A GRADIENT FLOW SCHEME FOR NONLINEAR FOURTH ORDER EQUATIONS

Abstract. We propose a method for numerical integration of Wasserstein gradient flows based on the classical minimizing movement scheme. In each time step, the discrete approximation is obtained as the solution of a constrained quadratic minimization problem on a finite-dimensional function space. Our method is applied to the nonlinear fourth-order Derrida-Lebowitz-Speer-Spohn equation, which arises in quantum semiconductor theory. We prove well-posedness of the scheme and derive a priori estimates on the discrete solution. Furthermore, we present numerical results which indicate second-order convergence and unconditional stability of our scheme. Finally, we compare these results to those obtained from different semi- and fully implicit finite difference discretizations.

Bertram Düring

Institut für Analysis und Scientific Computing
Technische Universität Wien, 1040 Wien, Austria

Daniel Matthes

Institut für Analysis und Scientific Computing
Technische Universität Wien, 1040 Wien, Austria

Josipa Pina Milišić

Faculty of Electrical Engineering and Computing
University of Zagreb, 10000 Zagreb, Croatia

1. **Introduction.** Evolution equations with an underlying gradient flow structure have since long been of special interest in analysis and mathematical physics. In particular, transport equations that allow for a variational formulation with respect to the L^2-Wasserstein metric have attracted a lot of attention after the linear Fokker-Planck equation has been put into gradient flow form in the seminal paper by Jordan, Kinderlehrer and Otto [19]. Various well-established non-linear evolution equations have subsequently been shown to constitute Wasserstein gradient flows, like the porous medium equation [29] and conservation equations for interacting gases [10]. The gradient flow formulation does not only allow for a geometric interpretation of the dynamics, but also paves the way for easy proofs of a priori

2000 *Mathematics Subject Classification.* Primary: 65M99; Secondary: 35K35, 35Q40.

Key words and phrases. Wasserstein gradient flow, Higher-order diffusion equation, numerical solution.

The authors acknowledge support from the Austrian-Croatian Project of the Austrian Exchange Service (ÖAD) and Ministry of Science, Education and Sports of the Republic of Croatia (MZOS). Bertram Düring is supported by the Deutsche Forschungsgemeinschaft, grant JU 359/6 (Forschergruppe 518). Daniel Matthes is supported by the Deutsche Forschungsgemeinschaft, grant JU 359/7. Bertram Düring and Daniel Matthes thank the Department of Applied Mathematics at the Faculty of Electrical Engineering and Computing, University of Zagreb, where a part of this research has been carried out, for the kind hospitality.

bounds, the estimation of equilibration rates etc. Moreover, the gradient flow formulation gives rise to a natural semi-discretization in time of the evolution by means of the minimizing movement scheme (see, e.g. [1]), which constitutes a time-discrete minimization problem for the (sum of kinetic and potential) energy.

On the other hand, nonlinear diffusion equations of fourth (and higher) order have become increasingly important in pure and applied mathematics. Many of them could be formulated in a variational framework as well, i.e., they have been interpreted as gradient flows with respect to some metric structure. Well-known examples are Cahn-Hilliard equations modeling gas mixtures [9, 16], lubrication equations describing the rupture of thin viscous films [4, 5, 13], and fluid-type quantum models for semiconductors device simulations [20]. Apart from their obvious relevance in theoretical physics and in engineering applications, these equations are of great interest in mathematical analysis, since the behavior of their solutions is typically much richer than that of the well-studied second order nonlinear diffusion processes.

When it comes to solve equations of a gradient flow type numerically, it is natural to ask for appropriate schemes that respect the equation's special structure in some way. Such adapted schemes have been designed for various second-order equations; they guarantee for example strict monotonicity of the energy functional on the time-discrete level. See e.g. the recent paper [11] and references therein. Surprisingly, in the numerical approximation of fourth order equations, gradient flow structures have seemingly been neglected so far. In this paper, we propose and study a fully discrete variant of the minimizing movement scheme for numerical solution of the nonlinear fourth order Derrida-Lebowitz-Speer-Spohn equation (or DLSS equation for short),

$$\partial_t u = -\big(u(\ln u)_{xx}\big)_{xx} = -2\Big(u\Big(\frac{(\sqrt{u})_{xx}}{\sqrt{u}}\Big)_x\Big)_x, \quad x \in \Omega,\, t > 0. \tag{1}$$

The variational character of (1) has been investigated only recently in [17], where it is proven that (1) — with appropriate boundary conditions — constitutes the L^2-Wasserstein gradient flow of the Fisher information

$$\mathcal{F}[u] = \frac{1}{2}\int_\Omega u(x)\big(\ln u(x)\big)_x^2\,dx = 2\int_\Omega \big(\sqrt{u(x)}\big)_x^2\,dx. \tag{2}$$

The numerical discretization of (1) is performed on the one-dimensional torus $\Omega = \mathbb{T}$, i.e., on the interval $\mathbb{T} \cong [0,1]$ with periodic boundary conditions, and for even solutions, i.e., $u(x) = u(1-x)$. This helps to avoid various technical difficulties, that would only obscure the key ideas behind the numerical scheme.

Let us emphasize that good numerical schemes for solution of (1) and other fourth order diffusion equations are of importance, even in one spatial dimension. Despite many advances in the study of existence, regularity and long-time behavior, fundamental questions on qualitative properties of the solutions to fourth order equations remain open. For instance, their solutions do not obey comparison principles in general, and thus the formation of vacuum from an initially strictly positive density cannot be ruled out a priori. Even in the case of the well-studied thin film equation, the analytical results on (the absence of) film rupture are currently rather unsatisfactory from the physical perspective; see [28] for a summary of the state of the art. Essentially, one has to rely on suitable numerical experiments [2, 3, 6] for qualitative investigations.

Before we describe our numerical method in more detail, some remarks are in order concerning the history of the DLSS equation and on previous analytical and numerical results. Equation (1) has been originally derived by Derrida, Lebowitz, Speer and Spohn in a non-variational context: in [15], the solution u was related to fluctuations (about the straight diagonal) of the interface between the regions of predominantly positive and negative spins in the anchored Toom model. Equation (1) also arises as a particular scaling limit of the quantum drift diffusion model that is well-established in semiconductor physics [20]; see [14] for a recent derivation.

An overview on the numerous analytical results for the DLSS equation (1), also in multiple dimensions, is provided in the review article [23]. We just mention that global non-negative, spatially periodic weak solutions u to (1) of regularity $\sqrt{u} \in L^2(0,\infty;H^2(\mathbb{T}))$ have been constructed in [22]; see [25, 17] for related results with other boundary conditions. Moreover, if $u_0 \in H^1(\mathbb{T})$ is strictly positive, then the weak solution $u(t)$ is classical and C^∞-smooth in space, at least up to the first time at which $u(t)$ loses strict positivity [7]. Various smallness conditions on u_0 have been identified [7, 27] under which the classical solution is actually global.

Numerical approximation of solutions to (1) is a non-trivial task, due to the equation's fourth order in x, because of the strong nonlinearity, and also since non-negativity of u needs to be preserved by the numerical method. Finite difference schemes have been proposed at least in three different places: The most direct (but computational very expensive) approach has been taken in [18], where (1) has been discretized with a fully implicit scheme. In [26], the fourth order equation is rewritten as a system of a second order diffusion equation and a second order elliptic problem,

$$\partial_t u = (uF_x)_x, \qquad -2\frac{\sqrt{u}_{xx}}{\sqrt{u}} = F.$$

This system is then solved by an implicit Euler method. The resulting scheme is shown to preserve strict positivity and to be convergent. Another scheme has been devised in [12]. Introducing the logarithmic variable $v = \ln u$, the DLSS equation is equivalently rewritten as

$$U(v)\partial_t v = -\big(U(v)v_{xx}\big)_{xx} \quad \text{with} \quad U(v) = e^v.$$

This equation is numerically solved by a semi-implicit time integration, employing v from the previous time step to define the coefficient function $U(v)$. In addition to being positivity preserving, this scheme was shown to dissipate the physical entropy.

Let us now describe the ingredients of our own method. As mentioned before, it is a practical implementation of the minimizing movement scheme. Our method is iterative and consists of an outer and an inner loop. In each time step of the outer loop a constrained quadratic optimization problem for the Fisher information (2) is solved on a finite-dimensional space of ansatz functions. These subproblems are solved iteratively in the inner loop by applying Newton's method to the optimality system, leading to a sequential quadratic programming method.

Similarly as in [11], the discrete equations are formulated in Lagrangian coordinates, since this allows to compute the Wasserstein distance between two densities in an easy way. On the other hand, there is a fundamental difference between our approach and the one developed in [11], since discretization and minimization are performed in opposite order: we generate a genuine solution of the minimizing movement scheme on a restricted (finite-dimensional) function space, whereas in [11], the (spatially continuous) Euler-Lagrange equations for the minimizing movement are

discretized and solved. Choosing suitable ansatz functions for the Lagrangian map, our method is able to deal with fourth order equations like (1).

In result, we obtain (apparently for the first time) a fully practical numerical scheme for a non-linear fourth order equation that respects its Wasserstein gradient flow structure. It is constructed in such a way that the discrete solution dissipates the Fisher information "as fast as possible", just like the original gradient flow. Some benefits of this adapted scheme are the built-in conservation of mass and non-negativity of the solution. Also, the formulation in Lagrangian coordinates induces a time-dependent numerical resolution of x-space, with a grid that is naturally adaptive to the movement of the density. Moreover, since our method derives from the implicit minimizing movement scheme, it does not suffer from restrictions on the time step size. Indeed, the numerical results suggest unconditional stability of the scheme.

We compare the numerical properties of our newly developed method with those of two classes of finite difference approximations. First, we use a fully implicit Euler scheme in the spirit of [18], that is computationally expensive, but is assumed to deliver the most reliable results for small mesh sizes. Second, we employ two semi-implicit schemes, which are much less expensive and can provide a high order of approximation for sufficiently smooth solutions.

It should be emphasized that although our presentation is focused on the DLSS equation, the proposed numerical method easily adapts to a wide class of Wasserstein gradient flows, replacing the Fisher information by other potentials. For instance, second-order diffusion or aggregation equations are easily dealt with. In principle, also potentials containing two (or more) derivatives of u are admissible, corresponding to evolution equations of sixth (or higher) order — see [24] for an example. In such cases, the ansatz space for the Lagrangian transformation needs to be modified accordingly, but there is no principal difficulty.

The structure of this paper is as follows. In Section 2 below, we define our novel numerical scheme and prove its well-posedness. In Section 3, we introduce the (standard) finite difference schemes, which are subsequently used for evaluation of the numerical quality of the new method. The results from numerical experiments are summarized in Section 4.

2. Definition of the discrete gradient flow scheme.

2.1. Wasserstein distances for periodic functions.

To start with, we recall and establish some basic facts about mass transportation and Wasserstein distances. We start with the fundamental notions: Let X be a Riemannian manifold with distance function $d : X \times X \to \mathbb{R}_{\geq 0}$, and denote by $\mathcal{P}_M(X)$ the convex set of measures with fixed mass $M > 0$ on X.

Definition 2.1. The L^2-Wasserstein distance of two measures $\mu_1, \mu_2 \in \mathcal{P}_M(X)$ is defined by

$$\mathbf{W}[\mu_1, \mu_2]^2 = \inf_{\pi \in \Pi(\mu_1, \mu_2)} \int_{X \times X} d(x, y)^2 \, d\pi(x, y), \tag{3}$$

where Π is the set of *transport plans* connecting μ_1 to μ_2, i.e. the set of all measures on $X \times X$ with respective marginals μ_1 and μ_2,

$$\pi(A \times X) = \mu_1(A), \quad \pi(X \times B) = \mu_2(B)$$

for all measurable sets A, $B \subset X$.

If the measures μ_1 and μ_2 possess densities u_1 and u_2 with respect to a fixed background measure on X, then we write — by abuse of notation — $\mathbf{W}[u_1, u_2]$ instead of $\mathbf{W}[\mu_1, \mu_2]$.

For later reference, we remark that the infimum in (3) is indeed a minimum, realized by some *optimal plan* $\pi_{opt} \in \Pi(\mu_1, \mu_2)$.

If $X = (a, b) \subset \mathbb{R}$ is an (possibly infinite) interval, then there exists an explicit formula to calculate the L^2-Wasserstein distance $\mathbf{W}$. For two measures $\mu_1, \mu_2 \in \mathcal{P}_M(X)$, define their distribution functions U_1 and U_2 by

$$U_i : (a, b) \to [0, M], \quad U_i(x) = \mu_i\big((a, x]\big) \qquad (i = 1, 2). \tag{4}$$

As the U_i are right-continuous and monotonically increasing, they possess right-continuous increasing pseudo-inverse functions $G_i : [0, M] \to [a, b]$,

$$G_i(\omega) = \inf\{x \in (a, b) \,|\, U(x) > \omega\}.$$

Then

$$\mathbf{W}[u_1, u_2]^2 = \int_0^M |G_1(\omega) - G_2(\omega)|^2 \, d\omega. \tag{5}$$

This formula does *not* extend to $X = \mathbb{T} \cong (0, 1)$, because of the topology induced by the periodic boundary conditions,

$$d(x, y) = \min\big(|x - y|, 1 - |x - y|\big).$$

The problem is that mass can be transported from $x \in \mathbb{T}$ to $y \in \mathbb{T}$ in two ways, either clock- or counter-clockwise. For example, consider two Dirac distributions, $\mu_1 = \delta_{1/5}$ and $\mu_2 = \delta_{4/5}$. Defining U_1 and U_2 by (4) with $a = 0$, it is immediately checked that $G_1(\omega) = 1/5$ and $G_2(\omega) = 4/5$ for all $\omega \in (0, M)$; thus the integral in formula (5) evaluates to $9/25$. On the other hand, $\pi := \delta_{1/5} \otimes \delta_{4/5}$ obviously belongs to $\Pi(\mu_1, \mu_2)$, so the infimum in (3) does not exceed $4/25$.

Fortunately, the point-symmetry of our densities u allows to extend (5) also to the situation at hand.

Lemma 2.2. *Assume that densities u_1 and u_2 in $\mathcal{P}_M(\mathbb{T})$ are even, i.e., $u_i(x) = u_i(1 - x)$ for $x \in (0, 1)$. Then formula* (5) *continues to hold, where the functions $G_i : [0, M] \to [0, 1]$ are the inverses of*

$$U_i(x) = \int_0^x u_i(y) \, dx. \tag{6}$$

The proof of this lemma has been deferred to Appendix B.

2.2. **The semi-discrete Euler scheme.** The next step is to introduce the (semi-discrete) minimizing movement scheme. In order to clarify the underlying idea, consider for the moment a finite-dimensional gradient flow

$$\dot{x} = -\nabla\phi(x) \tag{7}$$

for $x(t) \in \mathbb{R}^d$, with a smooth potential $\phi : \mathbb{R}^d \to \mathbb{R}$. We discretize equation (7) by the following inductive scheme of step size $\tau > 0$:

$$x^{n+1} := \operatorname{argmin}_{x \in \mathbb{R}^d} \Phi^\tau(x^n; x), \quad \Phi^\tau(x^*; x) := \frac{1}{2\tau}\|x - x^*\|^2 + \phi(x). \tag{8}$$

In fact, since the minimizer x^{n+1} satisfies the condition of a critical point,

$$0 = \nabla_x \Phi^\tau(x^n; x^{n+1}) = \frac{1}{\tau}(x^{n+1} - x^n) + \nabla\phi(x^{n+1}), \tag{9}$$

formula (8) is nothing but the implicit Euler scheme for (7), i.e., x^{n+1} satisfies

$$\frac{1}{\tau}(x^{n+1}-x^n) = -\nabla\phi(x^{n+1}).$$

The same general idea is applicable to gradient flows in the L^2-Wasserstein distance: For the DLSS equation, the scheme (8) turns into

$$u^{n+1} = \operatorname{argmin}_{u\in\mathcal{P}_M(\mathbb{T})}\Phi^\tau(u^n;u), \quad \Phi^\tau(u^*;u) := \frac{1}{2\tau}\mathbf{W}[u^*,u]^2 + \mathcal{F}[u]. \tag{10}$$

In Appendix A, the connection between (10) and the semi-discretization of (1) is established on a formal level. Notice that the Euler-Lagrange equation for (10) is much more complicated than in (9) because of the non-linear nature of the Wasserstein distance.

2.3. **Lagrangian coordinates.** Subsequently, we shall *not* make use of the Euler-Lagrange equation (39) (given in Appendix A), but instead perform the inductive minimization (10) for u^n on a suitable (finite-dimensional) subset of $\mathcal{P}_M(\mathbb{T})$. Before introducing the spatial discretization, we rewrite $\Phi^\tau(u^*;u)$ in a more explicit manner, using Lagrangian coordinates, i.e. the inverse distribution functions G and G^* of u and u^*, respectively. By Lemma 2.2, this allows to calculate the Wasserstein distance between u and u^* efficiently.

For a consistent change of variables, we need to express the Fisher information $\mathcal{F}$ in terms of Lagrangian coordinates as well. By the formula for the differential of an inverse function, it follows with $\omega = U(x)$,

$$u(x) = \partial_x U(x) = \frac{1}{\partial_\omega G(\omega)}, \quad \partial_x u(x) = -\frac{\partial_{\omega\omega}G(\omega)}{\big(\partial_\omega G(\omega)\big)^2}\cdot\partial_x U(x) = -\frac{\partial_{\omega\omega}G(\omega)}{\big(\partial_\omega G(\omega)\big)^3}.$$

Performing a change of variables $x = G(\omega)$ under the first integral in (2), we obtain

$$\mathcal{F}[u] = \int_{\mathbb{T}} \frac{\big(\partial_{\omega\omega}G(\omega)\big)^2}{2\big(\partial_\omega G(\omega)\big)^5}\,dx = \frac{1}{2}\int_0^M \left(\frac{\partial_{\omega\omega}G(\omega)}{\big(\partial_\omega G(\omega)\big)^2}\right)^2 d\omega = \frac{1}{2}\int_0^M \Big(\frac{1}{\partial_\omega G(\omega)}\Big)_\omega^2\,d\omega.$$

Introducing the derivative function $g := \partial_\omega G : [0,M]\to\mathbb{R}_+$, it follows that

$$\mathcal{F}[u] = \frac{1}{2}\int_0^M \Big(\frac{1}{g(\omega)}\Big)_\omega^2\,d\omega. \tag{11}$$

In terms of g, the expression for the Wasserstein distance becomes more lengthy,

$$\begin{aligned}
\mathbf{W}[u,u^*]^2 &= \int_0^M \big(G(\omega)-G^*(\omega)\big)^2\,d\omega \\
&= \int_0^M \left(\int_0^\omega \big(g(\eta)-g^*(\eta)\big)\,d\eta\right)^2 d\omega \\
&= \iiint_{[0,M]^3} \mathbf{1}_{\omega\geq\max(\eta,\eta')}\big(g(\eta)-g^*(\eta)\big)\big(g(\eta')-g^*(\eta')\big)\,d\eta\,d\eta'\,d\omega \\
&= \iint_{[0,M]^2} \big(M-\max(\eta,\eta')\big)\big(g(\eta)-g^*(\eta)\big)\big(g(\eta')-g^*(\eta')\big)\,d\eta\,d\eta'.
\end{aligned}$$

Despite its complicated appearance, this expression is simply a quadratic form in $g-g^*$ and thus easy to implement in the numerical scheme.

We summarize our results so far.

Lemma 2.3. *Assume that the initial condition $u_0 : \mathbb{T} \to \mathbb{R}_+$ is even. Then the solution u^n to the implicit Euler scheme* (39) *is in one-to-one correspondence to the solution g^n obtained from the inductive scheme*

$$g^{n+1} = \operatorname{argmin}_g \Psi^\tau(g^n; g), \tag{12}$$

with initial condition $g^0(\omega) := 1/u \circ G(\omega)$. The argmin *in* (12) *runs over all measurable functions $g : [0, M] \to \mathbb{R}_+$, satisfying the mass constraint*

$$\int_0^M g(\omega)\, d\omega = 1, \tag{13}$$

while Ψ^τ is given by

$$\begin{aligned}\Psi^\tau(g^*; g) =& \frac{1}{2\tau} \iint_{[0,M]^2} \big(M - \max(\eta, \eta')\big)\big(g(\eta) - g^*(\eta)\big)\big(g(\eta') - g^*(\eta')\big)\, d\eta\, d\eta' \\ &+ \int_0^M \Big(\frac{1}{g}\Big)_\omega^2 d\omega.\end{aligned} \tag{14}$$

In fact, the functional values agree, $\Phi^\tau(u^n; u^{n+1}) = \Psi^\tau(g^n; g^{n+1})$, and the functions u^n and g^n are related by

$$u^n(x_\omega) = \frac{1}{g^n(\omega)}, \quad \textit{where} \quad x_\omega = \int_0^\omega g^n(\eta)\, d\eta. \tag{15}$$

2.4. **Spatial discretization.** We shall now approximate the infinite-dimensional variational problem (12) by a finite-dimensional one. Minimization in (12) is not performed over all functions $g : [0, M] \to \mathbb{R}_+$, but only over a finite-dimensional subclass $\mathcal{G}_M^n$, referred to as *ansatz space* in the following.

Let a mesh $\mathbf{\Omega}_n = \{\omega_0, \omega_1, \dots, \omega_n\} \subset [0, M]$ be given, containing $n+1$ points ω_k with $\omega_0 = 0$, $\omega_k < \omega_{k+1}$ and $\omega_n = M$, which satisfy the symmetry property that $\omega_{n-k} = M - \omega_k$. Denote by

$$\delta_k = \omega_k - \omega_{k-1}, \quad \Delta_k = \frac{\omega_{k+1} - \omega_{k-1}}{2} \tag{16}$$

single and the double gaps, respectively. In particular, $\Delta_0 = \Delta_n = \omega_1$. Now, introduce the hat functions $\phi_1 : \mathbb{T} \to \mathbb{R}$ to $\phi_{n-1} : \mathbb{T} \to \mathbb{R}$ by

$$\phi_k(\omega) = \begin{cases} \frac{\omega - \omega_{k-1}}{\omega_k - \omega_{k-1}} & \text{if } \omega \in [\omega_{k-1}, \omega_k], \\ \frac{\omega_{k+1} - \omega}{\omega_{k+1} - \omega_k} & \text{if } \omega \in [\omega_k, \omega_{k+1}], \\ 0 & \text{otherwise,} \end{cases} \tag{17}$$

while $\phi_n : \mathbb{T} \to \mathbb{R}$ is given by

$$\phi_n(\omega) = \begin{cases} \frac{\omega_1 - \omega}{\omega_1} & \text{if } \omega \in [0, \omega_1], \\ \frac{\omega - \omega_{n-1}}{M - \omega_{n-1}} & \text{if } \omega \in [\omega_{n-1}, M], \\ 0 & \text{otherwise.} \end{cases} \tag{18}$$

Define the ansatz space $\mathcal{G}_M^n$ as the set of positive, *piecewise linear* functions $g : [0, M] \to \mathbb{R}_+$ of the form

$$g(\omega) = \sum_{k=1}^n g_k \phi_k(\omega). \tag{19}$$

We call $\mathbf{g} := (g_1, \cdots, g_n) \in \mathbb{R}_+^n$ the associated weight vector. By definition of the ϕ_k, clearly $g(\omega_k) = g_k$, and also $g(0) = g(M) = g_n$. For notational convenience, we put $g_0 = g_n$.

Definition 2.4. For given mass $M > 0$ and grid $\mathbf{\Omega}_n \subset [0, M]$, define $\mathbb{G}_M^n \subset \mathbb{R}_+^n$ as the set of weight vectors $\mathbf{g}$ for which the associated linear interpolation g from (19) satisfies the mass constraint (13), i.e.

$$\sum_{k=1}^{n} \Delta_k g_k = 1. \tag{20}$$

Let us interpret $\mathbf{g} \in \mathbb{G}_M^n$ in terms of its corresponding density function.

Lemma 2.5. *Given a vector $\mathbf{g} \in \mathbb{G}_M^n$ and a grid $\mathbf{\Omega}_n \subset [0, M]$, denote by $g \in \mathcal{G}_M^n$ the piecewise linear interpolation* (19)*. Moreover, define x_0, $x_1, \dots, x_n \in \mathbb{T}$ by*

$$x_k := \int_0^{\omega_k} g(\omega)\, d\omega = \frac{1}{2} \sum_{j=1}^{k} \delta_j (g_j + g_{j-1}). \tag{21}$$

Then the (continuous) density function $u \in \mathcal{P}_M(\mathbb{T})$, defined piecewisely on the intervals $[x_{k-1}, x_k] \subset \mathbb{T}$ by

$$u(x) = \Big(\frac{\delta_k (g_k + g_{k-1})}{2} \Big)^{1/2} \cdot \big(g_k^2 (x - x_{k-1}) + g_{k-1}^2 (x_k - x) \big)^{-1/2}, \tag{22}$$

is related to g by (15)*, i.e.*

$$u(x_\omega) = \frac{1}{g(\omega)}, \quad \text{where} \quad x_\omega = \int_0^{\omega} g(\eta)\, d\eta.$$

In particular, $g_k = 1/u(x_k)$ for $k = 0, 1, \dots, n$.

The proof of this lemma is found in Appendix B.

2.5. **Discretization of the functionals.** The next step is to obtain a more explicit representation of the functional Ψ^τ defined in (14) on the set $\mathcal{G}_M^n$. More precisely, we want to obtain an expression for Ψ^τ only depending on the weight vectors $\mathbf{g} = (g_1, \dots, g_n) \in \mathbb{G}_M^n$, and $\mathbf{g}^* = (g_1^*, \dots, g_n^*) \in \mathbb{G}_M^n$.

2.5.1. *Wasserstein distance.* The part of Ψ^τ corresponding to the Wasserstein distance is obviously a quadratic form in $\mathbf{g} - \mathbf{g}^*$,

$$\mathbf{W}[u, u^*]^2 = \sum_{j,k=1}^{n} a_{j,k} (g_j - g_j^*)(g_k - g_k^*), \tag{23}$$

where the $a_{j,k} = a_{k,j}$ are the elements of a symmetric $n \times n$-matrix A. In Appendix B, we prove the following lemma.

Lemma 2.6. *The matrix A implicitly defined in* (23) *is symmetric positive-definite, and its elements $a_{j,k}$ with $1 \le j \le k < n$ are given by*

$$\begin{aligned}
a_{j,j} &= \Delta_j^2 (M - \sigma_j) - \frac{\Delta_j}{60} (12 \Delta_j^2 + \delta_j^2 + \delta_{j+1}^2), \\
a_{j,j+1} &= \Delta_j \Delta_{j+1} (M - \sigma_{j+1}) - \frac{1}{120} \delta_{j+1}^3 && \text{for } j + 1 < n, \\
a_{j,k} &= \Delta_j \Delta_k (M - \sigma_k) && \text{if } k \ge j + 2,
\end{aligned}$$

and

$$\begin{aligned}
a_{1,n} &= \frac{1}{2}\Delta_1\Delta_n\Big(M - \frac{1}{3}\omega_2\Big) - \frac{1}{120}\Delta_n^3,\\
a_{j,n} &= \frac{1}{2}\Delta_j\Delta_n\Big(M - \sigma_j + \frac{1}{3}\Delta_n\Big) \qquad\qquad \text{if } 2 \le j \le n-2,\\
a_{n-1,n} &= \frac{1}{2}\Delta_{n-1}\Delta_n\Big(M - \frac{\omega_{n-2}+2\omega_{n-1}}{3}\Big) - \frac{1}{120}\Delta_n^3,\\
a_{n,n} &= \frac{1}{4}M\Delta_n^2 + \frac{1}{10}\Delta_n^3,
\end{aligned}$$

with $\sigma_k = \frac{\omega_{k-1}+\omega_k+\omega_{k+1}}{3}$ *for* $k < n$.

2.5.2. *Fisher information.* A function $g \in \mathcal{G}_M^n$ with weight vector $\mathbf{g} \in \mathbb{G}_M^n$ satisfies

$$g(\omega) = \frac{g_k(\omega - \omega_{k-1}) + g_{k-1}(\omega_k - \omega)}{\omega_k - \omega_{k-1}} \quad \text{for} \quad \omega \in [\omega_{k-1}, \omega_k].$$

By the representation (11) of the Fisher information,

$$\mathcal{F}[u] = \frac{1}{2}\int_0^M \Big(\frac{1}{g}\Big)_\omega^2 d\omega = \frac{1}{2}\sum_{k=1}^n \mathbf{F}_k[\mathbf{g}], \tag{24}$$

where we define

$$\begin{aligned}
\mathbf{F}_k[\mathbf{g}] &:= \delta_k^2(g_k - g_{k-1})^2 \int_{\omega_{k-1}}^{\omega_k} \frac{d\omega}{(g_k(\omega - \omega_{k-1}) + g_{k-1}(\omega_k - \omega))^4}\\
&= \frac{\delta_k^{-1}}{3}\Big(\frac{1}{g_k} - \frac{1}{g_{k-1}}\Big)^2\Big(1 + \frac{g_{k-1}}{g_k} + \frac{g_k}{g_{k-1}}\Big),
\end{aligned}$$

with the usual convention $g_0 = g_n$. For later reference, we calculate the (non-trivial) derivatives of $\mathbf{F}$,

$$\frac{\partial \mathbf{F}_k[\mathbf{g}]}{\partial g_k} = \delta_k^{-1}\frac{2 + (\frac{g_k}{g_{k-1}})^3 - 3\frac{g_{k-1}}{g_k}}{3g_k^3}, \quad \frac{\partial \mathbf{F}_k[\mathbf{g}]}{\partial g_{k-1}} = \delta_k^{-1}\frac{2 + (\frac{g_{k-1}}{g_k})^3 - 3\frac{g_k}{g_{k-1}}}{3g_{k-1}^3}.$$

The derivatives of second order are

$$\frac{\partial^2 \mathbf{F}_k[\mathbf{g}]}{\partial g_k \partial g_k} = 2\delta_k^{-1}\frac{2g_{k-1} - g_k}{g_k^5}, \quad \frac{\partial^2 \mathbf{F}_k[\mathbf{g}]}{\partial g_{k-1}\partial g_{k-1}} = 2\delta_k^{-1}\frac{2g_k - g_{k-1}}{g_{k-1}^5},$$
$$\frac{\partial^2 \mathbf{F}_k[\mathbf{g}]}{\partial g_k \partial g_{k-1}} = -\delta_k^{-1}\Big(\frac{1}{g_{k-1}^4} + \frac{1}{g_k^4}\Big).$$

2.6. **Minimization procedure.** To define our numerical scheme, we replace the original (infinite-dimensional) variational problem (12) by the following finite-dimensional one:

$$\mathbf{g}^{n+1} := \operatorname{argmin}_{\mathbf{g}\in\mathbb{G}_M^n} \Psi_n^\tau(\mathbf{g}^*; \mathbf{g}), \tag{25}$$

where the functional $\Psi_n^\tau : \mathbb{G}_M^n \times \mathbb{G}_M^n \to \mathbb{R}$ has the representation

$$\Psi^\tau(\mathbf{g}^*; \mathbf{g}) = \frac{1}{2\tau}\sum_{j,k=1}^n a_{j,k}(g_j - g_j^*)(g_k - g_k^*) + \frac{1}{2}\sum_{k=1}^n \mathbf{F}_k[\mathbf{g}]. \tag{26}$$

Notice that Ψ_n^τ above and the Ψ^τ originally introduced in (14) are related by $\Psi_n^\tau(\mathbf{g}^*; \mathbf{g}) = \Psi^\tau(g^*; g)$, with piecewise linear functions g^*, $g \in \mathcal{G}_M^n$ defined from $\mathbf{g}^*$, $\mathbf{g}$ by (19).

2.6.1. *Existence of a discrete minimizer.* Below, we derive a priori estimates that yield the existence of a global minimizer for Ψ_n^τ.

Theorem 2.7. *For given $\mathbf{g}^* \in \mathbb{G}_M^n$, there exists a global minimizer $\bar{\mathbf{g}} \in \mathbb{G}_M^n$ of $\Psi_n^\tau(\mathbf{g}^*, \mathbf{g})$ defined in (26). In particular, $\bar{\mathbf{g}}$ corresponds to a strictly positive density function on $\mathbb{T}$ of total mass M. Moreover, there is a finite constant C that only depends on properties of the mesh $\mathbf{\Omega}_n$, such that*

$$\frac{1}{M + \sqrt{2MF}} \le \bar{g}_k \le \max_j g_j^* + C\sqrt{\tau F} \qquad (k = 1, \dots, n), \tag{27}$$

where $F := \frac{1}{2}\sum_{j=1}^n \mathbf{F}_j[\mathbf{g}^*]$ is the Fisher information associated to $\mathbf{g}^*$.

Proof. Denote by $\bar{g} : [0, M] \to \mathbb{R}_+$ the piecewise linear interpolation for $\bar{\mathbf{g}}$ on $\mathbf{\Omega}_n$, and by $\bar{G} : [0, M] \to [0, 1]$ the associated inverse distribution function, i.e., $\bar{G}' = \bar{g}$; introduce g^* and G^*, respectively. Observe that any minimizer $\bar{\mathbf{g}} \in \mathbb{G}_M^n$ of $\Psi_n^\tau(\mathbf{g}^*, \mathbf{g})$ necessarily satisfies

$$\frac{1}{2\tau}\int_0^M \big|\bar{G}(\omega) - G^*(\omega)\big|^2 \, d\omega + \int_0^M \Big(\frac{1}{\bar{g}}\Big)_\omega^2 d\omega = \Psi^\tau(g^*, \bar{g}) \le \Psi^\tau(g^*, g^*) = F. \tag{28}$$

We begin by proving the lower a priori bound in (27). Since $1 = \bar{G}(M) = \int_0^M \bar{g}(\omega)d\omega$, it follows that that $\bar{g}(\bar{\omega}) \ge 1/M$ for some $\bar{\omega} \in [0, M]$. For arbitrary $\omega \in [0, M]$,

$$\begin{aligned}\frac{1}{\bar{g}(\omega)} &= \frac{1}{\bar{g}(\bar{\omega})} + \int_{\bar{\omega}}^\omega \Big(\frac{1}{\bar{g}(\eta)}\Big)_\eta \, d\eta \\ &\le \frac{1}{\bar{g}(\bar{\omega})} + \big|\omega - \bar{\omega}\big|^{1/2}\bigg(\int_{\bar{\omega}}^\omega \Big(\frac{1}{\bar{g}(\eta)}\Big)_\eta^2 d\eta\bigg)^{1/2} \le M + \sqrt{2M\mathcal{F}[\bar{g}]}.\end{aligned}$$

In particular, this estimate holds at the grid points $\omega_k \in \mathbf{\Omega}_n \subset [0, M]$. Now apply (28) to conclude the lower bound in (27).

The derivation of the upper bound is more technical, and we will only sketch the argument. Consider the class $S \subset L^2[0, M]$ of quadratic splines $s : [0, M] \to \mathbb{R}$ w.r.t. the grid $\mathbf{\Omega}_n \subset [0, M]$ which also satisfy homogeneous Dirichlet conditions $s(0) = s(M) = 0$. By the finite mesh size of $\mathbf{\Omega}_n$, an inverse Poincaré inequality holds on S. Denote by $c > 0$ a constant such that

$$\int_0^M s(\omega)^2 \, d\omega \ge c\int_0^M s'(\omega)^2 \, d\omega \qquad \text{for all } s \in S. \tag{29}$$

Now, notice that $\bar{G} - G^* \in S$. By (29) and (28),

$$F \ge \frac{1}{2\tau}\int_0^M \big|\bar{G}(\omega) - G^*(\omega)\big|^2 \, d\omega \ge \frac{c}{2\tau}\int_0^M \big|\bar{g}(\omega) - g^*(\omega)\big|^2 \, d\omega \ge \frac{c\Delta_k}{2\tau}\big|\bar{g}_k - g_k^*\big|^2,$$

where Δ_k is given by (16). From here, the upper bound in (27) follows immediately.

Finally, we conclude the *existence* of a minimizer. To this end, observe that the a priori estimates define a compact subset K of $\mathbb{R}^n$. It is immediately seen that $K \cap \mathbb{G}_M^n$ is closed, hence also compact. Finally, this intersection is non-empty, since it contains $\mathbf{g}^*$. Indeed, $\mathbf{g}^* \in \mathbb{G}_M^n$ by hypothesis, the upper bound in (27) is trivially satisfied, and the lower bound is obtained by repeating the calculations above with $\mathbf{g}^*$ in place of $\bar{\mathbf{g}}$ (recall that we have only used the boundedness of the Fisher information by F). □

Remark 1. Uniqueness of the minimizer is not clear. In fact, the finite-dimensional problem (25) is not convex (with respect to $\mathbf{g} \in \mathbb{G}_M^n$), although the original minimization problem (10) is known [17] to be strictly convex (with respect to $u \in \mathcal{P}_M(\mathbb{T})$). Notice that the restriction of g to the linear set $\mathcal{G}_M^n$ corresponds to impose a complicated non-linear constraint on the admissible densities u in the original problem (10), which thus loses its convex character.

2.6.2. *Fully discrete Euler-Lagrange equations.* By definition of $\mathbb{G}_M^n$, the sought minimizer $\mathbf{g}^{n+1}$ in (25) is subject to the mass constraint (20). Instead of working directly on the set $\mathbb{G}_M^n$, it is more convenient to consider (25) as a constrained minimization problem for $\mathbf{g}$ on the larger set $\mathbb{R}^n$. Accordingly, we introduce a Lagrange multiplier λ and the associated Lagrangian functional

$$L^\tau(\mathbf{g}^*; \mathbf{g}, \lambda) := \Psi^\tau(\mathbf{g}^*; \mathbf{g}) - \lambda\Big(1 - \sum_{k=1}^n \Delta_k g_k\Big).$$

By the classical theory of variations, solutions $\bar{\mathbf{g}}$ of (25) give rise to critical points $(\bar{\mathbf{g}}, \bar{\lambda})$ of L^τ.

A critical point $(\mathbf{g}, \lambda)$ of L^τ necessarily satisfies the n conditions

$$\begin{aligned}
0 = \mathbf{G}_k &:= \frac{\partial L^\tau}{\partial g_k} \\
&= \frac{1}{\tau}\sum_{j=1}^n a_{j,k}(g_j - g_j^*) + \frac{1}{2}\frac{\partial \mathbf{F}_k[\mathbf{g}]}{\partial g_k} + \frac{1}{2}\frac{\partial \mathbf{F}_{k+1}[\mathbf{g}]}{\partial g_k} - \lambda\Delta_k \\
&= \frac{1}{\tau}\sum_{j=1}^n a_{j,k}(g_j - g_j^*) - \lambda\Delta_k \\
&\quad + \frac{1}{6}\Big[\Big(\frac{1}{\delta_k g_{k-1}^3} + \frac{1}{\delta_{k+1} g_{k+1}^3}\Big) + \frac{2}{g_k^3}(\delta_k^{-1} + \delta_{k+1}^{-1}) - \frac{3}{g_k^4}(\delta_k^{-1} g_{k-1} + \delta_{k+1}^{-1} g_{k+1})\Big]
\end{aligned}$$

for $k = 1, \dots, n$, as well as the constraint

$$0 = \mathbf{G}_{n+1} := \frac{\partial L^\tau}{\partial \lambda} = 1 - \sum_{k=1}^n \Delta_k g_k.$$

Notice that $\mathbf{G} \in \mathbb{R}^{n+1}$ above denotes the gradient of $L^\tau(\mathbf{g}^*; \mathbf{g}, \lambda)$ with respect to $(\mathbf{g}, \lambda) \in \mathbb{R}^{n+1}$. To approximate a critical point numerically, we apply Newton's method to the first-order optimality condition

$$0 = \mathbf{G}[g^*; g, \lambda].$$

This leads to a sequential quadratic programming method, since at every iteration step a quadratic subproblem is solved. The entries of the Hessian $(n+1)\times(n+1)$-matrix $\mathbf{H}[g^*; g, \lambda]$ of $L^\tau[g^*; g, \lambda]$ are

$$\begin{aligned}
\mathbf{H}_{k,k} &= \frac{1}{\tau}a_{k,k} + \frac{1}{g_k^5}\Big(2\delta_k^{-1} g_{k-1} + 2\delta_{k+1}^{-1} g_{k+1} \\
&\qquad - (\delta_k^{-1} + \delta_{k+1}^{-1}) g_k\Big) && (1 \le k \le n), \\
\mathbf{H}_{k-1,k} = \mathbf{H}_{k,k-1} &= \frac{1}{\tau}a_{k-1,k} - \frac{1}{2\delta_k}\Big(\frac{1}{g_k^4} + \frac{1}{g_{k-1}^4}\Big) && (1 \le k \le n), \\
\mathbf{H}_{k,j} = \mathbf{H}_{j,k} &= \frac{1}{\tau}a_{k,j} && (1 \le k\,, k+2 \le j \le n),
\end{aligned}$$

while $\mathbf{H}_{k,n+1} = \mathbf{H}_{n+1,k} = -\Delta_k$ for $1 \le k \le n$, and $\mathbf{H}_{n+1,n+1} = 0$.

2.7. **Implementation.** With all the formulas at hand, the implementation is now straight-forward. Given $\mathbf{g}^n$, the solution at the nth time step, let $\mathbf{g}^{(0)} := \mathbf{g}^n$ and $\lambda^{(0)} := 0$. Now define inductively

$$\mathbf{g}^{(s+1)} := \mathbf{g}^{(s)} + \delta\mathbf{g}^{(s+1)}, \quad \lambda^{(s+1)} := \lambda^{(s)} + \delta\lambda^{(s+1)},$$

where $(\delta\mathbf{g}^{(s+1)}, \delta\lambda^{(s+1)})$ is the solution to the linear system

$$\mathbf{H}[\mathbf{g}^n; \mathbf{g}^{(s)}, \lambda^{(s)}](\delta\mathbf{g}^{(s+1)}, \delta\lambda^{(s+1)}) = -\mathbf{G}[\mathbf{g}^n; \mathbf{g}^{(s)}, \lambda^{(s)}].$$

As soon as the norm of $(\delta\mathbf{g}^{(s+1)}, \delta\lambda^{(s+1)})$ drops below a certain threshold, define $\mathbf{g}^{n+1} := \mathbf{g}^{(s+1)}$, $\lambda^{n+1} := \lambda^{(s+1)}$ as values in the $n+1$st time step.

2.8. **Choice of the initial condition.** We recall that the Wasserstein gradient flow scheme is written in Lagrangian coordinates, with the consequence that the grid points $\omega_k \in \mathbf{\Omega}_n \subset [0, M]$ are fixed, whereas the corresponding sites

$$x_k^n = G^n(\omega_k) = \frac{1}{2}\sum_{\ell=1}^{k} \delta_\ell (g_\ell^n + g_{\ell-1}^n) \tag{30}$$

on $\mathbb{T}$ are moving as the solution evolves. This leads to a certain ambiguity in the prescription of the initial datum, since not only the values g_k^0 of the vector $\mathbf{g}^0 \in \mathbb{G}_M^n$, but also the lattice $\mathbf{\Omega}_n$ needs to be suitably defined. The seemingly canonical choice of an equidistant grid $\omega_k = Mk/n$ is generally *not* advisable since it leads — by the correspondence (21) — to a poor resolution of x-space exactly in the interesting regions where the initial density u_0 is small.

For our numerical experiments with the Wasserstein gradient flow scheme, we mimic the direct sampling typically used for discretizations by finite differences, i.e. we set

$$g_k^0 := \frac{1}{u_0(k/n)}, \quad \text{and inductively} \quad \omega_{k+1} := \omega_k + \frac{2}{n}(g_k^0 + g_{k+1}^0)^{-1}. \tag{31}$$

It follows that for $k = 1, 2, \dots, n$

$$x_k^0 = \frac{1}{2}\sum_{\ell=1}^{k} \delta_k (g_k^0 + g_{k-1}^0) = \sum_{\ell=1}^{k} \frac{1}{n} = \frac{k}{n}.$$

In other words, the grid points on $\mathbb{T}$ are equidistantly distributed at $t = 0$, with respective function values $1/g_k^0 = u_0(x_k^0)$. In particular, the mass constraint is fulfilled in the sense that $x_n^0 = 1$. However, we need to emphasize that $\omega_n \neq M$ in general, so the reconstructed density by means of (22) does not have the same mass as the initial condition u_0.

We would like to mention that — depending on the specific application — other choices of the initial conditions may be more appropriate that (31). For instance, one may choose to discretize the relation between x-space and ω-space at time $t = 0$ exactly, i.e. to assure

$$\omega_k = U^0(x_k), \qquad k = 1, 2, \dots, n. \tag{32}$$

This implies in particular that $\omega_n = U^0(1) = M$, i.e. the mass constraint holds *precisely* with the mass M from u^0.

The most direct attempt to realize (32) would be to set $x_k = k/n$ and $\omega_k = U^0(x_k)$, implying that the vector $\mathbf{g}^0 \in \mathbb{G}_M^n$ satisfies

$$\frac{k}{n} = \frac{1}{2}\sum_{\ell=1}^{k} \delta_\ell (g_\ell^0 + g_{\ell-1}^0). \tag{33}$$

The inverse distribution function G^0 associated to this $\mathbf{g}^0$ would then be a genuine interpolant of the inverse of U^0 with respect to the grid $\mathbf{\Omega}_n$. However, the solution of (33) poses severe difficulties. For odd n, the solution $\mathbf{g}^0$ is uniquely determined, but its components might become negative, no matter how refined the grid $\mathbf{\Omega}_n$ is. For even n, the solution is not unique, and there appears to exist no general procedure to select a particular $\mathbf{g}^0$, which has all components positive and constitutes a good approximation of $1/u_0$ at the x_k.

A better choice, which is also easy to implement, is the following.

Lemma 2.8. *Assume that the initial datum u_0 is C^2-smooth and strictly positive, i.e. $\inf_{\mathbb{T}} u_0 > 0$. For given $n \geq 2$, define the grid $\mathbf{\Omega}_n$ by $\omega_k = U^0(k/n)$ for $k = 0, 1, 2, \dots, n$. Further, define $\mathbf{g}^0 \in \mathbb{R}_+^n$ by $g_k^0 = 1/(n\Delta_k)$, where Δ_k is given by (16). Then $\mathbf{g} \in \mathbb{G}_M^n$, i.e., the mass constraint (20) is satisfied, and moreover,*

$$x_k^0 = \frac{k}{n} + \mathcal{O}\big(n^{-2}\big) \quad \text{and} \quad g(\omega_k) = \frac{1}{u(k/n)} + \mathcal{O}\big(n^{-2}\big) \tag{34}$$

uniformly in $k = 0, 1, \dots, n$ as $n \to \infty$.

Proof. Since u_0 is positive, all Δ_k are positive numbers. Thus, the weights $\mathbf{g}^0$ and the corresponding linear interpolant $g^0 \in \mathcal{G}_M^n$ are well-defined and positive. Now observe that

$$x_k = \int_0^{\omega_k} g^0(\eta)\, d\eta = \frac{1}{2n}\sum_{j=1}^{k} \delta_j\big(\Delta_j^{-1} + \Delta_{j-1}^{-1}\big) = \frac{1}{n}\Big(\frac{\delta_1}{2\Delta_0} + \sum_{j=1}^{k-1} \frac{\delta_j + \delta_{j+1}}{2\Delta_j} + \frac{\delta_k}{2\Delta_k}\Big). \tag{35}$$

Using that $2\Delta_j = \delta_j + \delta_{j+1}$ for $1 \leq j < n$ by definition, and $\Delta_0 = \delta_1$ by symmetry,

$$x_k = \frac{k}{n} + n^{-2} R_1, \quad \text{with } |R_1| = \Big|\frac{n(\delta_{k+1} - \delta_k)}{2(\delta_{k+1} + \delta_k)}\Big| \leq \frac{1}{4}\frac{\sup_{\mathbb{T}} |(u_0)_x|}{\inf \mathbb{T} u_0}.$$

The last inequality follows by a Taylor expansion of $u_0 \in C^2$ around $x = n/k$. This proves the first claim in (34). Moreover, with $2\Delta_0 = 2\Delta_n = \delta_1 + \delta_n$, equation (35) for $k = n$ yields the mass constraint (20).

Another application of Taylor's theorem implies

$$\begin{aligned}\Delta_k &= \frac{1}{2}\Big(U^0\Big(\frac{k+1}{n}\Big) - U^0\Big(\frac{k-1}{n}\Big)\Big) \\ &= n^{-1} u_0(k/n) + n^{-3} R, \quad \text{with } |R| \leq \frac{1}{3}\sup_{\mathbb{T}} |(u_0)_{xx}|.\end{aligned}$$

Moreover,

$$g_k^0 = \frac{1}{n\Delta_k} = \frac{1}{u_0(k/n)}\Big(1 + \frac{R}{n^2 u_0(k/n)}\Big)^{-1} = \frac{1}{u_0(k/n)} + n^{-2} r, \quad \text{with } |r| \leq \frac{2R}{\inf_{\mathbb{T}} u_0}.$$

This shows the second claim in (34). □

3. **Fully implicit and semi-implicit finite difference schemes.** For nonlinear heat transfer problems, semi-implicit finite difference discretizations have been proven to deliver a useful and stable compromise between accuracy and computation time. Since the numerical treatment of fourth order parabolic problems as (1) is not well-developed, it is not clear if this assertion carries over to fourth order problems. In this section we consider a fully implicit and two different, semi-implicit finite difference discretizations of (1), one classical centered backward Euler scheme (BTCS), and a higher order Mehrstellenformel (MSF).

For the finite difference discretization consider N uniformly distributed grid points $x_i = ih$ in $[0,1]$, with space step $h = 1/N$ and constant time step k.

Let Δ_2 denote the standard central difference operator for approximating the second derivative. Let U_i^n denote the approximate solution of (1) in x_i at the time $t_n = nk$ and let $U^n = (U_i^n)_{i=0}^N$. We say a scheme is of order (m,n) if it is formally consistent of order m in time and of order n in space or, more precisely, the truncation error is of order $\mathcal{O}(k^m + h^n)$.

3.1. **Fully implicit scheme (FIMP).** For the computations, we discretize (1) in time by the implicit Euler scheme and in space by central finite differences:

$$\frac{U^{n+1} - U^n}{k} = -\Delta_2\Big(U^{n+1}\Delta_2\big(\ln U^{n+1}\big)\Big).$$

The discrete nonlinear system is solved by the Newton method. The initial guess is chosen as the solution of the previous time level.

3.2. **Backward time central space (BTCS).** Semi-implicit discretization of (1) is given by

$$\frac{U^{n+1} - U^n}{k} = -\Delta_2\Big(U^{n+1}\Delta_2\big(\ln U^{n}\big)\Big).$$

The scheme is consistent of order (1,2). A different, semi-implicit discretization was proposed in [12]. There, the logarithm term was evaluated implicitly and the linear term explicitly, which also leads to a scheme of order (1,2).

3.3. **Mehrstellenformel (MSF).** To obtain a higher order scheme, one can employ a differencing stencil of higher order, e.g. one can approximate the second derivative by the centered five point stencil

$$(u_{xx})_i = \frac{-U_{i+2} + 16U_{i+1} - 30U_i + 16U_{i-1} - U_{i-2}}{12h^2} + \mathcal{O}(h^4).$$

Denote this discretization by the operator $\widetilde{\Delta}_2$. A scheme that is consistent of order (1,4) is then given by

$$\frac{U^{n+1} - U^n}{k} = -\widetilde{\Delta}_2\Big(U^{n+1}\widetilde{\Delta}_2\big(\ln U^{n}\big)\Big),$$

To increase the time resolution along with the better space discretization, one can employ a time discretization of higher order, e.g. one can use the backward differencing formula of order 2 (BDF-2). This yields

$$U^{n+1} - \frac{4}{3}U^n + \frac{1}{3}U^{n-1} = -\frac{2}{3}k\widetilde{\Delta}_2\Big(U^{n+1}\widetilde{\Delta}_2\big(\ln U^{n}\big)\Big),$$

which is consistent of order (2,4). For the first step in time, one uses an (semi-)implicit Euler step.

4. **Numerical experiments.**

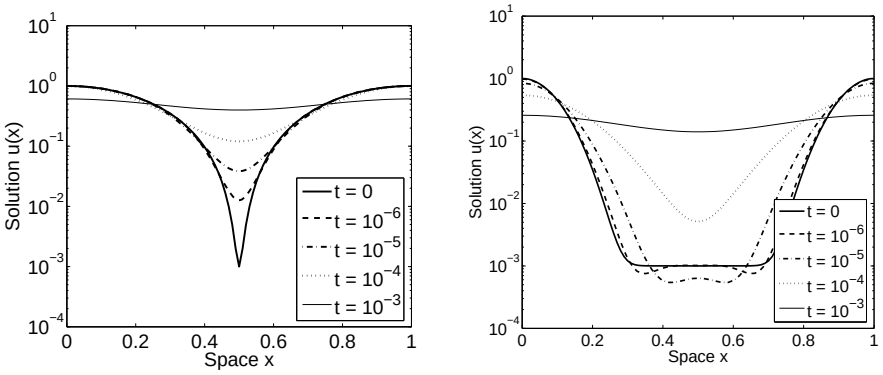

FIGURE 1. Time evolution of the solution $u(x,t)$ with initial condition (36) using the Wasserstein gradient flow scheme: left: $m = 1$, right: $m = 8$

4.1. **Calculation of a transient solution.** First, we illustrate the time evolution of typical solutions to (1). We choose an initial datum similar to that used in [7, 26],

$$u_0(x) = \epsilon + \cos^{2m}(\pi x), \quad x \in (0,1), \tag{36}$$

where m is a positive integer and $\epsilon > 0$ is a small parameter. We choose the parameter $\epsilon = 10^{-3}$ and consider the cases $m = 1$ and $m = 8$. As terminal time we have chosen $T = 5 \cdot 10^{-6}$. The time step size is $k = 10^{-8}$ and the number of grid points is $N = 100$. Figure 1 shows the evolution of the numerical solution of the DLSS equation using the Wasserstein gradient flow scheme (WGF) for $m = 1$ and $m = 8$, respectively. The solution for $m = 1$ increases rapidly in the interior of the domain until it reaches the constant steady state. On the other hand, the solution for $m = 8$ starts with one higher-order minimum, bifurcates into two local minima and reduces to one extremum again. In particular, it violates the minimum principle since the initial condition provides no lower bound for the solution.

4.1.1. *Particle trajectories.* Apart from the figures of the density u shown above, the Wasserstein gradient flow scheme provides a different view on the solution u: Since the DLSS equation is a gradient flow in the Wasserstein metric, the associated evolution can (and should!) be interpreted as a process of redistribution of particles with spatio-temporal density $u(t;x)$ on the domain $\mathbb{T}$ under the influence of a non-linear particle interaction (described by $\mathcal{F}$). It is thus natural to trace the trajectories of "test particles" on $\mathbb{T}$ in order to obtain a picture of the way in which the initial density u_0 is deformed in time.

In Figure 2, the trajectories of one hundred such "test particles" are displayed for the solutions to (1) for initial condition (36) with $m = 8$ and $\epsilon = 10^{-1}$ (left) or $\epsilon = 10^{-3}$ (right), respectively. The test particles have been distributed uniformly on $\mathbb{T}$ at time $t = 0$ for better visualization; we emphasize that the density of trajectories is generally no indication for the spatial density u of the solution.

4.1.2. *Decay of the Fisher information.* We compute the Fisher information (2) of the numerical solutions obtained using the fully implicit finite difference scheme and the Wasserstein gradient flow scheme using different numbers of mesh points. The definite integral in formula (2) is evaluated using the mid-point rule. We use

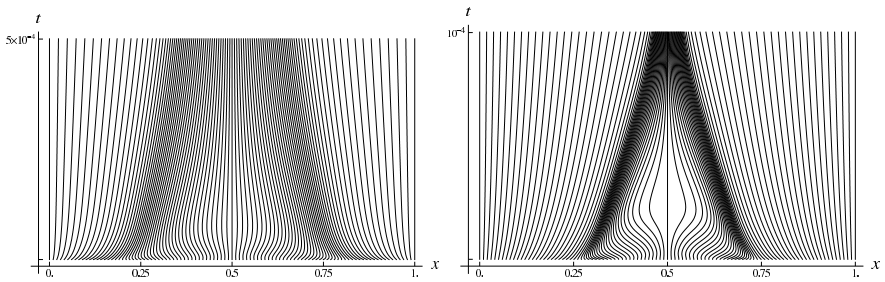

FIGURE 2. Particle trajectories in the Wasserstein gradient flow scheme, corresponding to solutions with $m = 8$ and $\epsilon = 10^{-1}$ (left) or $\epsilon = 10^{-3}$ (right).

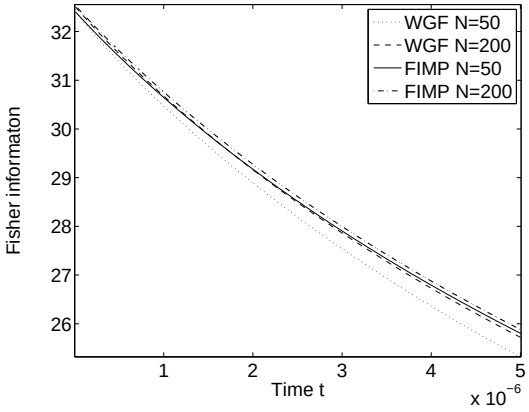

FIGURE 3. Decay of the Fisher information over time for the fully implicit finite difference scheme (FIMP) and the Wasserstein gradient flow scheme (WGF) using different numbers of grid points.

initial condition (36) and choose the parameters $\epsilon = 10^{-3}$, $m = 8$ and $T = 5 \cdot 10^{-6}$. Figure 3 shows the decay of the Fisher information $\mathcal{F}[u(t)]$ over time. For both schemes, the Fisher information decreases monotonically in time. As the number of mesh points increases both schemes necessarily approach the behavior of the true solution. Already for $N = 200$ the Fisher information resulting from both schemes are very close to each other.

4.2. **Numerical convergence.** To study the numerical convergence of the different schemes we are interested in the l_2-error ε_2 of the numerical solutions,

$$\varepsilon_2 = \left(h \sum_{i=0}^{N} \left| U_i^{k_T} - u(x_i, T) \right|^2 \right)^{\frac{1}{2}}$$

where $k_T \in \mathbb{N}$ is such that $T = k_T k$. We expect the pointwise error to behave like $\varepsilon_2 \approx Ch^s$ asymptotically in the limit of vanishing mesh size $h \downarrow 0$, where $s > 0$ is the (numerical) order of the scheme. The constant s is conveniently read off from

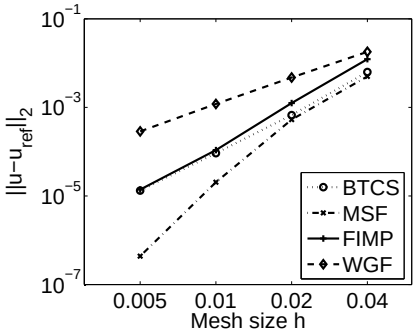

FIGURE 4. Numerical Convergence: l_2-error ε_2 vs. $h = 1/N$

the slope in the doubly logarithmic plot of ε_2 against h, since

$$\ln \varepsilon_2 \approx \ln C + s \ln h.$$

We use a solution on a fine spatial grid as reference solution. Figure 4 shows the convergence rates in l_2-norm for the different schemes. We use $T = 5 \cdot 10^{-6}$ and the constant parabolic mesh ratio $k/h^4 = 0.1$. Table 1 summarizes the average numerical convergence rates obtained using linear regression.

	s
WGF	1.99
FIMP	3.29
BTCS	2.95
MSF	4.51

TABLE 1. Numerically obtained convergence rates

Since equation (1) is of fourth order, the parabolic mesh ratio is given by $\alpha = k/h^4$. Hence time step restrictions have a severe impact on the practicality of the numerical schemes. The semi-implicit finite difference schemes are easy to implement, and computationally cheap, since at each step in time only a linear system has to be solved. However, they are not unconditionally stable, i.e. to have stability of the scheme the parabolic mesh ratio has to be chosen sufficiently small. In our experiments α could not be chosen much larger than 0.1, which renders the semi-implicit schemes impractical for fine grids. The fully implicit finite difference method and the Wasserstein gradient flow method are both implicit formulations and expected to be unconditionally stable. This is confirmed by our numerical experiments. Figure 5 shows the error for the fully implicit finite difference scheme and the Wasserstein gradient flow scheme using the above data and initial condition (36) with $m = 8$. We use a fixed number of grid points $N = 100$ and vary the time step size k. For both schemes one obtains stable solutions for all time step sizes. Interestingly, while the error deteriorates with increasing k for the fully implicit finite difference scheme, it remains on the same (although generally higher) level for the Wasserstein gradient flow scheme. This is commendable, since the new scheme

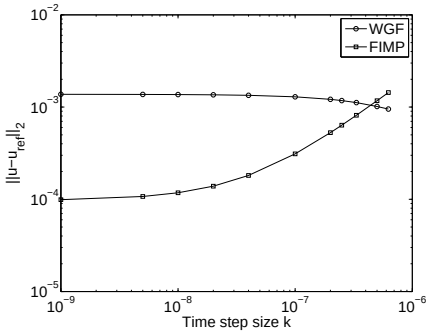

FIGURE 5. Comparison between the fully implicit finite difference scheme (FIMP) and the Wasserstein gradient flow (WGF) scheme: error vs. time step size using $N = 100$ grid points and initial condition (36) with $m = 8$.

seems to be a promising step toward an efficient solution of nonlinear diffusion equations. On the other hand it seems to suggest that the error using the Wasserstein gradient flow scheme in the current form is dominated by the spatial error. This effect and an improved version of the method are currently under investigation by the authors.

5. **Conclusions.** We have proposed a novel method for numerical integration of the fourth order DLSS equation, which easily adapts to a variety of other partial differential equations that constitute gradient flows in the L^2-Wasserstein metric. The implicit discrete scheme is based on successive solution of constrained minimization problems. We have proven the existence of the minimizers and have derived a priori estimates. Quadratic convergence of our method has been verified numerically. Moreover, our numerical experiments suggest that the scheme is unconditionally stable. In a forthcoming paper, the authors plan to study a multi-dimensional extension of the method.

Appendix A. **Formal derivation of the Euler-Lagrange equation.** We will briefly indicate how to deduce (at least formally) the Euler-Lagrange equations associated to the minimizing movement scheme (10).

To begin with, we recall the notion of the *push-forward* of a density u on $\mathbb{T}$ by a measurable map $T : \mathbb{T} \to \mathbb{T}$: the push-forward $T_{\#}u$ is again a density (of the same mass as u), which is uniquely characterized by the property

$$\int_{\mathbb{T}} \varphi(x)(T_{\#}u)(x)\,dx = \int_{\mathbb{T}} \varphi(T(x))u(x)\,dx$$

for all continuous test functions $\varphi : \mathbb{T} \to \mathbb{R}$. It is a fundamental fact from optimal transportation theory that the optimal transport between two densities (this is not true for measures with concentrations — in general, the optimal transport can only be described by a plan π_{opt}) is realized by the push-forward via a suitable map. In the case at hand, there exists some $T : \mathbb{T} \to \mathbb{T}$ with $u^n = T_{\#}u^{n+1}$ (notice the direction of transport!) such that

$$\mathbf{W}[u^n, u^{n+1}]^2 = \int_{\mathbb{T}} |T(x) - x|^2\, u^{n+1}(x)\,dx.$$

To obtain the Euler-Lagrange equations associated to (10), we follow [19]; see also [1, Proof of Thm.5.8]. The minimizing density u^{n+1} is perturbed in the form $u_\epsilon^{n+1} = S^\epsilon_{\#}u^{n+1}$, where the maps $S^\epsilon : \mathbb{T} \to \mathbb{T}$ constitute an ϵ-dependent family with $S^\epsilon(x) = x + \epsilon\xi(x)$, and $\xi : \mathbb{T} \to \mathbb{R}$ is a arbitrarily prescribed smooth function. By minimality of u^{n+1}, it is a critical point of $\Phi^\tau(u^n;\cdot)$, thus one expects

$$\frac{1}{2\tau}\mathbf{W}[u^n, u_\epsilon^{n+1}]^2 + \mathcal{F}[u_\epsilon^{n+1}] = \frac{1}{2\tau}\mathbf{W}[u^n, u^{n+1}]^2 + \mathcal{F}[u^{n+1}] + \mathcal{O}\big(\epsilon^2\big). \tag{37}$$

By definition of the Wasserstein distance as an infimum, it is clear that

$$\begin{aligned}\mathbf{W}[u^n, u_\epsilon^{n+1}]^2 &\le \int_{\mathbb{T}} \big|S^\epsilon(x) - T(x)\big|^2\, u^{n+1}(x)\,dx \\ &= \mathbf{W}[u^n, u^{n+1}]^2 + 2\epsilon\int_{\mathbb{T}} \xi(x)(x - T(x))u^{n+1}(x)\,dx + \mathcal{O}\big(\epsilon^2\big).\end{aligned}$$

In fact, since this inequality is true independent of the sign of ϵ, it shows that

$$\frac{1}{2\tau}\big(\mathbf{W}[u^n, u_\epsilon^{n+1}]^2 - \mathbf{W}[u^n, u^{n+1}]^2\big) = \frac{\epsilon}{\tau}\int_{\mathbb{T}} \xi(x)\cdot\big(T(x) - x\big)\,u^{n+1}(x)\,dx + \mathcal{O}\big(\epsilon^2\big).$$

On the other hand, a direct, but rather lengthy computation [17] reveals

$$\mathcal{F}[u_\epsilon^{n+1}] - \mathcal{F}[u^{n+1}] = \epsilon\int_{\mathbb{T}} \big(\xi_{xx}(u^{n+1})_x + 4\xi_x(\sqrt{u^{n+1}})_x^2\big)\,dx + \mathcal{O}\big(\epsilon^2\big).$$

Assuming that u^{n+1} is sufficiently regular, repeated integration by parts leads to

$$\int_{\mathbb{T}} \big(\xi_{xx}(u^{n+1})_x + 4\xi_x(\sqrt{u^{n+1}})_x^2\big)\,dx = 2\int_{\mathbb{T}} \xi\, u^{n+1}\Big(\frac{(\sqrt{u^{n+1}})_{xx}}{\sqrt{u^{n+1}}}\Big)_x\,dx.$$

Altogether, the variational formula (37) implies

$$\frac{1}{\tau}\int_{\mathbb{T}} \xi(x)(T(x) - x)\,u^{n+1}(x)\,dx = 2\int_{\mathbb{T}} \xi\, u^{n+1}\Big(\frac{(\sqrt{u^{n+1}})_{xx}}{\sqrt{u^{n+1}}}\Big)_x\,dx. \tag{38}$$

To identify the expression on the left-hand side in terms of $u^{n+1}-u^n$, integrate this difference against a smooth test function $\varphi:\mathbb{T}\to\mathbb{R}$ and use $u^n=T_\#u^{n+1}$,

$$\begin{aligned}\int_{\mathbb{T}}\varphi(x)\big(u^{n+1}(x)-u^n(x)\big)\,dx&=\int_{\mathbb{T}}\big(\varphi(x)-\varphi(T(x))\big)\,u^{n+1}(x)\,dx\\&=\int_{\mathbb{T}}\varphi_x(x)(x-T(x))\,u^{n+1}(x)\,dx+\mathcal{O}\big(\tau^2\big).\end{aligned}$$

Above, we have assumed implicitly that $T(x)=x+\mathcal{O}(\tau)$, which can be justified by an energy estimate [1]. Choosing $\xi=\varphi_x$ in (38),

$$\frac{1}{\tau}\int_{\mathbb{T}}\varphi\big(u^{n+1}-u^n\big)\,dx=2\int_{\mathbb{T}}\varphi_x\,u^{n+1}\Big(\frac{(\sqrt{u^{n+1}})_{xx}}{\sqrt{u^{n+1}}}\Big)_x\,dx+\mathcal{O}(\tau),\tag{39}$$

which is (a weak formulation of) the implicit Euler scheme for (1), with an $\mathcal{O}(\tau)$-correction.

Appendix B. Proofs of Lemma 2.2, Lemma 2.5, and Lemma 2.6.

Proof of Lemma 2.2. Without loss of generality, we assume that the optimal plan π_{opt} connecting u_1 to u_2 is symmetric on $\mathbb{T}\times\mathbb{T}$ with respect to reflection about the point $(1/2,1/2)$. Indeed, given an arbitrary (possibly asymmetric) optimal plan π_1, define the measure π_2 as the point reflection of π_1 about $(1/2,1/2)$. By symmetry of u_1 and u_2 with respect to $x=1/2$, also π_2 is a connecting plan; it is also optimal since it gives the same value for the integral in (3). And clearly, the *point-symmetric* convex combination $\pi_{opt}:=(\pi_1+\pi_2)/2$ is optimal as well.

Next, we argue that $\pi_{opt}(Z)=0$ for $Z:=(0,1/2)\times(1/2,1)\cup(1/2,1)\times(0,1/2)\subset\mathbb{T}^2$, i.e. that no mass is transferred from the interval $(0,1/2)$ to $(1/2,1)$, and vice versa. To see this, define the measure γ_1 on $\mathbb{T}\times\mathbb{T}$ by restriction of π to Z, and let γ_2 be the measure supported on $\mathbb{T}\times\mathbb{T}\setminus Z$, obtained by reflection of γ_1 about the axis $y=1/2$. Then $\tilde\pi_{opt}:=\pi_{opt}-\gamma_1+\gamma_2$ is a measure on $\mathbb{T}\times\mathbb{T}$, and (using the point-symmetry of π) it has u_1 and u_2 as its respective marginal densities. Hence, $\tilde\pi_{opt}$ is a transport plan connecting u_1 to u_2, and it satisfies $\tilde\pi_{opt}(Z)=0$. (Intuitively, whenever π_{opt} transports some piece of mass from $x\in(0,1/2)$ to $y\in(1/2,1)$, the plan $\tilde\pi_{opt}$ transports the same piece of mass from x to $\tilde y=1-y$.) Since $d(x,y)\geq d(x,1-y)$ for $(x,y)\in Z$, it follows that

$$\int_{\mathbb{T}\times\mathbb{T}\setminus Z}d(x,y)^2\,d\gamma_2(x,y)=\int_Z d(x,1-y')^2\,d\gamma_1(x,y')\leq\int_Z d(x,y')^2\,d\gamma_1(x,y').$$

This implies that the plan $\tilde\pi_{opt}$ is optimal, since

$$\begin{aligned}\int_{\mathbb{T}\times\mathbb{T}}d(x,y)^2\,d\tilde\pi_{opt}(x,y)&=\int_{\mathbb{T}\times\mathbb{T}}d(x,y)^2\,d\pi_{opt}(x,y)\\&\quad+\int_{\mathbb{T}\times\mathbb{T}\setminus Z}d(x,y)^2\,d\gamma_2(x,y)-\int_Z d(x,y)^2\,d\gamma_1(x,y)\\&\leq\int_{\mathbb{T}\times\mathbb{T}}d(x,y)^2\,d\pi_{opt}(x,y).\end{aligned}$$

In conclusion, the optimal transport problem on $\mathbb{T}$ decouples into two independent problems, one on $(0,1/2)$ and one on $(1/2,1)$. Formula (5) applies to each of these problems separately. It is now easily seen (using again the symmetry of the u_i about $x=1/2$) that formula (5) can simply be used on the whole interval. □

Proof of Lemma 2.5. Define $g \in \mathcal{G}_M^n$ from $\mathbf{g}$ by (19); it holds

$$g(\omega) = g_k \frac{\omega - \omega_{k-1}}{\delta_k} + g_{k-1} \frac{\omega_k - \omega}{\delta_k}, \quad \text{for} \quad \omega \in [\omega_{k-1}, \omega_k]. \tag{40}$$

Defining x_ω as in (21), it follows for $k = 1, 2, \dots, n$ that

$$\begin{aligned} x_\omega - x_{k-1} &= \frac{1}{\delta_k} \int_{\omega_{k-1}}^{\omega} \big(g_k(\eta - \omega_{k-1}) + g_{k-1}(\omega_k - \eta)\big)\, d\eta \\ &= \frac{\omega - \omega_{k-1}}{\delta_k} \big(g_k(\omega - \omega_{k-1}) + g_{k-1}(2\omega_k - \omega_{k-1} - \omega)\big), \end{aligned}$$

and similarly,

$$x_k - x_\omega = \frac{\omega_k - \omega}{\delta_k} \big(g_k(\omega + \omega_k - 2\omega_{k-1}) + g_{k-1}(\omega_k - \omega)\big).$$

A straight-forward calculation reveals that

$$\begin{aligned} g_k^2(x_\omega - x_{k-1}) + g_{k-1}^2(x_k - x_\omega) &= \frac{g_k + g_{k-1}}{2\delta_k} \big(g_k(\omega - \omega_{k-1}) + g_{k-1}(\omega_k - \omega)\big)^2 \\ &= \frac{1}{2}(g_k + g_{k-1})\delta_k g(\omega_k)^2. \end{aligned}$$

This shows that u defined in (22) indeed satisfies the relation (21). □

Proof of Lemma 2.6. In view of (14), the matrix elements are given by

$$a_{j,k} = \iint_{[0,M]^2} \big(M - \max(\eta, \eta')\big) \phi_j(\eta) \phi_k(\eta')\, d\eta\, d\eta'.$$

By the definition of the functions ϕ_ℓ in (17) and (18), it is easily verified that for $1 \le j \le k \le n$

$$\int_0^M \phi_j(\eta)\, d\eta = \Delta_j \quad \text{and} \quad \int_0^M \eta' \phi_k(\eta')\, d\eta' = \Delta_k \sigma_k,$$

with $\sigma_k = (\omega_{k-1} + \omega_k + \omega_{k+1})/3$ for $k < n$ and $\sigma_n = M/2$. First, assume that $k < n$. Then, by Fubini's theorem,

$$\begin{aligned} a_{j,k} &= M \int_0^M \phi_j(\eta)\, d\eta \int_0^M \phi_k(\eta')\, d\eta' - \int_0^M \phi_j(\eta)\, d\eta \int_0^M \eta' \phi_k(\eta')\, d\eta' \\ &\quad - \underbrace{\iint_{[0,M]^2} (\eta - \eta')_+ \phi_j(\eta) \phi_k(\eta')\, d\eta\, d\eta'}_{=: J_{j,k}} \\ &= \Delta_j \Delta_k (M - \sigma_k) - J_{j,k}. \end{aligned}$$

As usual, $(\eta - \eta')_+$ equals $\eta - \eta'$ for $\eta > \eta'$ and is zero otherwise. If $j + 2 \le k < n$, then $\eta \le \eta'$ on the support of $\phi_j(\eta)\phi_k(\eta')$, so $J_{j,k}$ vanishes. Now assume that $k = j + 1 < n$; then

$$\begin{aligned} J_{j,j+1} &= \iint_{[0,M]^2} (\eta - \eta')_+ \phi_j(\eta) \phi_{j+1}(\eta')\, d\eta\, d\eta' \\ &= \int_{\omega_{j-1}}^{\omega_{j+1}} \Big(\phi_j(\eta) \int_{\omega_j}^{\max(\eta, \omega_j)} (\eta - \eta') \phi_{j+1}(\eta')\, d\eta'\Big)\, d\eta \\ &= \delta_{j+1}^{-2} \int_{\omega_j}^{\omega_{j+1}} \Big((\omega_{j+1} - \eta) \int_{\omega_j}^{\eta} (\eta - \eta')(\eta' - \omega_j)\, d\eta'\Big)\, d\eta = \frac{1}{120} \delta_{j+1}^3. \end{aligned}$$

On the other hand, for $k = j < n$ we find

$$\begin{aligned}
J_{j,j} &= \iint_{[0,M]^2} (\eta - \eta')_+ \phi_j(\eta)\phi_j(\eta')\,d\eta\,d\eta' \\
&= \int_{\omega_{j-1}}^{\omega_{j+1}} \Big(\phi_j(\eta) \int_{\omega_{j-1}}^{\eta} (\eta - \eta')\phi_j(\eta')\,d\eta'\Big)\,d\eta \\
&= \delta_j^{-2} \int_{\omega_{j-1}}^{\omega_j} \Big((\eta - \omega_{j-1}) \int_{\omega_{j-1}}^{\eta} (\eta - \eta')(\eta' - \omega_{j-1})\,d\eta'\Big)\,d\eta \\
&\quad + \delta_{j+1}^{-2} \int_{\omega_j}^{\omega_{j+1}} \Big((\omega_{j+1} - \eta) \int_{\omega_j}^{\eta} (\eta - \eta')(\omega_{j+1} - \eta')\,d\eta'\Big)\,d\eta \\
&\quad + (\delta_j\delta_{j+1})^{-1} \int_{\omega_j}^{\omega_{j+1}} \Big((\omega_{j+1} - \eta) \int_{\omega_{j-1}}^{\omega_j} (\eta - \eta')(\eta' - \omega_{j-1})\,d\eta'\Big)\,d\eta \\
&= \frac{\Delta_j}{60}(12\Delta_j^2 + \delta_j^2 + \delta_{j+1}^2).
\end{aligned}$$

For $k = n$, we proceed in an analogous manner, taking into account that ϕ_n is supported in $[0, \omega_1] \cup [\omega_{n-1}, M]$:

$$\begin{aligned}
a_{j,n} &= M \int_0^M \phi_j(\eta)\,d\eta \int_0^M \phi_n(\eta')\,d\eta' \\
&\quad - \int_0^M \eta\phi_j(\eta)\,d\eta \int_0^{\omega_1} \phi_n(\eta')\,d\eta' - \int_0^M \phi_j(\eta)\,d\eta \int_{\omega_{n-1}}^M \eta'\phi_n(\eta')\,d\eta' \\
&\quad - \underbrace{\int_0^M \int_0^{\omega_1} (\eta' - \eta)_+ \phi_j(\eta)\phi_n(\eta')\,d\eta'\,d\eta}_{=:K_j^+} - \underbrace{\int_0^M \int_{\omega_{n-1}}^M (\eta - \eta')_+ \phi_j(\eta)\phi_n(\eta')\,d\eta'\,d\eta}_{=:K_j^-} \\
&= \frac{1}{2}\Delta_j\Delta_n(M - \sigma_j + \Delta_n/3) - K_j^+ - K_j^-.
\end{aligned}$$

Above, we have used that $\delta_1 = \delta_n = \omega_1 = \Delta_n$ by symmetry of the lattice $\{\omega_k\} \subset [0, M]$. For $2 \le j \le n-2$, the support of ϕ_j is contained in $[\omega_1, \omega_{n-1}]$, thus $K_j^+ = K_j^- = 0$. If $j = 1$, then $K_1^- = 0$ while

$$K_1^+ = \delta_1^{-2} \int_0^{\omega_1} \Big(\eta \int_\eta^{\omega_1} (\eta' - \eta)(\omega_1 - \eta')\,d\eta'\Big)\,d\eta = \frac{\Delta_n^3}{120}.$$

If, on the other hand, $j = n-1$, then $K_{n-1}^+ = 0$ and

$$K_{n-1}^- = \delta_n^{-2} \int_{\omega_{n-1}}^M \Big((M - \eta) \int_{\omega_{n-1}}^{\eta} (\eta - \eta')(\eta' - \omega_{n-1})\,d\eta'\Big)\,d\eta = \frac{\Delta_n^3}{120}.$$

The most complicated case is met when $j = k = n$,

$$\begin{aligned}
K_n^+ &= \delta_1^{-2} \int_0^{\omega_1} \Big((\omega_1 - \eta) \int_\eta^{\omega_1} (\eta' - \eta)(\omega_1 - \eta')\,d\eta'\Big)\,d\eta = \frac{\Delta_n^3}{30}, \\
K_n^- &= \delta_n^{-2} \int_{\omega_{n-1}}^M \Big((\eta - \omega_{n-1}) \int_{\omega_{n-1}}^{\eta} (\eta - \eta')(\eta' - \omega_{n-1})\,d\eta'\Big)\,d\eta = \frac{\Delta_n^3}{30}.
\end{aligned}$$

□

REFERENCES

[1] L. Ambrosio, G. Savaré. *Gradient flows of probability measures.* Handbook of Evolution Equations, Dafermos, C. and Feireisl, E. (ed.); Vol. 3, Elsevier (2006) 1–136.

[2] J.W. Barrett, J.F. Blowey, and H. Garcke. *Finite element approximation of a fourth order nonlinear degenerate parabolic equation.* Numer. Math. **80** (1998), no. 4, 525–556.

[3] J. Becker, G. Grün, M. Lenz, and M. Rumpf. *Numerical methods for fourth order nonlinear degenerate diffusion problems.* Mathematical theory in fluid mechanics (Paseky, 2001). Appl. Math. **47** (2002), no. 6, 517–543.

[4] F. Bernis and A. Friedman. *Higher order nonlinear degenerate parabolic equations.* J. Differential Equations **83** (1990), no. 1, 179–206.

[5] A.L. Bertozzi and M. Pugh. *The lubrication approximation for thin viscous films: regularity and long-time behavior of weak solutions.* Comm. Pure Appl. Math., **49** (1996), no. 2, 85–123.

[6] A.L. Bertozzi and L. Zhornitskaya. *Positivity-preserving numerical schemes for lubrication-type equations.* SIAM J. Numer. Anal. **37** (2000), no. 2, 523–555

[7] P. Bleher, J. Lebowitz, and E. Speer. *Existence and positivity of solutions of a fourth-order nonlinear PDE describing interface fluctuations.* Commun. Pure Appl. Math., **47** (1994), 923–942.

[8] M. Cáceres, J.A. Carrillo, and G. Toscani. *Long-time behavior for a nonlinear fourth-order parabolic equation.* Trans. Amer. Math. Soc. **357** (2005), no. 3, 1161–1175

[9] J.W. Cahn and J.E. Hilliard. *Spinodal decomposition: A reprise.* Acta Metallurgica **19** (1971), 151–161.

[10] J.A. Carrillo, R.J. McCann, and C. Villani. *Contractions in the 2-Wasserstein length space and thermalization of granular media.* Arch. Rational Mech. Anal. **179**, no. 2, 217–263.

[11] J.A. Carrillo and J.S. Moll. *Numerical simulation of diffusive and aggregation phenomena in nonlinear continuity equations by evolving diffeomorphisms.* To appear in: SIAM J. Sci. Computing (2009).

[12] J.A. Carrillo, A. Jüngel, and S. Tang. *Positive entropic schemes for a nonlinear fourth-order equation.* Discrete Contin. Dynam. Sys. B **3** (2003), 1–20.

[13] R. Dal Passo, H. Garcke, and G. Grün. *On a fourth-order degenerate parabolic equation: global entropy estimates, existence, and qualitative behavior of solutions.* SIAM J. Math. Anal., **29** (1998), no. 2, 321–342

[14] P. Degond, F. Méhats, and C. Ringhofer. *Quantum energy-transport and drift-diffusion models.* J. Stat. Phys. **118** (2005), no. 3-4, 625–667.

[15] B. Derrida, J. Lebowitz, E. Speer, and H. Spohn. *Fluctuations of a stationary nonequilibrium interface.* Phys. Rev. Lett. **67** (1991), 165–168.

[16] C.M. Elliott and H. Garcke. *On the Cahn-Hilliard equation with degenerate mobility.* SIAM J. Math. Anal. **27** (1996), no. 2, 404–423.

[17] U. Gianazza, G. Savaré, and G. Toscani. *The Wasserstein gradient flow of the Fisher information and the quantum drift-diffusion equation.* To appear in Arch. Rat. Mech., 2009.

[18] M.P. Gualdani, A. Jüngel, and G. Toscani. *A nonlinear fourth-order parabolic equation with nonhomogeneous boundary conditions.* SIAM J. Math. Anal. **37** (2006), no. 6, 1761–1779.

[19] R. Jordan, D. Kinderlehrer, and F. Otto. *The variational formulation of the Fokker-Planck equation.* SIAM J. Math. Anal. **29** (1998), no. 1, 1–17.

[20] A. Jüngel. *Transport Equations for Semiconductors.* Lecture Notes in Physics No. 773. Springer, Berlin, 2009.

[21] A. Jüngel and D. Matthes. *An algorithmic construction of entropies in higher-order nonlinear PDEs.* Nonlinearity **19** (2006), no. 3, 633–659.

[22] A. Jüngel and D. Matthes. *The Derrida-Lebowitz-Speer-Spohn equation: existence, nonuniqueness, and decay rates of the solutions.* SIAM J. Math. Anal. **39** (2008), no. 6, 1996–2015.

[23] A. Jüngel and D. Matthes. *A review on results for the Derrida-Lebowitz-Speer-Spohn equation.* To appear in Proceedings of the Conference Equadiff 2007, 2009.

[24] A. Jüngel and J.P. Milisic. *A sixth-order nonlinear parabolic equation for quantum systems.* To appear in SIAM J. Math. Anal., 2009.

[25] A. Jüngel and R. Pinnau. *Global nonnegative solutions of a nonlinear fourth-order parabolic equation for quantum systems.* SIAM J. Math. Anal. **32** (2000), no. 4, 760–777.

[26] A. Jüngel and R. Pinnau. *A positivity preserving numerical scheme for a nonlinear fourth-order parabolic equation.* SIAM J. Numer. Anal. **39** (2001), 385–406.

[27] A. Jüngel and I. Violet. *First-order entropies for the Derrida-Lebowitz-Speer-Spohn equation.* Discrete Contin. Dyn. Syst. Ser. B **8** (2007), no. 4, 861–877

[28] R.S. Laugesen. *New dissipated energies for the thin fluid film equation.* Commun. Pure Appl. Anal. **4** (2005), no. 3, 613–634.

[29] F. Otto. *The geometry of dissipative evolution equations: the porous medium equation.* Comm. Partial Differential Equations **26**, no. 1–2, 101–174.

[30] C. Villani. *Optimal transport. Old and new.* Grundlehren der Mathematischen Wissenschaften, 338. Springer-Verlag, Berlin, 2009.

E-mail address: bduering@anum.tuwien.ac.at
E-mail address: matthes@asc.tuwien.ac.at
E-mail address: pina@webmail.fer.hr