. \end{aligned}$$

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

$$\begin{aligned}
& \sup_{t \in [\underline{t}, \bar{t}]} \left| \int_0^M (\{\widehat{u}_\Delta\}_\tau(t; \{X_\Delta\}_\tau(t; \xi)) - \{\widehat{u}_\Delta\}_\tau(t; X_*(t; \xi))) \, d\xi \right| \\
& \leq \sup_{t \in [\underline{t}, \bar{t}]} \int_0^M \int_{X_*(t, \xi)}^{\{X_\Delta\}_\tau(t; \xi)} |\partial_x \{\widehat{u}_\Delta\}_\tau(t; y)| \, dy \, d\xi \\
& \leq \sup_{t \in [\underline{t}, \bar{t}]} \int_0^M \|\{\widehat{u}_\Delta\}_\tau\|_{H^1(\Omega)} |X_* - \{X_\Delta\}_\tau|(t, \xi)^{1/2} \, d\xi \\
& \leq \sup_{t \in [\underline{t}, \bar{t}]} \left( \|\{\widehat{u}_\Delta\}_\tau(t)\|_{H^1(\Omega)} \|X_*(t) - \{X_\Delta\}_\tau(t)\|_{L^2([0, M])}^{1/4} \right).
\end{aligned}$$

A similar reasoning applies to the integral involving  $u_*$  in place of  $\{\widehat{u}_\Delta\}_\tau$ . Together, this proves (62), and it further proves that  $u_* \in L^\infty(\mathbb{R}_+[\underline{t}, \bar{t}]; H^1(\Omega))$ , since the uniform bound on  $\widehat{u}_\Delta$  from 55 is inherited by the limit.

Now the Gagliardo-Nirenberg inequality (107) provides the estimate

$$\|\{\widehat{u}_\Delta\}_\tau(t) - u^*(t)\|_{C^{1/6}(\Omega)} \leq C \|\{\widehat{u}_\Delta\}_\tau(t) - u^*(t)\|_{H^1(\Omega)}^{2/3} \|\{\widehat{u}_\Delta\}_\tau(t) - u^*(t)\|_{L^2(\Omega)}^{1/3}. \quad (63)$$

Combining the convergence in  $L^2(\Omega)$  by (62) with the boundedness in  $H^1(\Omega)$  from (55), it readily follows that  $\widehat{u}_\Delta(t) \rightarrow u_*(t)$  in  $C^{1/6}(\Omega)$ , uniformly in  $t \in [\underline{t}, \bar{t}]$ . This clearly implies that  $\{\widehat{u}_\Delta\}_\tau \rightarrow u_*$  uniformly on  $[\underline{t}, \bar{t}] \times \Omega$ .  $\square$

**Proposition 20.** *Under the hypotheses and with the notations of Proposition 19, we have that  $\sqrt{u_*} \in L^2(\mathbb{R}_{\geq 0}; H^1(\Omega))$ , and*

$$\left\{ \sqrt{\widehat{u}_\Delta} \right\}_\tau \rightarrow \sqrt{u_*} \quad \text{strongly in } L^2_{loc}(\mathbb{R}_+; H^1(\Omega)) \quad (64)$$

as  $\Delta \rightarrow 0$ .

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

*Proof.* Fix  $[\underline{t}, \bar{t}] \in \mathbb{R}_+$ . By definition (45) of the total variation,

$$\|\partial_x f\|_{L^2(\Omega)}^2 = \int_\Omega \partial_x f(x)^2 \, dx \leq \left( \sup_{x \in \Omega} |f(x)| \right) \text{TV}[\partial_x f]$$

holds for every Lipschitz function  $f : \Omega \rightarrow \mathbb{R}$ . The functions  $\sqrt{\widehat{u}_\Delta^n}$  are obviously Lipschitz continuous. Moreover, thanks to weak lower semi-continuity of the total variation, it follows from (47) that

$$\int_0^\infty \text{TV}[\partial_x \sqrt{u_*}]^2 \, dt \leq 10(b-a)\overline{\mathcal{H}}.$$

In particular, the weak derivative  $x \mapsto \partial_x \sqrt{u_*}(t; x)$  is in  $L^\infty(\Omega)$  — and thus  $x \mapsto \sqrt{u_*}(t; x)$  is Lipschitz — for almost every  $t > 0$ . Using that  $\text{TV}[f - g] \leq \text{TV}[f] + \text{TV}[g]$ , we obtain that

$$\begin{aligned}
& \int_{\underline{t}}^{\bar{t}} \left\| \partial_x \left( \left\{ \sqrt{\widehat{u}_\Delta} \right\}_\tau - \sqrt{u_*} \right) \right\|_{L^2(\Omega)}^2 \, dt \\
& \leq (\bar{t} - \underline{t})^{1/2} \left( \sup_{[\underline{t}, \bar{t}] \times \Omega} \left| \left\{ \sqrt{\widehat{u}_\Delta} \right\}_\tau - \sqrt{u_*} \right| \right) \left( 2 \int_0^\infty \left( \text{TV}[\partial_x \{\sqrt{u_\Delta}\}_\tau]^2 + \text{TV}[\partial_x \sqrt{u_*}]^2 \right) \, dt \right)^{1/2}.
\end{aligned}$$

For  $\Delta \rightarrow 0$ , the first term on the right-hand side converges to zero by (59), and the second term remains bounded by (47). This proves (64).

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 yields that

$$4 \int_{\Omega} \left( \partial_x \sqrt{\widehat{u}_{\Delta}^n} \right)^2 (x) dx = \int_0^M \frac{\partial_{\xi} \widehat{z}_{\Delta}^n(\xi)^2}{\widehat{z}_{\Delta}^n(\xi) \partial_{\xi} X_{\Delta}^n(\xi)} d\xi.$$

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

$$\frac{1}{\widehat{z}_{\Delta}^n(\xi) \partial_{\xi} X_{\Delta}^n(\xi)} \leq \frac{z_{k+\frac{1}{2}}^n}{z_{k-\frac{1}{2}}^n} + \frac{z_{k-\frac{1}{2}}^n}{z_{k+\frac{1}{2}}^n}$$

for all  $\xi \in (\xi_{k-\frac{1}{2}}, \xi_{k+\frac{1}{2}})$ . Therefore,

$$\begin{aligned} & 4 \int_0^T \left( \int_{\Omega} \left\{ \sqrt{\widehat{u}_{\Delta}} \right\}_{\tau}^2 (t; x) dx \right) dt \\ & \leq \tau \sum_{n=1}^N \delta \sum_{k \in \mathbb{I}_K^+} \left( \frac{z_{k+\frac{1}{2}}^n - z_{k-\frac{1}{2}}^n}{\delta} \right)^2 \left( \frac{z_{k+\frac{1}{2}}^n}{z_{k-\frac{1}{2}}^n} + \frac{z_{k-\frac{1}{2}}^n}{z_{k+\frac{1}{2}}^n} \right) \\ & \leq 2 \left( \tau \sum_{n=1}^N \delta \sum_{k \in \mathbb{I}_K^+} \left[ \frac{z_{k+\frac{1}{2}}^n - z_{k-\frac{1}{2}}^n}{\delta} \right]^4 \right)^{1/2} \left( \tau \sum_{n=1}^N \delta \sum_{k \in \mathbb{I}_K^+} \left[ \left( \frac{z_{k+\frac{1}{2}}^n}{z_{k-\frac{1}{2}}^n} \right)^2 + \left( \frac{z_{k-\frac{1}{2}}^n}{z_{k+\frac{1}{2}}^n} \right)^2 \right] \right)^{1/2}. \end{aligned}$$

The two sums are  $\Delta$ -uniformly bounded, thanks to the estimates (42) and (44). 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,  $(\bar{x}_{\Delta}^n)_{n=0}^{\infty}$  with its derived functions  $\bar{u}_{\Delta}$ ,  $\widehat{u}_{\Delta}$ ,  $X_{\Delta}$  is a (sub)sequence for which the convergence results stated in Proposition 19 and Proposition 20 holds. We continue to assume (36) and (37). The goal of this section is to prove the following.

**Proposition 21.** *For every  $\rho \in C^{\infty}(\Omega)$  with  $\rho'(a) = \rho'(b) = 0$ , and for every  $\psi \in C_c^{\infty}(\mathbb{R}_{\geq 0})$ ,*

$$\begin{aligned} & \int_0^{\infty} \psi'(t) \left( \int_{\Omega} \rho(x) u_*(t; x) dx \right) dt + \psi(0) \int_{\Omega} \rho(x) u^0(x) dx \\ & + \int_0^{\infty} \psi(t) \left( \int_{\Omega} [\rho'''(x) \partial_x u_*(t; x) + 4\rho''(x) \partial_x \sqrt{u_*(t; x)}^2] dx \right) dt = 0. \end{aligned} \quad (65)$$

For definiteness, fix a spatial test function  $\rho \in C^{\infty}(\Omega)$  with  $\rho'(a) = \rho'(b) = 0$ , and a temporal test function  $\psi \in C_c^{\infty}(\mathbb{R}_{\geq 0})$  with  $\text{supp } \psi \subset [0, T)$  for a suitable  $T > 0$ . Let  $B > 0$  be chosen such that

$$\|\rho\|_{C^4(\Omega)} \leq B, \quad \|\psi\|_{C^1(\mathbb{R}_+)} \leq B. \quad (66)$$

For convenience, we assume  $\delta < 1$  and  $\tau < 1$ . Further, we introduce the short-hand notation

$$\rho'(\bar{x}_{\Delta}^n) = (\rho'(x_1^n), \dots, \rho'(x_{K-1}^n)) \in \mathbb{R}^{K-1}. \quad (67)$$

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

estimates, respectively:

$$\begin{aligned} e_{1,\Delta} := & \left| \int_0^T \left( \psi'(t) \int_{\Omega} \rho(x) \{\bar{u}_{\Delta}\}_{\tau}(t; x) dx + \psi(t) \{ \langle \rho'(\bar{x}_{\Delta}), \nabla_{\delta} \mathbf{F}_{\delta}(\bar{x}_{\Delta}) \rangle_{\delta} \}_{\tau}(t) \right) dt \right. \\ & \left. + \psi(0) \int_{\Omega} \rho(x) \bar{u}_{\Delta}^0(x) dx \right| \leq C((\delta \mathbf{F}_{\delta}(\bar{x}_{\Delta}^0))^{1/2} + (\tau \mathbf{F}_{\delta}(\bar{x}_{\Delta}^0))), \end{aligned} \quad (68)$$

and

$$\begin{aligned} e_{2,\Delta} := & \left| \int_0^T \psi(t) \left( \int_{\Omega} [\rho'''(x) \partial_x \{\widehat{u}_{\Delta}\}_{\tau}(t; x) + 4\rho''(x) \partial_x \{ \sqrt{\widehat{u}_{\Delta}} \}_{\tau}(t; x)^2] dx \right. \right. \\ & \left. \left. - \{ \langle \rho'(\bar{x}_{\Delta}^n), \nabla_{\delta} \mathbf{F}_{\delta}(\bar{x}_{\Delta}^n) \rangle_{\delta} \}_{\tau}(t) \right) dt \right| \leq C\delta^{1/4}. \end{aligned} \quad (69)$$

We proceed by proving (68) and (69). At the end of this section, it is shown how the claim (65) follows from (68)&(69) on basis of the convergence for  $\{\bar{u}_{\Delta}\}_{\tau}$  obtained previously.

*Proof of (68).* 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”:

$$\begin{aligned} - \int_0^T \psi'(t) \left( \int_{\Omega} \rho(x) \{\bar{u}_{\Delta}\}_{\tau}(t; x) dx \right) dt &= - \sum_{m=1}^{N_{\tau}} \left( \int_{(m-1)\tau}^{m\tau} \psi'(t) dt \int_{\Omega} \rho(x) \bar{u}_{\Delta}^m(x) dx \right) \\ &= -\tau \sum_{m=1}^{N_{\tau}} \left( \frac{\psi(m\tau) - \psi((m-1)\tau)}{\tau} \int_0^M \rho \circ X_{\Delta}^m(\xi) d\xi \right) \\ &= \tau \sum_{n=1}^{N_{\tau}} \left( \psi((n-1)\tau) \int_0^M \frac{\rho \circ X_{\Delta}^n - \rho \circ X_{\Delta}^{n-1}}{\tau}(\xi) d\xi \right) + \psi(0) \int_0^M \rho \circ X_{\Delta}^0(\xi) d\xi. \end{aligned} \quad (70)$$

A Taylor expansion of the term in the inner integral yields

$$\frac{\rho \circ X_{\Delta}^n - \rho \circ X_{\Delta}^{n-1}}{\tau} = \rho' \circ X_{\Delta}^n \left( \frac{X_{\Delta}^n - X_{\Delta}^{n-1}}{\tau} \right) + \frac{\tau}{2} \rho'' \circ \tilde{X} \left( \frac{X_{\Delta}^n - X_{\Delta}^{n-1}}{\tau} \right)^2. \quad (71)$$

where  $\tilde{X}$  symbolizes suitable “intermediate values” in  $[0, M]$ . We analyze the first term on the right-hand side of (71): using the representation (18) of  $X_{\Delta}$  in terms of hat functions  $\theta_k$ , we can write its integral as follows,

$$\int_0^M \rho' \circ X_{\Delta}^n \left( \frac{X_{\Delta}^n - X_{\Delta}^{n-1}}{\tau} \right) d\xi = \sum_{k \in \mathbb{I}_K^+} \left( \frac{x_k^n - x_k^{n-1}}{\tau} \right) \int_{\xi_{k-1}}^{\xi_{k+1}} \rho' \circ X_{\Delta}^n \theta_k d\xi. \quad (72)$$

On the other hand, since

$$\int_{\xi_{k-1}}^{\xi_{k+1}} \theta_k d\xi = \delta, \quad (73)$$

the discrete evolution equation (32) yields that

$$- \langle \rho'(\bar{x}_{\Delta}^n), \nabla_{\delta} \mathbf{F}_{\delta}(\bar{x}_{\Delta}^n) \rangle_{\delta} = \left\langle \rho'(\bar{x}_{\Delta}^n), \frac{\bar{x}_{\Delta}^n - \bar{x}_{\Delta}^{n-1}}{\tau} \right\rangle_{\delta} = \sum_{k \in \mathbb{I}_K^+} \left( \frac{x_k^n - x_k^{n-1}}{\tau} \right) \int_{\xi_{k-1}}^{\xi_{k+1}} \rho(x_k^n) \theta_k(\xi) d\xi. \quad (74)$$

Finally, observing that

$$|X_\Delta^n(\xi) - x_k^n| \leq (x_{k+1}^n - x_{k-1}^n) \quad \text{for each } \xi \in (\xi_{k-1}, \xi_{k+1}),$$

we can estimate the difference of the terms in (72) and (74) with the help of the bound (66) on  $\rho$  as follows:

$$\begin{aligned} & \left| \int_0^M \rho' \circ X_\Delta^n(\xi) \left( \frac{X_\Delta^n - X_\Delta^{n-1}}{\tau} \right) (\xi) d\xi - \langle \rho'(\bar{x}_\Delta^n), \nabla_\delta \mathbf{F}_\delta(\bar{x}_\Delta^n) \rangle_\delta \right| \\ & \leq \sum_{k \in \mathbb{I}_K^+} \left| \frac{x_k^n - x_k^{n-1}}{\tau} \right| \int_{\xi_{k-1}}^{\xi_{k+1}} |\rho' \circ X_\Delta^n(\xi) - \rho'(x_k^n)| \theta_k(\xi) d\xi \\ & \leq B\delta \sum_{k \in \mathbb{I}_K^+} \left| \frac{x_k^n - x_k^{n-1}}{\tau} \right| (x_{k+1}^n - x_{k-1}^n). \end{aligned} \quad (75)$$

As a final preparation for the proof of (68), observe that

$$\begin{aligned} R' & := \left| \int_0^T \psi(t) \{ \langle \rho'(\bar{x}_\Delta), \nabla_\delta \mathbf{F}_\delta(\bar{x}_\Delta) \rangle_\delta \}_\tau(t) dt - \tau \sum_{n=1}^{N_\tau} \psi((n-1)\tau) \langle \rho'(\bar{x}_\Delta^n), \nabla_\delta \mathbf{F}_\delta(\bar{x}_\Delta^n) \rangle_\delta \right| \\ & \leq \left( \tau \sum_{n=1}^{N_\tau} \left| \frac{1}{\tau} \int_{(n-1)\tau}^{n\tau} \psi(t) dt - \psi((n-1)\tau) \right|^2 \right)^{1/2} \left( \tau \sum_{n=1}^{\infty} B^2 \|\nabla_\delta \mathbf{F}_\delta(\bar{x}_\Delta^n)\|_\delta^2 \right)^{1/2} \\ & \leq ((T+1)B^2\tau^2)^{1/2} (2B^2\mathbf{F}_\delta(\bar{x}_\Delta^0))^{1/2} = C'\mathbf{F}_\delta(\bar{x}_\Delta^0)^{1/2}\tau, \end{aligned}$$

using the energy estimate (40). We are now ready to estimate  $e_{1,\Delta}$  in (68):

$$\begin{aligned} e_{1,\Delta} & \stackrel{(70)}{\leq} R' + \tau \sum_{n=1}^{N_\tau} \left( |\psi((n-1)\tau)| \left| \int_0^M \frac{\rho \circ X_\Delta^n - \rho \circ X_\Delta^{n-1}}{\tau} (\xi) d\xi - \langle \rho'(\bar{x}_\Delta^n), \nabla_\delta \mathbf{F}_\delta(\bar{x}_\Delta^n) \rangle_\delta \right| \right) \\ & \stackrel{(71)}{\leq} R' + B\tau \sum_{n=1}^{N_\tau} \left( \left| \int_0^M \rho' \circ X_\Delta^n(\xi) \left( \frac{X_\Delta^n - X_\Delta^{n-1}}{\tau} \right) (\xi) d\xi - \langle \rho'(\bar{x}_\Delta^n), \nabla_\delta \mathbf{F}_\delta(\bar{x}_\Delta^n) \rangle_\delta \right| \right. \\ & \quad \left. + \frac{B\tau}{2} \int_0^M \left( \frac{X_\Delta^n - X_\Delta^{n-1}}{\tau} \right)^2 (\xi) d\xi \right) \\ & \stackrel{(75)}{\leq} R' + B^2 \left( \tau \sum_{n=1}^{\infty} \delta \sum_{k \in \mathbb{I}_K^+} \left( \frac{x_k^n - x_k^{n-1}}{\tau} \right)^2 \right)^{1/2} \left( \tau \sum_{n=1}^{N_\tau} \delta \sum_{k \in \mathbb{I}_K^+} (x_{k+1}^n - x_{k-1}^n)^2 \right)^{1/2} \\ & \quad + \frac{B^2\tau}{2} \sum_{n=1}^{\infty} \left\| \frac{X_\Delta^n - X_\Delta^{n-1}}{\tau} \right\|_{L^2([0,M])}^2 \\ & \leq C'(\tau\mathbf{F}_\delta(\bar{x}_\Delta^0)) + B^2(2(b-a)^2T)^{1/2}(\delta\mathbf{F}_\delta(\bar{x}_\Delta^0))^{1/2} + B^2(\tau\mathbf{F}_\delta(\bar{x}_\Delta^0)), \end{aligned}$$

where we have used the energy estimate (40) and the the bound (101).  $\square$

The proof of (69) 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 (67), one has that*

$$-\langle \rho'(\bar{x}_\Delta^n), \nabla_\delta \mathbf{F}_\delta(\bar{x}_\Delta^n) \rangle_\delta = A_1^n - A_2^n + A_3^n + A_4^n, \quad (76)$$

where

$$\begin{aligned}
A_1^n &= \delta \sum_{k \in \mathbb{I}_K^+} \left( \frac{z_{k+\frac{1}{2}}^n - z_{k-\frac{1}{2}}^n}{\delta} \right)^2 \left( \frac{z_{k+\frac{1}{2}}^n + z_{k-\frac{1}{2}}^n}{2} \right) \left( \frac{\rho'(x_{k+1}^n) - \rho'(x_{k-1}^n)}{\delta} \right), \\
A_2^n &= \delta \sum_{k \in \mathbb{I}_K^+} \left( \frac{z_{k+\frac{1}{2}}^n - z_{k-\frac{1}{2}}^n}{\delta} \right)^2 \left( \frac{(z_{k+\frac{1}{2}}^n)^2 + (z_{k-\frac{1}{2}}^n)^2}{2z_{k+\frac{1}{2}}^n z_{k-\frac{1}{2}}^n} \right) \rho''(x_k^n), \\
A_3^n &= \delta \sum_{k \in \mathbb{I}_K^+} \left( \frac{z_{k+\frac{1}{2}}^n - z_{k-\frac{1}{2}}^n}{\delta} \right) \left( \frac{(z_{k+\frac{1}{2}}^n)^2 + (z_{k-\frac{1}{2}}^n)^2}{2} \right) \left( \frac{\rho'(x_{k+1}^n) - \rho'(x_k^n) - (x_{k+1}^n - x_k^n) \rho''(x_k^n)}{\delta^2} \right), \\
A_4^n &= \delta \sum_{k \in \mathbb{I}_K^+} \left( \frac{z_{k+\frac{1}{2}}^n - z_{k-\frac{1}{2}}^n}{\delta} \right) \left( \frac{(z_{k+\frac{1}{2}}^n)^2 + (z_{k-\frac{1}{2}}^n)^2}{2} \right) \left( \frac{\rho'(x_{k-1}^n) - \rho'(x_k^n) - (x_{k-1}^n - x_k^n) \rho''(x_k^n)}{\delta^2} \right).
\end{aligned}$$

*Proof.* Fix some time index  $n \in \mathbb{N}$  (omitted in the calculations below). Recall the representation of  $\nabla_\delta \mathbf{F}_\delta$  from (30). By definition of  $\rho'(\vec{x}_\Delta)$ , it follows via a ‘‘summation by parts’’ that

$$\begin{aligned}
-\langle \nabla_{\vec{x}} \mathbf{F}_\delta(\vec{x}_\Delta), \rho'(\vec{x}_\Delta) \rangle_\delta &= -\delta \sum_{\kappa \in \mathbb{I}_K^{1/2}} \frac{1}{\delta} \left( \frac{z_{\kappa+1} - z_\kappa}{\delta} - \frac{z_\kappa - z_{\kappa-1}}{\delta} \right) z_\kappa^2 \left( \frac{\rho'(x_{\kappa+\frac{1}{2}}) - \rho'(x_{\kappa-\frac{1}{2}})}{\delta} \right) \\
&= \delta \sum_{k \in \mathbb{I}_K^+} \left( \frac{z_{k+\frac{1}{2}} - z_{k-\frac{1}{2}}}{\delta} \right) \frac{1}{\delta} \left( z_{k+\frac{1}{2}} \frac{\rho'(x_{k+1}) - \rho'(x_k)}{\delta} - z_{k-\frac{1}{2}} \frac{\rho'(x_k) - \rho'(x_{k-1})}{\delta} \right).
\end{aligned}$$

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

$$\alpha_+ \beta_+ - \alpha_- \beta_- = \frac{\alpha_+ + \alpha_-}{2} (\beta_+ - \beta_-) + (\alpha_+ - \alpha_-) \frac{\beta_+ + \beta_-}{2},$$

we obtain further:

$$\begin{aligned}
&-\langle \nabla_{\vec{x}} \mathbf{F}_\delta(\vec{x}_\Delta), \rho'(\vec{x}_\Delta) \rangle_\delta \\
&= \delta \sum_{k \in \mathbb{I}_K^+} \left( \frac{z_{k+\frac{1}{2}} - z_{k-\frac{1}{2}}}{\delta} \right) \left( \frac{z_{k+\frac{1}{2}}^2 - z_{k-\frac{1}{2}}^2}{2\delta} \right) \left( \frac{\rho'(x_{k+1}) - \rho'(x_{k-1})}{\delta} \right) \tag{77}
\end{aligned}$$

$$\begin{aligned}
&+ \delta \sum_{k \in \mathbb{I}_K^+} \left( \frac{z_{k+\frac{1}{2}} - z_{k-\frac{1}{2}}}{\delta} \right) \left( \frac{z_{k+\frac{1}{2}}^2 + z_{k-\frac{1}{2}}^2}{2} \right) \left( \frac{\rho'(x_{k+1}) - 2\rho'(x_k) + \rho'(x_{k-1}))}{\delta^2} \right). \tag{78}
\end{aligned}$$

The sum in (77) equals to  $A_1^n$ . In order to see that the sum in (78) equals to  $-A_2^n + A_3^n + A_4^n$ , simply observe that the identity

$$\frac{x_{k+1} - x_k}{\delta} + \frac{x_{k-1} - x_k}{\delta} = \frac{1}{z_{k+\frac{1}{2}}} - \frac{1}{z_{k-\frac{1}{2}}} = -\frac{z_{k+\frac{1}{2}} - z_{k-\frac{1}{2}}}{z_{k+\frac{1}{2}} z_{k-\frac{1}{2}}},$$

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

**Lemma 23.** *There is a constant  $C_1 > 0$  such that for each  $N$  with  $N\tau < T$ , one has*

$$R_1 := \tau \sum_{n=1}^N \left| A_1^n - 2 \int_0^M \partial_\xi \widehat{z}_\Delta^n(\xi)^2 \rho'' \circ X_\Delta^n(\xi) \, d\xi \right| \leq C_1 \delta^{1/4}.$$

*Proof.* First, observe that by definition of  $\widehat{z}$ ,

$$\int_0^M \partial_\xi \widehat{z}_\Delta^n(\xi)^2 \rho'' \circ X_\Delta^n(\xi) \, d\xi = \sum_{k \in \mathbb{I}_K^+} \left( \frac{z_{k+\frac{1}{2}}^n - z_{k-\frac{1}{2}}^n}{\delta} \right)^2 \int_{\xi_{k-\frac{1}{2}}}^{\xi_{k+\frac{1}{2}}} \rho'' \circ X_\Delta^n(\xi) \, d\xi,$$

and therefore, by Hölder's inequality,

$$R_1 \leq R_{1a}^{1/2} R_{1b}^{1/2}, \quad (79)$$

with, recalling (42),

$$R_{1a} = \tau \sum_{n=1}^N \delta \sum_{k \in \mathbb{I}_K^+} \left( \frac{z_{k+\frac{1}{2}}^n - z_{k-\frac{1}{2}}^n}{\delta} \right)^4 \leq \tau \sum_{n=1}^\infty \|\widehat{z}_\Delta^n\|_{L^4(\Omega)}^4 \leq 9\overline{\mathcal{H}}, \quad (80)$$

$$R_{1b} = \tau \sum_{n=1}^N \delta \sum_{k \in \mathbb{I}_K^+} \left[ \frac{z_{k+\frac{1}{2}}^n + z_{k-\frac{1}{2}}^n}{2} \frac{\rho'(x_{k+1}^n) - \rho'(x_{k-1}^n)}{\delta} - \frac{2}{\delta} \int_{\xi_{k-\frac{1}{2}}}^{\xi_{k+\frac{1}{2}}} \rho'' \circ X_\Delta^n \, d\xi \right]^2. \quad (81)$$

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

$$\begin{aligned} \frac{\rho'(x_{k+1}) - \rho'(x_{k-1})}{\delta} &= \frac{\rho'(x_{k+1}) - \rho'(x_k)}{\delta} + \frac{\rho'(x_k) - \rho'(x_{k-1})}{\delta} \\ &= \rho''(\tilde{x}_k^+) \frac{x_{k+1} - x_k}{\delta} + \rho''(\tilde{x}_k^-) \frac{x_{k+1} - x_k}{\delta} = \frac{\rho''(\tilde{x}_k^+)}{z_{k+\frac{1}{2}}} + \frac{\rho''(\tilde{x}_k^-)}{z_{k-\frac{1}{2}}}. \end{aligned}$$

For each  $k \in \mathbb{I}_K^+$ , we have that — recalling (73) —

$$\begin{aligned} &\frac{z_{k+\frac{1}{2}} + z_{k-\frac{1}{2}}}{2} \left( \frac{\rho''(\tilde{x}_k^+)}{z_{k+\frac{1}{2}}} + \frac{\rho''(\tilde{x}_k^-)}{z_{k-\frac{1}{2}}} \right) - \frac{2}{\delta} \int_{\xi_{k-\frac{1}{2}}}^{\xi_{k+\frac{1}{2}}} \rho'' \circ X_\Delta \, d\xi \\ &= \frac{1}{2} \left[ \left( \frac{z_{k-\frac{1}{2}}}{z_{k+\frac{1}{2}}} + 1 \right) \rho''(\tilde{x}_k^+) + \left( \frac{z_{k+\frac{1}{2}}}{z_{k-\frac{1}{2}}} + 1 \right) \rho''(\tilde{x}_k^-) \right] - \frac{2}{\delta} \int_{\xi_{k-\frac{1}{2}}}^{\xi_{k+\frac{1}{2}}} \rho'' \circ X_\Delta \, d\xi \\ &= \frac{1}{2} \left[ \left( \frac{z_{k-\frac{1}{2}}}{z_{k+\frac{1}{2}}} - 1 \right) \rho''(\tilde{x}_k^+) + \left( \frac{z_{k+\frac{1}{2}}}{z_{k-\frac{1}{2}}} - 1 \right) \rho''(\tilde{x}_k^-) \right] \\ &\quad - \frac{2}{\delta} \int_{\xi_k}^{\xi_{k+\frac{1}{2}}} [\rho'' \circ X_\Delta - \rho''(\tilde{x}_k^+)] \, d\xi - \frac{2}{\delta} \int_{\xi_{k-\frac{1}{2}}}^{\xi_k} [\rho'' \circ X_\Delta - \rho''(\tilde{x}_k^-)] \, d\xi. \end{aligned}$$

Since  $X_\Delta(\xi) \in [x_k, x_{k+\frac{1}{2}}]$  for each  $\xi \in [\xi_k, \xi_{k+\frac{1}{2}}]$ , and  $\tilde{x}_k^+ \in [x_k, x_{k+1}]$ , it follows that  $|X_\Delta(\xi) - \tilde{x}_k^+| \leq x_{k+1} - x_k$ , and therefore

$$\frac{2}{\delta} \int_{\xi_k}^{\xi_{k+\frac{1}{2}}} |\rho'' \circ X_\Delta(\xi) - \rho''(\tilde{x}_k^+)| \, d\xi \leq B(x_{k+1} - x_k). \quad (82)$$

A similar estimate holds for the other integral. Thus

$$R_{1b} \leq B^2 \tau \sum_{n=1}^N \delta \sum_{k \in \mathbb{I}_K^+} \left[ \left( \frac{z_{k-\frac{1}{2}}^n}{z_{k+\frac{1}{2}}^n} - 1 \right)^2 + \left( \frac{z_{k+\frac{1}{2}}^n}{z_{k-\frac{1}{2}}^n} - 1 \right)^2 + 2(x_{k+1}^n - x_{k-1}^n)^2 \right].$$

Recalling the estimates (44) and (101), we further conclude that

$$R_{1b} \leq B^2 (6(b-a)^2 (\overline{\mathcal{H}} T \delta)^{1/2} + 4T(b-a)^2 \delta). \quad (83)$$

In combination with (79) and (80), this proves the claim.  $\square$

**Lemma 24.** *There is a constant  $C_2 > 0$  such that for each  $N$  with  $N\tau < T$ , one has*

$$R_2 := \tau \sum_{n=1}^N \left| A_2^n - \int_0^M \partial_\xi \widehat{z}_\Delta^n(\xi)^2 \rho'' \circ X_\Delta^n(\xi) \, d\xi \right| \leq C_2 \delta^{1/4}.$$

*Proof.* The proof is almost identical to (and even easier than) the one for Lemma 23 above. Again, we have a decomposition of the form

$$R_2 \leq R_{2a}^{1/2} R_{2b}^{1/2},$$

where  $R_{2a}$  equals  $R_{1a}$  from (80), and

$$R_{2b} = \tau \sum_{n=1}^N \delta \sum_{k \in \mathbb{I}_K^+} \left[ \frac{z_{k+\frac{1}{2}}^n + z_{k-\frac{1}{2}}^n}{2z_{k+\frac{1}{2}}^n z_{k-\frac{1}{2}}^n} \rho''(x_k^n) - \frac{1}{\delta} \int_{\xi_{k-\frac{1}{2}}}^{\xi_{k+\frac{1}{2}}} \rho'' \circ X_\Delta^n \, d\xi \right]^2.$$

By writing

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

and observing — in analogy to (82) — that

$$\frac{1}{\delta} \int_{\xi_{k-\frac{1}{2}}}^{\xi_{k+\frac{1}{2}}} |\rho'' \circ X_\Delta^n(\xi) - \rho''(x_k)| \, d\xi \leq B(x_{k+\frac{1}{2}} - x_{k-\frac{1}{2}}),$$

we obtain the same bound on  $R_{2b}$  as the one on  $R_{1b}$  from (83).  $\square$

**Lemma 25.** *There is a constant  $C_3 > 0$  such that for each  $N$  with  $N\tau \leq T$ , one has*

$$R_3 := \tau \sum_{n=1}^N \left| A_3^n - \frac{1}{2} \int_0^M \partial_\xi \widehat{z}_\Delta^n(\xi) \rho''' \circ X_\Delta^n(\xi) \, d\xi \right| \leq C_3 \delta^{1/4}.$$

*Proof.* Arguing like in the previous proofs, we first deduce — now by means of Hölder's inequality instead of the Cauchy-Schwarz inequality — that

$$R_3 \leq R_{3a}^{1/4} R_{3b}^{3/4},$$

where  $R_{3a} = R_{1a}$ , and

$$R_{3b} = \tau \sum_{n=1}^N \delta \sum_{k \in \mathbb{I}_K^+} \left| \left( \frac{(z_{k+\frac{1}{2}}^n)^2 + (z_{k-\frac{1}{2}}^n)^2}{2} \right) \left( \frac{\rho'(x_{k+1}^n) - \rho'(x_k^n) - (x_{k+1}^n - x_k^n) \rho''(x_k^n)}{\delta^2} \right) - \frac{1}{2\delta} \int_{\xi_{k-\frac{1}{2}}}^{\xi_{k+\frac{1}{2}}} \rho''' \circ X_\Delta^n \, d\xi \right|^{4/3}.$$

Introduce intermediate values  $\tilde{x}_k^+$  such that

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

Thus we have that

$$\begin{aligned}
& \left( \frac{(z_{k+\frac{1}{2}}^n)^2 + (z_{k-\frac{1}{2}}^n)^2}{2} \right) \left( \frac{\rho'(x_{k+1}^n) - \rho'(x_k^n) - (x_{k+1}^n - x_k^n)\rho''(x_k^n)}{\delta^2} \right) - \frac{1}{2\delta} \int_{\xi_{k-\frac{1}{2}}}^{\xi_{k+\frac{1}{2}}} \rho''' \circ X_\delta^n \, d\xi \\
&= \frac{1}{4} \left( \left( \frac{z_{k-\frac{1}{2}}^n}{z_{k+\frac{1}{2}}^n} \right)^2 + 1 \right) \rho'''(\tilde{x}_k^+) - \frac{1}{2\delta} \int_{\xi_{k-\frac{1}{2}}}^{\xi_{k+\frac{1}{2}}} \rho''' \circ X_\delta^n \, d\xi \\
&= \frac{1}{4} \left( \frac{z_{k-\frac{1}{2}}^n}{z_{k+\frac{1}{2}}^n} + 1 \right) \left( \frac{z_{k-\frac{1}{2}}^n}{z_{k+\frac{1}{2}}^n} - 1 \right) \rho'''(\tilde{x}_k^+) - \frac{1}{2\delta} \int_{\xi_{k-\frac{1}{2}}}^{\xi_{k+\frac{1}{2}}} [\rho''' \circ X_\Delta^n - \rho'''(\tilde{x}_k^+)] \, d\xi.
\end{aligned}$$

By the analogue of (82), it follows further that

$$\begin{aligned}
R_{3b} &\leq 2B^{4/3} \tau \sum_{n=1}^N \delta \sum_{k \in \mathbb{I}_K^+} \left[ \left( \frac{z_{k-\frac{1}{2}}^n}{z_{k+\frac{1}{2}}^n} + 1 \right)^{4/3} \left( \frac{z_{k-\frac{1}{2}}^n}{z_{k+\frac{1}{2}}^n} - 1 \right)^{4/3} + (x_{k+1}^n - x_{k-1}^n)^{4/3} \right] \\
&\leq 2B^{4/3} \left( \tau \sum_{n=1}^N \delta \sum_{k \in \mathbb{I}_K^+} \left( \frac{z_{k-\frac{1}{2}}^n}{z_{k+\frac{1}{2}}^n} + 1 \right)^4 \right)^{1/3} \left( \tau \sum_{n=1}^N \delta \sum_{k \in \mathbb{I}_K^+} \left( \frac{z_{k-\frac{1}{2}}^n}{z_{k+\frac{1}{2}}^n} - 1 \right)^2 \right)^{2/3} \\
&\quad + 2B^{4/3} T(b-a)^{4/3} \delta,
\end{aligned}$$

where we have used (101). At this point, the estimates (43) and (44) are used to control the first and the second sum, respectively.  $\square$

Along the same lines, one proves the analogous estimate for  $A_4$  in place of  $A_3$ .

It remains to identify the integral expressions inside  $R_1$  to  $R_3$  with those in the weak formulation (65).

**Lemma 26.** *One has that*

$$\int_0^M \partial_\xi \widehat{z}_\Delta^n(\xi) \rho''' \circ X_\Delta^n(\xi) \, d\xi = \int_\Omega \partial_x \widehat{u}_\Delta^n(x) \rho'''(x) \, dx, \quad (84)$$

$$R_5 := \tau \sum_{n=1}^N \left| \int_0^M \partial_\xi \widehat{z}_\Delta^n(\xi)^2 \rho'' \circ X_\Delta^n(\xi) \, d\xi - 4 \int_\Omega \left( \partial_x \sqrt{\widehat{u}_\Delta^n} \right)^2(x) \rho''(x) \, dx \right| \leq C_5 \delta^{1/4}, \quad (85)$$

where (85) holds for each  $N$  with  $N\tau \leq T$ .

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

$$\widehat{z}_\Delta^n(\xi) = \widehat{u}_\Delta^n \circ X_\Delta^n(\xi) \quad (86)$$

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

$$\partial_\xi \widehat{z}_\Delta^n(\xi) = \partial_x \widehat{u}_\Delta^n \circ X_\Delta^n(\xi) \partial_\xi X_\Delta^n(\xi).$$

Substitute this expression for  $\partial_\xi \widehat{z}_\Delta^n(\xi)$  into the left-hand side of (84), and perform a change of variables  $x = X_\Delta^n(\xi)$  to obtain the integral on the right.

Next, take the square root in (86) before differentiation, then calculate the square and divide by  $\partial_\xi X_\Delta^n(\xi)$  afterwards:

$$\frac{\partial_\xi \widehat{z}_\Delta^n(\xi)^2}{4\widehat{z}_\Delta^n(\xi) \partial_\xi X_\Delta^n(\xi)} = (\partial_x \sqrt{\widehat{u}_\Delta^n})^2 \circ X_\Delta^n(\xi) \partial_\xi X_\Delta^n(\xi).$$

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

$$\int_0^M \frac{\partial_\xi \widehat{z}_\Delta^n(\xi)^2}{\widehat{z}_\Delta^n(\xi) \partial_\xi X_\Delta^n(\xi)} \rho'' \circ X_\Delta^n(\xi) \, d\xi = 4 \int_\Omega \left( \partial_x \sqrt{\widehat{u}_\Delta^n} \right)^2(x) \rho''(x) \, dx. \quad (87)$$

It remains to estimate the difference between the  $\xi$ -integrals in (85) and in (87), respectively. To this end, observe that for each  $\xi \in (\xi_k, \xi_{k+\frac{1}{2}})$  with some  $k \in \mathbb{I}_K^+$ , one has  $\partial_\xi X_\Delta^n(\xi) = 1/z_{k+\frac{1}{2}}^n$  and  $\widehat{z}_\Delta^n(\xi) \in [z_{k-\frac{1}{2}}, z_{k+\frac{1}{2}}]$ . Hence, for those  $\xi$ ,

$$\left| 1 - \frac{1}{\widehat{z}_\Delta^n(\xi) \partial_\xi X_\Delta^n(\xi)} \right| \leq \left| 1 - \frac{z_{k+\frac{1}{2}}^n}{z_{k-\frac{1}{2}}^n} \right|.$$

If instead  $\xi \in (\xi_{k-\frac{1}{2}}, \xi_k)$ , then this estimate holds with the roles of  $z_{k+\frac{1}{2}}^n$  and  $z_{k-\frac{1}{2}}^n$  interchanged. Consequently, using once again (42) and (44),

$$\begin{aligned} & \tau \sum_{n=1}^N \left| \int_0^M \partial_\xi \widehat{z}_\Delta^n(\xi)^2 \rho'' \circ X_\Delta^n(\xi) \, d\xi - \int_0^M \frac{\partial_\xi \widehat{z}_\Delta^n(\xi)^2}{\widehat{z}_\Delta^n(\xi) \partial_\xi X_\Delta^n(\xi)} \rho'' \circ X_\Delta^n(\xi) \, d\xi \right| \\ & \leq B\tau \sum_{n=1}^N \int_0^M \partial_\xi \widehat{z}_\Delta^n(\xi)^2 \left| 1 - \frac{1}{\widehat{z}_\Delta^n(\xi) \partial_\xi X_\Delta^n(\xi)} \right| \, d\xi \\ & \leq B \left( \tau \sum_{n=1}^{\infty} \|\partial_\xi \widehat{z}_\Delta^n\|_{L^4}^4 \right)^{1/2} \left( \tau \sum_{n=1}^N \delta \sum_{k \in \mathbb{I}_K^+} \left[ \left( 1 - \frac{z_{k+\frac{1}{2}}^n}{z_{k-\frac{1}{2}}^n} \right)^2 + \left( 1 - \frac{z_{k-\frac{1}{2}}^n}{z_{k+\frac{1}{2}}^n} \right)^2 \right] \right)^{1/2} \\ & \leq 3\overline{\mathcal{H}}^{1/2} (6(b-a)^2 T^{1/2} \overline{\mathcal{H}}^{1/2} \delta^{1/2})^{1/2}, \end{aligned}$$

since  $N\tau \leq T$  by hypothesis. This shows (85).  $\square$

*Proof of (69).* Again, let  $N_\tau \in \mathbb{N}$  be such that  $N_\tau \tau \in (T, T+1)$ . Combining the discrete weak formulation (76), the change of variables formulae (84)&(85), and the definitions of  $R_1$  to  $R_5$ , it follows that

$$\begin{aligned} e_{2,\Delta} & \leq BR_5 + B\tau \sum_{n=1}^{N_\tau} \left| \int_0^M [\partial_\xi \widehat{z}_\Delta^n \rho''' \circ X_\Delta^n(\xi) + \partial_\xi \widehat{z}_\Delta^n(\xi)^2 \rho'' \circ X_\Delta^n(\xi)] \, d\xi - (A_1^n - A_2^n + A_3^n + A_4^n) \right| \\ & \leq B(R_1 + R_2 + R_3 + R_4 + R_5) \leq B(C_1 + C_2 + C_3 + C_4 + C_5) \delta^{1/4}. \end{aligned}$$

This implies the desired inequality (69).  $\square$

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

*Proof of Proposition 21.* Thanks to (68)&(69), we know that

$$\begin{aligned} & \left| \int_0^T \psi'(t) \int_\Omega \rho(x) \{\bar{u}_\Delta\}_\tau(t; x) \, dx \, dt + \psi(0) \int_\Omega \rho(x) \bar{u}_\Delta^0(x) \, dx \right. \\ & \quad \left. + \int_0^T \psi(t) \int_\Omega [\rho'''(x) \partial_x \{\widehat{u}_\Delta\}_\tau(t; x) + 4\rho''(x) \partial_x \{\sqrt{\widehat{u}_\Delta}\}_\tau(t; x)^2] \, dx \, dt \right| \\ & \leq e_{1,\Delta} + e_{2,\Delta} \leq C((\tau \mathbf{F}_\delta(\bar{x}_\Delta^0)) + (\delta \mathbf{F}_\delta(\bar{x}_\Delta^0))^{1/2} + \delta^{1/4}). \end{aligned}$$

By our assumption (37) on  $\mathbf{F}_\delta(\bar{x}_\Delta^0)$ , the expression on the right hand side vanishes as  $\Delta \rightarrow 0$ . To obtain (65) in the limit  $\Delta \rightarrow 0$ , we still need to show the convergence of the integrals to their respective limits.

A technical tool is the observation that, for each  $p \in [1, 4]$ ,

$$Q_p := \sup_{\Delta} \tau \sum_{n=1}^{N_\tau} \delta \sum_{\kappa \in \mathbb{I}_K^{1/2}} (z_\kappa^n)^p < \infty,$$

thanks to the estimates (102) and (42). For the first integral, we use that  $\{\bar{u}_\Delta\}_\tau$  converges to  $u_*$  w.r.t.  $\mathbb{W}$ , locally uniformly with respect to  $t \in (0, T)$ . Thus clearly

$$\int_{\Omega} \rho(x) \{\bar{u}_\Delta\}_\tau(t; x) dx \rightarrow \int_{\Omega} \rho(x) u_*(t; x) dx$$

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

$$\begin{aligned} \int_0^T \left| \psi'(t) \int_{\Omega} \rho(x) \{\bar{u}_\Delta\}_\tau(t; x) dx \right|^2 dt &\leq B^2(b-a)\tau \sum_{n=1}^{N_\tau} \int_{\Omega} \bar{u}_\Delta^n(x)^2 dx \\ &= B^2(b-a)\tau \sum_{n=1}^{N_\tau} \delta \sum_{\kappa \in \mathbb{I}_K^{1/2}} z_\kappa^n \leq Q_1 B^2(b-a). \end{aligned}$$

Next, using the strong convergence from (64), it follows that

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

strongly in  $L^1(\Omega)$ , 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:

$$\begin{aligned} \int_0^T \left| \psi(t) \int_{\Omega} \rho'''(x) \partial_x \{\widehat{u}_\Delta\}_\tau dx \right|^2 dt &\leq B^2(b-a)\tau \sum_{n=1}^{N_\tau} \int_{\Omega} (\partial_x \widehat{u}_\Delta^n(x))^2 dx \\ &= B^2(b-a)\tau \sum_{n=1}^{N_\tau} \delta \sum_{k \in \mathbb{I}_K^+} \left( \frac{z_{k+\frac{1}{2}}^n - z_{k-\frac{1}{2}}^n}{\delta} \right)^2 \left( \frac{z_{k+\frac{1}{2}}^n + z_{k-\frac{1}{2}}^n}{2} \right) \\ &\leq B^2(b-a) \left( \tau \sum_{n=1}^{\infty} \sum_{k \in \mathbb{I}_K^+} \left( \frac{z_{k+\frac{1}{2}}^n - z_{k-\frac{1}{2}}^n}{\delta} \right)^4 \right)^{1/2} \left( \tau \sum_{n=1}^{N_\tau} \delta \sum_{\kappa \in \mathbb{I}_K^{1/2}} (z_\kappa^n)^2 \right)^{1/2} \leq 3\bar{\mathcal{H}}^{1/2} Q_2^{1/2} B^2(b-a), \end{aligned}$$

where we have used (42). Finally, the strong convergence implies (64) also implies that

$$(\partial_x \{\widehat{u}_\Delta\}_\tau)^2 \rightarrow (\partial_x \sqrt{u_*})^2$$

strongly in  $L^1(\Omega)$ , for almost every  $t \in (0, T)$ . One more time, we invoke Vitali's theorem, using that

$$\begin{aligned} \int_0^T \left| \psi(t) \int_{\Omega} \rho''(x) \partial_x \left\{ \sqrt{\widehat{u}_\Delta} \right\}_\tau^2(t; x) dx \right|^2 dt &\leq B^2\tau \sum_{n=1}^{N_\tau} \int_{\Omega} (\partial_x \sqrt{\widehat{u}_\Delta^n})^4(x) dx \\ &\leq \frac{1}{2} B^2\tau \sum_{n=1}^{N_\tau} \delta \sum_{k \in \mathbb{I}_K^+} \left( \frac{z_{k+\frac{1}{2}}^n - z_{k-\frac{1}{2}}^n}{\delta} \right)^2 \left[ \left( 1 - \frac{z_{k+\frac{1}{2}}^n}{z_{k-\frac{1}{2}}^n} \right)^2 + \left( 1 - \frac{z_{k-\frac{1}{2}}^n}{z_{k+\frac{1}{2}}^n} \right)^2 \right]. \\ &\leq B^2 \left( \tau \sum_{n=1}^{\infty} \delta \sum_{k \in \mathbb{I}_K^+} \left( \frac{z_{k+\frac{1}{2}}^n - z_{k-\frac{1}{2}}^n}{\delta} \right)^4 \right)^{1/2} \left( \tau \sum_{n=1}^{\infty} \delta \sum_{k \in \mathbb{I}_K^+} \left[ \left( 1 - \frac{z_{k+\frac{1}{2}}^n}{z_{k-\frac{1}{2}}^n} \right)^4 + \left( 1 - \frac{z_{k-\frac{1}{2}}^n}{z_{k+\frac{1}{2}}^n} \right)^4 \right] \right)^{1/2}. \end{aligned}$$

The two terms in the last line are uniformly controlled in view of (42) and (44), respectively.  $\square$

## 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  $\bar{u}_\Delta = \mathbf{u}_\delta[\bar{x}_\Delta]$ , with  $\mathbf{u}_\delta$  defined in (19). The monotonicity of  $\mathbf{H}_\delta$  and  $\mathbf{F}_\delta$  have been obtained in Lemma 12 and 11, respectively.

It remains to show the exponential decay (14) of  $\mathbf{H}_\delta$ . From (the proof of) Lemma 16, it follows for each  $n = 1, 2, \dots$  that

$$\mathcal{H}(\bar{u}_\Delta^n) - \mathcal{H}(\bar{u}_\Delta^{n-1}) = \mathbf{H}_\delta(\bar{x}_\Delta^n) - \mathbf{H}_\delta(\bar{x}_\Delta^{n-1}) \leq -\frac{\tau}{10(b-a)} \text{TV} \left[ \partial_x \sqrt{\widehat{u}_\Delta^n} \right]^2. \quad (88)$$

By the logarithmic Sobolev inequality on  $\Omega$  and thanks to the fact that  $\partial_x \sqrt{\widehat{u}_\Delta^n}(0) = 0$ , we further have that

$$\mathcal{H}(\widehat{u}_\Delta^n) \leq \frac{(b-a)^2}{2\pi^2} \int_\Omega \left( \sqrt{\widehat{u}_\Delta^n} \right)^2(x) dx \leq \frac{(b-a)^3}{2\pi^2} \text{TV} \left[ \partial_x \sqrt{\widehat{u}_\Delta^n} \right]^2. \quad (89)$$

Now combine (88) and (89) with the estimate (105) 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 (14) 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 (57) and (59). The regularity  $\sqrt{u_*} \in L^2(\mathbb{R}_{\geq 0}; H^1(\Omega))$  has been observed in Proposition 20. The strong convergence stated in the same proposition implies that  $\mathcal{F}(u_*)$  is “almost monotone”: indeed, thanks to (64) we may assume — passing to a further subsequence with  $\Delta \rightarrow 0$  if necessary — that

$$\left\{ \sqrt{\widehat{u}_\Delta} \right\}_\tau(t) \rightarrow \sqrt{u_*(t)} \quad \text{strongly in } H^1(\Omega), \text{ for a.e. } t > 0,$$

and therefore also

$$2 \int_\Omega \left\{ \partial_x \sqrt{\widehat{u}_\Delta} \right\}_\tau^2(t; x) dx \rightarrow \mathcal{F}(u_*(t)), \quad \text{for a.e. } t > 0. \quad (90)$$

On the other hand, arguing just like in the proof of (85), it follows that

$$\int_0^\infty \left| \left\{ \mathbf{F}_\delta(\bar{x}_\Delta) \right\}_\tau - 2 \int_\Omega \left\{ \partial_x \sqrt{\widehat{u}_\Delta} \right\}_\tau^2(t; x) dx \right| dt \leq C\delta^{1/4}. \quad (91)$$

Now combine (90) and (91) with the fact that  $\{\mathbf{F}_\delta(\bar{x}_\Delta)\}_\tau$  is decreasing in  $t$ , for each  $\Delta$ , and is  $\Delta$ -uniformly bounded above according to (54). By Helly’s selection principle, there exists a monotone  $f: \mathbb{R}_+ \rightarrow \mathbb{R}_+$  such that  $\mathcal{F}(u_*(t)) = f(t)$  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 (59).

Finally, the weak formulation (16) has been shown in Proposition 21. Simply observe that any  $\varphi \in C_c^\infty(\mathbb{R}_{\geq 0} \times \Omega)$  can be approximated by linear combinations of products  $\psi(t)\rho(x)$  with functions  $\psi \in C^\infty(\mathbb{R}_{\geq 0})$  and  $\rho \in C^\infty(\Omega)$ .  $\square$

## 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 27.** *Suppose that  $X \in C^\infty(\mathbb{R}_{\geq 0} \times [0, M])$  is a classical solution to*

$$\partial_t X = \partial_\xi (Z^2 \partial_\xi^2 Z), \quad (92)$$

*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  $(\bar{x}_\Delta)$  of  $X$  to the respective meshes, given by  $x_k^n := X(n\tau; \xi_k)$  for  $n \in \mathbb{N}$  and  $k \in \{1, \dots, K\}$ , satisfy (13) with an error  $O(\delta^2) + O(\tau)$  as  $\Delta \rightarrow 0$ .*

*Proof.* Given  $\Delta = (\tau; \delta)$ , introduce  $\tilde{Z} : \mathbb{R}_{\geq 0} \times [\delta/2, M - \delta/2] \rightarrow \mathbb{R}_+$  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

$$\partial_\xi^m \tilde{Z}(t; \xi) = \partial_\xi^m Z(t; \xi) + O(\delta^2), \quad (93)$$

for each  $m \in \mathbb{N}$  and locally uniformly in  $(t; \xi)$  as  $\Delta \rightarrow 0$ . Observe that, by definition of  $(\bar{x}_\Delta)$  as restriction of  $X$  to  $\Delta$ , one has

$$z_\kappa^n = \frac{\delta}{x_{\kappa+\frac{1}{2}}^n - x_{\kappa-\frac{1}{2}}^n} = \tilde{Z}(n\tau; \xi_\kappa). \quad (94)$$

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

$$z_* = \tilde{Z}(n\tau; \xi_k), \quad z'_* = \partial_\xi Z(n\tau; \xi_k), \quad \dots, \quad \dot{z}_* = \partial_t Z(n\tau; \xi_k).$$

Relation (94) and a standard Taylor expansion of  $\tilde{Z}$  around  $\xi = \xi_k$  yield

$$\begin{aligned} (z_{k+\frac{1}{2}}^n)^2 \left( \frac{z_{k+\frac{3}{2}}^n - 2z_{k+\frac{1}{2}}^n + z_{k-\frac{1}{2}}^n}{\delta^2} \right) &= (z_*^2 + \delta z_* z'_* + O(\delta^2)) (z_*'' + \frac{\delta}{2} z_*''' + O(\delta^2)) \\ &= z_*^2 z_*'' + \frac{\delta}{2} (2z_* z'_* z_*'' + z_*^2 z_*''') + O(\delta^2). \end{aligned}$$

The same expansion — with  $(-\delta)$  in place of  $\delta$  — is obtained for  $z_{k+\frac{3}{2}}^n, z_{k+\frac{1}{2}}^n$  and  $z_{k-\frac{1}{2}}^n$  replaced by  $z_{k+\frac{1}{2}}^n, z_{k-\frac{1}{2}}^n$  and  $z_{k-\frac{3}{2}}^n$ , respectively. Therefore,

$$\begin{aligned} \frac{1}{\delta} \left[ (z_{k+\frac{1}{2}}^n)^2 \left( \frac{z_{k+\frac{3}{2}}^n - 2z_{k+\frac{1}{2}}^n + z_{k-\frac{1}{2}}^n}{\delta^2} \right) - (z_{k-\frac{1}{2}}^n)^2 \left( \frac{z_{k+\frac{1}{2}}^n - 2z_{k-\frac{1}{2}}^n + z_{k-\frac{3}{2}}^n}{\delta^2} \right) \right] \\ = 2z_* z'_* z_*'' + z_*^2 z_*''' + O(\delta). \end{aligned} \quad (95)$$

Next, observe that the expression on the left-hand side remains invariant under the simultaneous exchange of  $z_{k+\frac{1}{2}}^n$  with  $z_{k-\frac{1}{2}}^n$  and of  $z_{k+\frac{3}{2}}^n$  with  $z_{k-\frac{3}{2}}^n$ . 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 (94) and (93), the term of order  $\delta^0$  can be written as

$$2z_* z'_* z''_* + z_*^2 z'''_* = \partial_\xi (\tilde{Z}(n\tau; \xi_k)^2 \partial_\xi^2 \tilde{Z}(n\tau; \xi_k)) = \partial_\xi (Z(n\tau; \xi_k)^2 \partial_\xi^2 Z(n\tau; \xi_k)) + O(\delta^2).$$

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

$$\frac{x_k^n - x_k^{n-1}}{\tau} = \frac{1}{\tau} (X(n\tau; \xi_k) - X((n-1)\tau; \xi_k)) = \partial_t X(n\tau; \xi_k) + O(\tau), \quad (96)$$

thanks to the smoothness of  $X$  in time. Comining (95)&(96) with the continuous equation (92), 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 are going to change our setting and allow *non-equidistant* mass grids. That is, the mass discretization of  $[0, M]$  is determined by a vector  $\vec{\delta} = (\xi_0, \xi_1, \xi_2, \dots, \xi_{K-1}, \xi_K)$ , with

$$0 = \xi_0 < \xi_1 < \dots < \xi_{K-1} < \xi_K = M,$$

and we introduce accordingly the distances

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

for  $\kappa \in \mathbb{I}_K^{1/2}$  and  $k \in \mathbb{I}_K^+$ , respectively. The piecewise constant density function  $\bar{u} \in \mathcal{P}_{\vec{\delta}}(\Omega)$  corresponding to a vector  $\vec{x} \in \mathbb{R}^{K-1}$  is now given by

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

The Wasserstein-like metric needs to be adapted as well: the scalar product  $\langle \cdot, \cdot \rangle_\delta$  is replaced by

$$\langle \vec{v}, \vec{w} \rangle_{\vec{\delta}} = \sum_{k \in \mathbb{I}_K^+} \delta_k v_k w_k.$$

Hence the metric gradient  $\nabla_{\vec{\delta}} f(\vec{x}) \in \mathbb{R}^{K-1}$  of a function  $f : \mathfrak{r}_{\vec{\delta}} \rightarrow \mathbb{R}$  at  $\vec{x} \in \mathfrak{r}_{\vec{\delta}}$  is given by

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

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

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

attains the form

$$\begin{aligned} \frac{x_k^n - x_k^{n-1}}{\tau} = \frac{1}{\delta_k} & \left[ \frac{(z_{k+\frac{1}{2}}^n)^2}{\delta_{k+\frac{1}{2}}} \left( \frac{z_{k+\frac{3}{2}}^n - z_{k+\frac{1}{2}}^n}{\delta_{k+1}} - \frac{z_{k+\frac{1}{2}}^n - z_{k-\frac{1}{2}}^n}{\delta_k} \right) \right. \\ & \left. - \frac{(z_{k-\frac{1}{2}}^n)^2}{\delta_{k-\frac{1}{2}}} \left( \frac{z_{k+\frac{1}{2}}^n - z_{k-\frac{1}{2}}^n}{\delta_k} - \frac{z_{k-\frac{1}{2}}^n - z_{k-\frac{3}{2}}^n}{\delta_{k-1}} \right) \right]. \end{aligned} \quad (97)$$

6.2.2. *Initial condition.* Our main experiments are carried out using the by now classical test case from [4], that is

$$u^0(x) = \epsilon + \cos^{16}(\pi x), \quad \text{on } \Omega = [0, 1], \quad (98)$$

with  $\epsilon = 10^{-3}$ . The mass grid  $\vec{\delta}$  is chosen in such a way that  $u_\Delta^0$  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  $\vec{x}_\Delta^0$  for  $\vec{x}$  attains the simple form

$$x_k^0 = a + \frac{b-a}{K}k. \quad (99)$$

To construct  $\vec{\delta}$ , we first calculate the cumulative distribution function  $U^0 : [0, 1] \rightarrow [0, M]$  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  $k = 0, 1, \dots, K$ .

**Remark 28.** *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 are 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  $\vec{x}_\Delta^0$ , the fully discrete solution is calculated inductively by solving the implicit Euler scheme (97) for  $\vec{x}_\Delta^n$ , given  $\vec{x}_\Delta^{n-1}$ . In each time step, a damped Newton iteration is performed, with the solution from the previous time step as initial guess. Slow convergence of the Newton iteration has been observed in situations where the density  $\bar{u}_\Delta^{n-1}$  has steep gradients and/or intervals of very low values.

Our reference solution is calculated with the scheme described in [20], which is fully variational as well, but uses different ansatz functions for the Lagrangian maps. Even without a rigorous result on uniqueness of weak solutions, it seems reasonable to expect that both schemes should approximate the same solution. A technical issue with the comparison of our solution to the reference solution is that both use a different way for the reconstruction of the density from the Lagrangian map. This difference camouflages the true approximation error in the plain  $L^2$ -differences. For a fair comparison, we calculate the  $L^2$ -difference of the linear interpolations of the values for the density with respect to the nodes of the Lagrangian maps.

6.2.4. *Observed rate of convergence.* Figure 1 provides a qualitative picture of the evolution with initial condition  $u^0$ : the plot on the left shows the density function  $\bar{u}_\Delta$  at several instances in time, the plot on the right visualizes the motion of the mesh points  $\{x_k\}_\tau$  associated to the Lagrangian maps  $X_\Delta$  in continuous 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 the comparison principle does obviously not hold for the DLSS equation. Both figures has been generated using  $K = 200$  spatial grid points and the time step size  $\tau = 10^{-6}$ .

For numerical analysis 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, using  $K = 25, 50, 100, 200$ . Figure 2/Left shows the corresponding  $L^2$ -error between the solution to our scheme and the reference solution, evaluated at time  $T = 10^{-5}$ . It is clearly seen that the error decays with an almost perfect rate of  $\delta^2 \propto K^{-2}$ . For the second series of experiments, we keep the spatial discretization parameter  $K = 800$  fixed and run our scheme with the time step

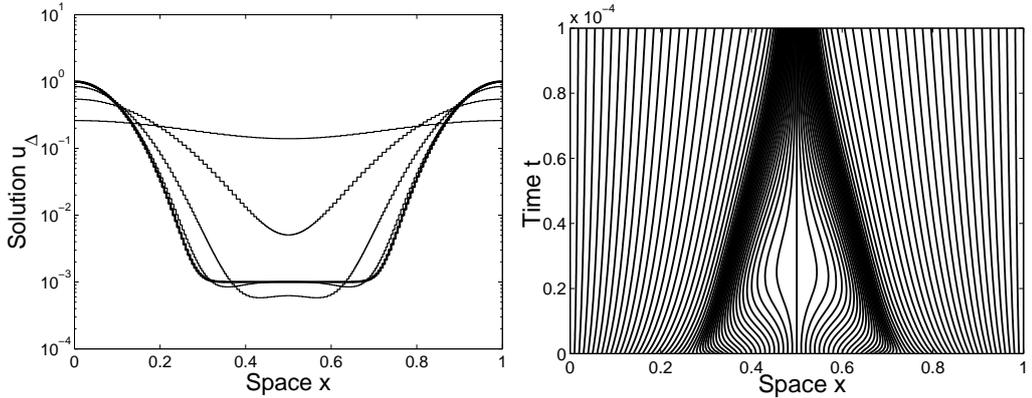


FIGURE 1. *Left:* snapshots of the densities  $\bar{u}_\Delta$  for the initial condition (98) at times  $t = 0$  and  $t = 10^i$ ,  $i = -6, \dots, -3$ , using  $K = 200$  grid points and the time step size  $\tau = 10^{-6}$ . *Right:* associated particle trajectories.

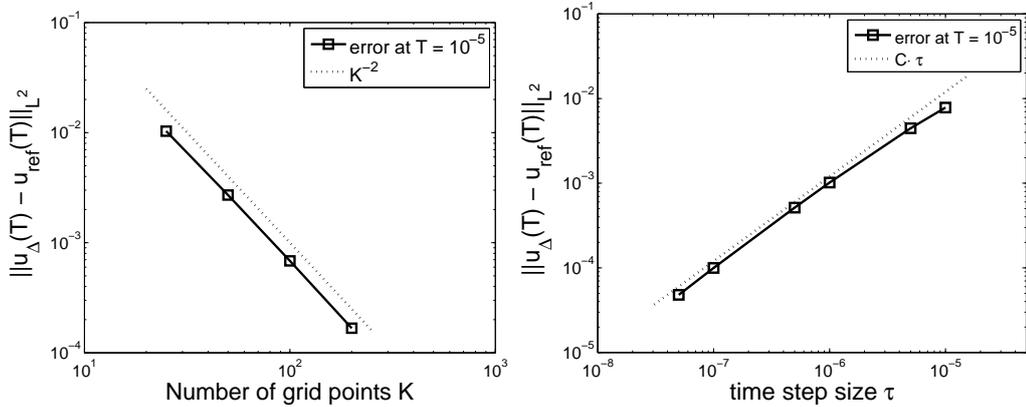


FIGURE 2. Numerical error analysis for  $u^0$  from (98). *Left:* fixed time step size  $\tau = 10^{-8}$  and  $K = 25, 50, 100, 200$  spatial grid points. The  $L^2$ -errors are evaluated at  $T = 5 \cdot 10^{-6}$ . *Right:* fixed  $K = 800$  using  $\tau = 10^{-5}, 5 \cdot 10^{-6}, 10^{-6}, 5 \cdot 10^{-7}, 10^{-7}, 5 \cdot 10^{-8}$ . The error is evaluated at  $T = 10^{-5}$ .

sizes  $\tau = 10^{-5}, 5 \cdot 10^{-6}, 10^{-6}, 5 \cdot 10^{-7}, 10^{-7}, 5 \cdot 10^{-8}$ , respectively. The corresponding  $L^2$ -error at  $T = 10^{-5}$  is plotted in Figure 2/Right. It is proportional to  $\tau$ .

6.2.5. *Discontinuous initial data.* One of the conclusions of Theorem 2 is that the discrete approximations  $u_\Delta$  converge also for (a large class of) non-regular initial data  $u^0$ . For illustration of this feature, we consider the discontinuous initial density function

$$u_{\text{discont}}^0 = \begin{cases} 1 & x \in [0, \frac{1}{3}] \cup [\frac{2}{3}, 1], \\ 10^{-3} & x \in (\frac{1}{3}, \frac{2}{3}) \end{cases} \quad (100)$$

instead of  $u^0$  from (98). According to our hypothesis (15), we need to use a sufficiently high spatial and temporal resolution. In practice, this is done in an adaptive way: the  $K$  points of

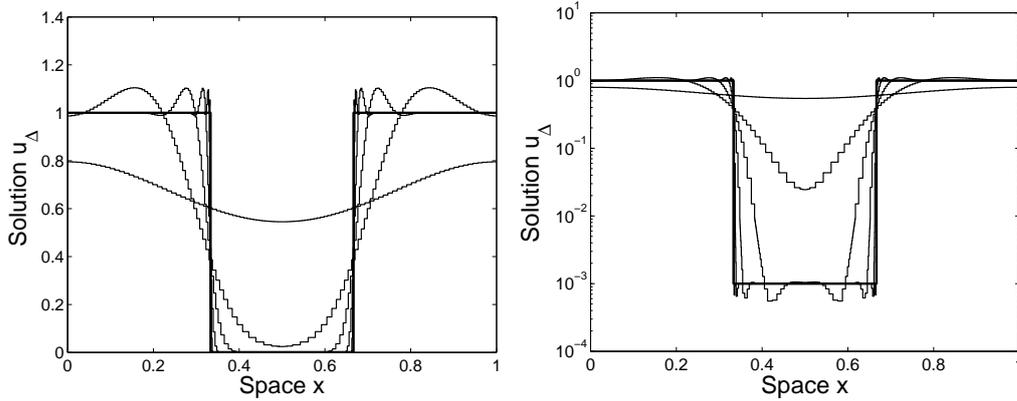


FIGURE 3. Snapshots of the densities  $\bar{u}_\Delta$  for the initial condition (100) at times  $t = 0$  and  $t = 10^i$ ,  $i = -13, -11, \dots, -5, -3$ , using  $K = 200$  grid points with linear (left) and logarithmic (right) scaling.

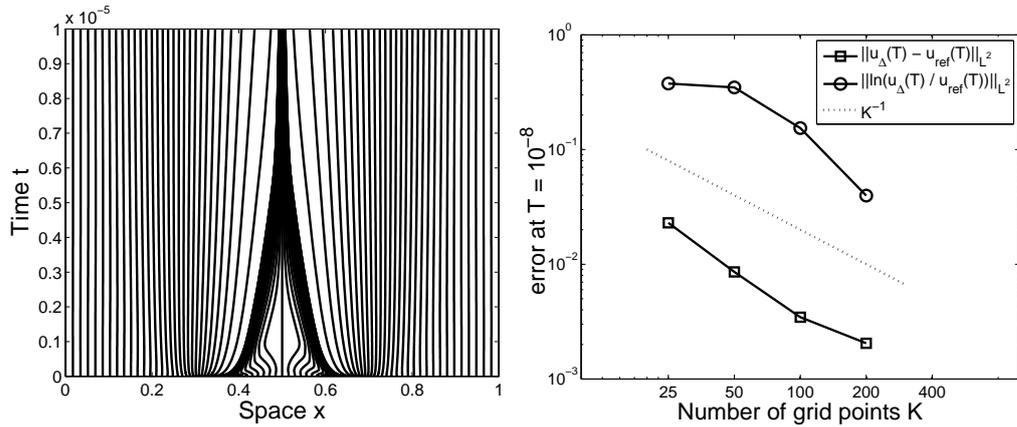


FIGURE 4. *Left*: associated particle trajectories of  $\bar{u}_\Delta$  using the initial condition (100). *Right*: Numerical error analysis for  $u_{\text{discont}}^0$  from (100) with fixed  $\tau$  and  $K = 25, 50, 100, 200$  spatial grid points. The  $L^2$ -errors are evaluated at  $T = 10^{-8}$ .

the initial grid  $\bar{x}_\Delta^0$  are not placed equidistantly, but with a higher refinement around the points of discontinuity; the applied time step  $\tau$  is extremely small (down to  $10^{-13}$ ) during the initial phase of the evolution, and is larger (up to  $10^{-9}$ ) at later times.

Figure 3 provides a qualitative picture of the fully discrete evolution for  $K = 200$  grid points: 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 Figure 4/Left. Notice the fast motion of the grid points near the discontinuities.

To estimate the rate of convergence, we performed a series of experiments using  $K = 25, 50, 100$  and 200 spatial grid points. For comparison, we calculated a highly refined solution of the following semi-implicit reference scheme,

$$\frac{u_{\text{ref}}^{n+1} - u_{\text{ref}}^n}{\tau} = -\Delta_2(u_{\text{ref}}^n \Delta_2 \ln(u_{\text{ref}}^{n+1})),$$

where  $\Delta_2$  is the standard central difference operator  $\Delta_2$ . The reference scheme is run with  $K = 800$  spatial grid point. An adaptive choice of the time step  $\tau$  needs to be made in order to avoid that the reference solution  $u_{\text{ref}}$  breaks down due to loss of positivity. The  $L^2$ -differences of the densities and of their logarithms have been evaluated at  $T = 10^{-8}$ , see Figure 4/Right. As expected, the rate of convergence is no longer quadratic in  $\delta \propto K^{-1}$ ; instead, the error decays approximately linearly.

#### APPENDIX A. SOME TECHNICAL LEMMAS

**Lemma 29.** *For each  $p > 1$  and  $\vec{x} \in \mathfrak{r}_\delta$  with  $\vec{z} = \mathbf{z}[\vec{x}]$ , one has that*

$$\sum_{\kappa \in \mathbb{I}_K^{1/2}} \left( \frac{\delta}{z_\kappa} \right)^p = \sum_{\kappa \in \mathbb{I}_K^{1/2}} (x_{\kappa+\frac{1}{2}} - x_{\kappa-\frac{1}{2}})^p \leq (b-a)^p. \quad (101)$$

*Proof.* The first equality is simply the definition (20) 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

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

**Lemma 30.** *For each  $\vec{x} \in \mathfrak{r}_\delta$  with  $\vec{z} = \mathbf{z}[\vec{x}]$ , one has that*

$$\frac{\delta}{b-a} \leq z_\kappa \leq M^{1-1/q} \left( \delta \sum_{k \in \mathbb{I}_K^+} \left| \frac{z_{k+\frac{1}{2}} - z_{k-\frac{1}{2}}}{\delta} \right|^q \right)^{1/q} + \frac{M}{b-a} \quad \text{for all } \kappa \in \mathbb{I}_K^{1/2}, \quad (102)$$

and consequently,

$$z_\kappa \leq (2M\mathbf{F}_\delta[\vec{x}])^{1/2} + \frac{M}{b-a} \quad \text{for all } \kappa \in \mathbb{I}_K^{1/2}. \quad (103)$$

*Proof.* The first estimate in (102) is an immediate consequence of the definition of  $z_\kappa$  in (20). To prove the second estimate, let  $\kappa_* \in \mathbb{I}_K^{1/2}$  be such that  $z_{\kappa_*} = \max z_k$ . Observe that there exists a  $\kappa_* \in \mathbb{I}_K^{1/2}$  such that

$$z_{\kappa_*} \leq \frac{M}{b-a} \leq z_{\kappa_*}. \quad (104)$$

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

$$z_{\kappa_*} - z_{\kappa_*} \leq \sum_{k \in \mathbb{I}_K^+} |z_{k+\frac{1}{2}} - z_{k-\frac{1}{2}}| \leq \left( \delta \sum_{k \in \mathbb{I}_K^+} 1 \right)^{1-1/q} \left( \delta \sum_{k \in \mathbb{I}_K^+} \left| \frac{z_{k+\frac{1}{2}} - z_{k-\frac{1}{2}}}{\delta} \right|^q \right)^{1/q}.$$

Now combine this with (104).  $\square$

**Lemma 31.** *With  $\hat{u}$  and  $\bar{u}$  being, respectively, the piecewise linear and the piecewise constant densities associated to a given vector  $\vec{x}$ , then*

$$\mathbf{H}_\delta(\vec{x}) = \mathcal{H}(\bar{u}) \leq \mathcal{H}(\hat{u}). \quad (105)$$

*Proof.* First observe that

$$\int_0^1 \ln(p(1-\lambda) + q\lambda) d\lambda = \frac{p \ln p - q \ln q}{p - q} - 1 \geq \frac{1}{2}(\ln p + \ln q), \quad (106)$$

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

$$\int_{\Omega} \bar{u}(x) \log \bar{u}(x) dx = \delta \sum_{\kappa \in \mathbb{I}_K^{1/2}} \log z_{\kappa} = \delta \frac{\log z_{1/2} + \log z_{K-1/2}}{2} + \delta \sum_{k \in \mathbb{I}_K^+} \frac{\log z_{k+\frac{1}{2}} + \log z_{k-\frac{1}{2}}}{2},$$

and on the other hand,

$$\begin{aligned} \int_{\Omega} \hat{u}(x) \log \hat{u}(x) dx &= \int_0^M \log \hat{z}(\xi) d\xi \\ &= \frac{\delta}{2} (\log z_0 + \log z_K) + \delta \sum_{k \in \mathbb{I}_K^+} \int_0^1 \ln(z_{k-\frac{1}{2}}(1-\lambda) + z_{k+\frac{1}{2}}\lambda) d\lambda \\ &\geq \delta \frac{\log z_{1/2} + \log z_{K-1/2}}{2} + \delta \sum_{k \in \mathbb{I}_K^+} \frac{\log z_{k+\frac{1}{2}} + \log z_{k-\frac{1}{2}}}{2}, \end{aligned}$$

where we have used (106). This clearly implies (105).  $\square$

**Lemma 32** (Gargliardo-Nirenberg inequality). *For each  $f \in H^1(\Omega)$ , one has that*

$$\|f\|_{C^{1/6}(\Omega)} \leq (9/2)^{1/3} \|f\|_{H^1(\Omega)}^{2/3} \|f\|_{L^2(\Omega)}^{1/3}. \quad (107)$$

*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

$$|f(x)^{3/2} - f(y)^{3/2}| \leq \frac{3}{2} \int_x^y 1 \cdot f(z)^{1/2} |f'(z)| dz \leq \frac{3}{2} |x - y|^{1/4} \|f\|_{L^2(\Omega)}^{1/2} \|f'\|_{L^2(\Omega)}.$$

Since  $f \geq 0$ , we can further estimate

$$|f(x) - f(y)| \leq |f(x)^{3/2} - f(y)^{3/2}|^{2/3} \leq (3/2)^{2/3} |x - y|^{1/6} \|f\|_{L^2(\Omega)}^{1/3} \|f\|_{H^1(\Omega)}^{1/3}.$$

This shows (107) for non-negative 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  $\|f_{\pm}\|_{H^1(\Omega)} \leq \|f\|_{H^1(\Omega)}$ ,

$$\|f\|_{C^{1/6}(\Omega)} \leq \|f_+\|_{C^{1/6}(\Omega)} + \|f_-\|_{C^{1/6}(\Omega)} \leq 2(3/2)^{2/3} \|f\|_{L^2(\Omega)}^{1/3} \|f\|_{H^1(\Omega)}^{1/3}.$$

This proves the claim.  $\square$

## 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] A. BLANCHET, 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.
- [4] 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.
- [5] 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.
- [6] 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.

- [7] 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.
- [8] M. J. CÁCERES, 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.
- [9] 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.
- [10] 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.
- [11] 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.
- [12] J. A. CARRILLO AND D. SLEPČEV, *Example of a displacement convex functional of first order*, *Calc. Var. Partial Differential Equations*, 36 (2009), pp. 547–564.
- [13] 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.
- [14] J. A. CARRILLO AND M.-T. WOLFRAM, *A finite element method for nonlinear continuity equations in lagrangian coordinates*. Working paper.
- [15] 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.
- [16] P. DEGOND, F. MÉHATS, AND C. RINGHOFER, *Quantum energy-transport and drift-diffusion models*, *J. Stat. Phys.*, 118 (2005), pp. 625–667.
- [17] 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.
- [18] 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.
- [19] ———, *Fluctuations of a stationary nonequilibrium interface*, *Phys. Rev. Lett.*, 67 (1991), pp. 165–168.
- [20] 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.
- [21] L. C. EVANS, O. SAVIN, AND W. GANGBO, *Diffeomorphisms and nonlinear heat flows*, *SIAM J. Math. Anal.*, 37 (2005), pp. 737–751.
- [22] 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.
- [23] ———, *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.
- [24] 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.
- [25] N. GIGLI AND J. MAAS, *Gromov-Hausdorff convergence of discrete transportation metrics*, *SIAM J. Math. Anal.*, 45 (2013), pp. 879–899.
- [26] 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).
- [27] ———, *Lagrangian numerical approximations to one-dimensional convolution-diffusion equations*, *SIAM J. Sci. Comput.*, 28 (2006), pp. 1203–1227 (electronic).
- [28] 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).
- [29] R. JORDAN, D. KINDERLEHRER, AND F. OTTO, *The variational formulation of the Fokker-Planck equation*, *SIAM J. Math. Anal.*, 29 (1998), pp. 1–17.
- [30] A. JÜNGEL, *Transport equations for semiconductors*, vol. 773 of *Lecture Notes in Physics*, Springer-Verlag, Berlin, 2009.
- [31] A. JÜNGEL AND D. MATTHES, *An algorithmic construction of entropies in higher-order nonlinear PDEs*, *Nonlinearity*, 19 (2006), pp. 633–659.
- [32] ———, *The Derrida-Lebowitz-Speer-Spohn equation: existence, nonuniqueness, and decay rates of the solutions*, *SIAM J. Math. Anal.*, 39 (2008), pp. 1996–2015.
- [33] 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).
- [34] ———, *A positivity-preserving numerical scheme for a nonlinear fourth order parabolic system*, *SIAM J. Numer. Anal.*, 39 (2001), pp. 385–406 (electronic).
- [35] 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.

- [36] 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.
- [37] D. KINDERLEHRER AND N. J. WALKINGTON, *Approximation of parabolic equations using the Wasserstein metric*, M2AN Math. Model. Numer. Anal., 33 (1999), pp. 837–852.
- [38] 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.
- [39] 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.
- [40] 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.
- [41] G. RUSSO, *Deterministic diffusion of particles*, Comm. Pure Appl. Math., 43 (1990), pp. 697–733.
- [42] S. SERFATY, *Gamma-convergence of gradient flows on Hilbert and metric spaces and applications*, Discrete Contin. Dyn. Syst., 31 (2011), pp. 1427–1451.
- [43] C. VILLANI, *Topics in optimal transportation*, vol. 58 of Graduate Studies in Mathematics, American Mathematical Society, Providence, RI, 2003.

DANIEL MATTHES, ZENTRUM MATHEMATIK, TU MÜNCHEN, BOLTZMANNSTR. 3, D-85748 GARCHING, GERMANY  
E-mail address: [matthes@ma.tum.de](mailto:matthes@ma.tum.de)

HORST OSBERGER, ZENTRUM MATHEMATIK, TU MÜNCHEN, BOLTZMANNSTR. 3, D-85748 GARCHING, GERMANY  
E-mail address: [osberger@ma.tum.de](mailto:osberger@ma.tum.de)