

A Lagrangian Scheme for the Solution of Nonlinear Diffusion Equations Using Moving Simplex Meshes

José A. Carrillo¹  · Bertram Düring²  · Daniel Matthes³ · David S. McCormick²

Received: 6 February 2017 / Revised: 17 October 2017 / Accepted: 27 October 2017 /
Published online: 7 November 2017
© The Author(s) 2017. This article is an open access publication

Abstract A Lagrangian numerical scheme for solving nonlinear degenerate Fokker–Planck equations in space dimensions $d \geq 2$ is presented. It applies to a large class of nonlinear diffusion equations, whose dynamics are driven by internal energies and given external potentials, e.g. the porous medium equation and the fast diffusion equation. The key ingredient in our approach is the gradient flow structure of the dynamics. For discretization of the Lagrangian map, we use a finite subspace of linear maps in space and a variational form of the implicit Euler method in time. Thanks to that time discretisation, the fully discrete solution inherits energy estimates from the original gradient flow, and these lead to weak compactness of the trajectories in the continuous limit. Consistency is analyzed in the planar situation, $d = 2$. A variety of numerical experiments for the porous medium equation indicates that the scheme is well-adapted to track the growth of the solution’s support.

JAC acknowledges support by the Engineering and Physical Sciences Research Council (EPSRC) under Grant No. EP/P031587/1, by the Royal Society and the Wolfson Foundation through a Royal Society Wolfson Research Merit Award and by the National Science Foundation (NSF) under Grant No. RNMS11-07444 (KI-Net). DM was supported by the DFG Collaborative Research Center TRR 109, “Discretization in Geometry and Dynamics”. BD and DSMcC were supported by the Leverhulme Trust research project grant “Novel discretisations for higher-order nonlinear PDE” (RPG-2015-69).

✉ José A. Carrillo
carrillo@imperial.ac.uk

Bertram Düring
bd80@sussex.ac.uk

Daniel Matthes
matthes@ma.tum.de

David S. McCormick
D.S.McCormick@sussex.ac.uk

¹ Department of Mathematics, Imperial College London, London SW7 2AZ, UK

² Department of Mathematics, University of Sussex, Pevensey II, Brighton BN1 9QH, UK

³ Zentrum Mathematik, Technische Universität München, Boltzmannstraße 3, 85747 Garching, Germany

Keywords Lagrangian method · Optimal transport · Variational schemes · Nonlinear diffusions

1 Introduction

1.1 Nonlinear Fokker–Planck Equations

We study a variational Lagrangian discretization of the following type of initial value problem:

$$\partial_t \rho = \Delta P(\rho) + \nabla \cdot (\rho \nabla V) \quad \text{on } \mathbb{R}_{>0} \times \mathbb{R}^d, \tag{1.1a}$$

$$\rho(\cdot, 0) = \rho^0 \quad \text{on } \mathbb{R}^d. \tag{1.1b}$$

This problem is posed for the time-dependent probability density function $\rho: \mathbb{R}_{\geq 0} \times \mathbb{R}^d \rightarrow \mathbb{R}_{\geq 0}$, with a given initial density ρ^0 . We assume that the pressure $P: \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$ can be written in the form

$$P(r) = rh'(r) - h(r) \quad \text{for all } r \geq 0, \tag{1.2}$$

for some non-negative and convex $h \in C^1(\mathbb{R}_{\geq 0}) \cap C^\infty(\mathbb{R}_{>0})$, and that $V \in C^2(\mathbb{R}^d)$ is a non-negative potential without loss of generality. Problem (1.1) encompasses a large class of diffusion equations, such as—for power-type nonlinearities $P(r) = r^m$ and vanishing potential $V \equiv 0$ —the heat equation ($m = 1$), porous medium equations ($m > 1$) and fast diffusion equations ($m < 1$). By a slight abuse of notation, we refer to (1.1) with more general P and non-vanishing V as nonlinear Fokker–Planck equations. In this paper, we assume a *degenerate diffusion*, that is $h(0) = h'(0) = 0$, and a *confining potential*, that is V is convex, not necessarily strict. For technical reasons, we further need to assume that

$$\lim_{s \rightarrow \infty} sh''(s) = +\infty. \tag{1.3}$$

Since our particular spatio-temporal discretization of the initial value problem (1.1) is based on the Lagrangian representation of its dynamics, and on its variational formulation, we briefly recall both of them now.

1.2 Lagrangian Formulation

Equation (1.1) can be written as a transport equation,

$$\partial_t \rho + \nabla \cdot (\rho \mathbf{v}[\rho]) = 0, \tag{1.4a}$$

with a velocity field \mathbf{v} that depends on the solution ρ itself,

$$\mathbf{v}[\rho] = -\nabla(h'(\rho) + V). \tag{1.4b}$$

Various further evolution equations can be written in the form (1.4a), such as non-local aggregation equations (see, e.g., Ambrosio et al. [1]); Keller–Segel type models (see, e.g., Blanchet et al. [5]); and also fourth order thin film equations (see, e.g., Otto [34]) or quantum equations (see, e.g., Gianazza et al. [21]). To simplify the presentation, we stick to equations of nonlinear Fokker–Planck type (1.1a).

The system (1.4) naturally induces a Lagrangian representation of the dynamics, which can be summarized as follows. Below, the *reference density* $\bar{\rho}$ is a probability density supported

on some compact set $K \subset \mathbb{R}^d$, and we use the notation $G_{\#}\bar{\rho}$ for the *push-forward* of $\bar{\rho}$ under a map $G : K \rightarrow \mathbb{R}^d$; the definition is recalled in (2.1).

Lemma 1.1 *Assume that $\rho : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}_{\geq 0}$ is a smooth positive solution of (1.1). Let $G^0 : K \rightarrow \mathbb{R}^d$ be a given map such that $G_{\#}^0\bar{\rho} = \rho^0$. Further, let $G : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ be the flow map associated to (1.4b), satisfying*

$$\partial_t G_t = \mathbf{v}[\rho_t] \circ G_t, \quad G(0, \cdot) = G^0, \tag{1.5}$$

where $\rho_t := \rho(t, \cdot)$ and $G_t := G(t, \cdot) : \mathbb{R}^d \rightarrow \mathbb{R}^d$. Then, at any $t \in [0, T]$,

$$\rho_t = (G_t)_{\#}\bar{\rho}. \tag{1.6}$$

In short, the solution G to (1.5) is a Lagrangian map for the solution ρ to (1.1). This fact is an immediate consequence of (1.4a); for convenience of the reader, we recall the proof in ‘‘Appendix A’’. Subsequently, (1.6) can be substituted for ρ in the expression (1.4b) for the velocity, which makes (1.5) an autonomous evolution equation for G :

$$\partial_t G_t = -\nabla \left[h' \left(\frac{\bar{\rho}}{\det DG_t} \right) \right] \circ G_t - \nabla V \circ G_t. \tag{1.7}$$

A more explicit form of (1.7) is derived in (5.2).

1.3 Variational Structure

It is well-known (see Otto [35] or Ambrosio et al. [1]) that (1.1) is a gradient flow for the relative Renyi entropy functional

$$\mathcal{E}(\rho) = \int_{\mathbb{R}^d} [h(\rho(x)) + V(x)\rho(x)] dx, \tag{1.8}$$

with respect to the L^2 -Wasserstein metric on the space $\mathcal{P}_2^{\text{ac}}(\mathbb{R}^d)$ of probability densities on \mathbb{R}^d with finite second moment. It appears to be less well known (see Evans et al. [20], Carrillo and Moll [13], or Carrillo and Lisini [12]) that also (1.7) is a gradient flow, namely for the functional

$$\mathbf{E}(G|\bar{\rho}) := \mathcal{E}(G_{\#}\bar{\rho}) = \int_K \left[\tilde{h} \left(\frac{\det DG}{\bar{\rho}} \right) + V \circ G \right] \bar{\rho} d\omega, \quad \tilde{h}(s) := s h(s^{-1}), \tag{1.9}$$

on the Hilbert space $L^2(K \rightarrow \mathbb{R}^d; \bar{\rho})$ of square integrable maps from K to \mathbb{R}^d . We shall discuss these gradient flow structures in more detail in Sect. 2 below.

1.4 Discretization and Approximation Results

Our discretization in space is based on the Lagrangian formulation. Instead of numerically integrating (1.1a) to obtain the density ρ directly, we approximate the associated Lagrangian maps G that satisfy (1.7): specifically, we assume that a simplicial decomposition \mathcal{T} of K is given, and we restrict G to the finite dimensional subspace $\mathcal{A}_{\mathcal{T}}$ of continuous maps from K to \mathbb{R}^d that are piecewise linear with respect to \mathcal{T} . A posteriori, we recover an approximation of ρ via (1.6). That ansatz for the Lagrangian maps corresponds to a simple geometric picture: the induced densities are piecewise constant on triangles whose vertices move in time.

For the discretization in time, we exploit the aforementioned variational structure of (1.7): namely, we adopt the celebrated *minimizing movement scheme* that is known to provide a robust approximation of gradient flows. In the context at hand, this scheme reads as follows:

let a time step $\tau > 0$ and an initial condition $G_{\boxplus}^0 \in \mathcal{A}_{\mathcal{T}}$ be given. (Here and below, \boxplus symbolizes the space-time mesh generated by \mathcal{T} on K and τ on $\mathbb{R}_{>0}$.) Then the n th time iterate $G_{\boxplus}^n \in \mathcal{A}_{\mathcal{T}}$ —that serves as our approximation of $G(n\tau; \cdot)$ —is chosen inductively for $n = 1, 2, \dots$ as the minimizer in the respective problem

$$\frac{1}{2\tau} \|G - G_{\boxplus}^n\|_{L^2(K \rightarrow \mathbb{R}^d; \bar{\rho})}^2 + \mathbf{E}(G|\bar{\rho}) \longrightarrow \min, \tag{1.10}$$

where the minimization is carried out over the finite dimensional space $\mathcal{A}_{\mathcal{T}}$. With the sequence $(G_{\boxplus}^n)_{n=0,1,\dots}$ of approximating Lagrangian maps at hand, we define piecewise-constant-in-time interpolations for the derived density $\tilde{\rho}_{\boxplus}$ and velocity $\tilde{\mathbf{v}}_{\boxplus}$ as usual via

$$\tilde{\rho}_{\boxplus}(t) = (G_{\boxplus}^n)_{\#}\bar{\rho}, \quad \tilde{\mathbf{v}}_{\boxplus}(t) = \frac{G_{\boxplus}^n - G_{\boxplus}^{n-1}}{\tau} \quad \text{with } n \text{ such that } t \in ((n-1)\tau, n\tau].$$

Our analytical results on the scheme can be summarized as follows.

- The sequence of fully discrete minimization problems (1.10) is well-posed: see Lemma 3.1. We thus obtain a sequence $(G_{\boxplus}^n)_{n=0,1,\dots}$ for each sufficiently fine discretization \boxplus .
- The G_{\boxplus}^n are entropy-diminishing and are \boxplus -uniformly Hölder continuous: see Lemma 4.1.
- Consequently, the induced densities $\tilde{\rho}_{\boxplus}$ converge weakly to an absolutely continuous limit trajectory ρ , and the fluxes $\tilde{\rho}_{\boxplus}\tilde{\mathbf{v}}_{\boxplus}$ converge weakly to a limit of the form $\rho\mathbf{v}$: see Theorem 4.2. The identification of the limit velocity \mathbf{v} , however, is only possible under strong additional hypotheses: see Corollary 4.5.
- In $d = 2$ dimensions, we prove numerical consistency in the sense that, if G is a smooth solution to (1.7), then its restriction to the mesh \boxplus satisfies the fully discrete Euler–Lagrange equations associated to (1.10), with a quantifiable error that vanishes in a suitable continuous limit: see Theorem 5.2.
- Our previously mentioned consistency results requires that the triangulation \mathcal{T} of K is almost ideally hexagonal: see Eq. (5.7). We discuss why consistency cannot be expected if that condition is violated: see Remark 5.4.

1.5 Comparison with Results in the Literature

The approach presented in this paper is an alternative to the one developed by Carrillo et al. [13, 15], where G is obtained by directly solving the PDE (1.7) numerically with finite differences or Galerkin approximation via finite element methods. In other words, while Carrillo et al. [13, 15] follows the strategy *minimize first then discretize*, our present approach is to *discretize first then minimize*. In the former approach, the minimization (1.10) is performed on the spatially continuous level, yielding Euler–Lagrange equations that are then discretized in space; in the present approach, the space of Lagrangian maps is approximated by the finite dimensional subspace $\mathcal{A}_{\mathcal{T}}$, and the minimization problem (1.10) on $\mathcal{A}_{\mathcal{T}}$ yields a nonlinear system of Euler–Lagrange equations that are directly solvable numerically.

Let us mention that other numerical methods have been developed to conserve particular properties of solutions of the gradient flow (1.1). Finite volume methods preserving the decay of energy at the semi-discrete level, along with other important properties like non-negativity and mass conservation, were proposed in the papers [4, 8, 10]. Particle methods based on suitable regularizations of the flux of the continuity Eq. (1.1) have been proposed in the papers [18, 27, 28, 37]. A particle method based on the steepest descent of a regularized internal part of the energy \mathcal{E} in (1.8) by substituting particles by non-overlapping blobs was proposed and analysed in Carrillo et al. [11, 14]. Deterministic particle methods

for diffusions have been recently explored, see [9] and the references therein. High-order relaxation schemes for nonlinear diffusion problems have been proposed in Cavalli et al. [16], while high-resolution schemes for nonlinear convection-diffusion problems are introduced in Kurganov et al. [26]. Moreover, the numerical approximation of the JKO variational scheme has already been tackled by different methods using pseudo-inverse distributions in one dimension (see [5, 7, 23, 40]) or solving for the optimal map in a JKO step (see [3, 25]). Finally, note that gradient-flow-based Lagrangian methods in one dimension for higher-order, drift diffusion and Fokker–Planck equations have recently been proposed in the papers [19, 31–33].

There are two main arguments in favour of our taking this indirect approach of solving (1.7) instead of solving (1.1). The first is our interest in *structure-preserving discretizations*: the scheme that we present builds on the non-obvious “secondary” gradient flow representation of (1.1) in terms of Lagrangian maps. The benefits include monotonicity of the transformed entropy functional \mathbf{E} and L^2 control on the metric velocity for our fully discrete solutions, that eventually lead to weak compactness of the trajectories in the continuous limit. We remark that our long-term goal is to design a numerical scheme that makes full use of the much richer “primary” variational structure of (1.1) in the Wasserstein distance, which is reviewed in Sect. 2 below. However, despite significant effort in the recent past—see, e.g., the references [3, 5, 14, 15, 19, 22, 25, 29, 36, 40]—it has not been possible so far to preserve features like metric contractivity of the flow under the discretization, except in the rather special situation of one space dimension (see Matthes and Osberger [29]). This is mainly due to the non-existence of finite-dimensional submanifolds of $\mathcal{P}_2^{\text{ac}}(\mathbb{R}^d)$ that are complete with respect to generalized geodesics.

The second motivation is that Lagrangian schemes are a natural choice for *numerical front tracking*, see, e.g., Budd [6] for first results on the numerical approximation of self-similar solutions to the porous medium equation. We recall that, due to the assumed degeneracy $P'(0) = 0$ of the diffusion in (1.1), solutions that are compactly supported initially remain compactly supported for all times. A numerically accurate calculation of the moving edge of support is challenging, since the solution can have a very complex behavior near that edge, like the waiting time phenomenon (see Vazquez [38]). Our simulation results for $\partial_t \rho = \Delta(\rho^3)$ —which possesses an analytically known, compactly supported, self-similar Barenblatt solution—indicate that our discretization is indeed able to track the edge of support quite accurately.

The expected convergence of our scheme, with implicit Euler stepping in time and piecewise linear approximation of the Lagrangian maps, is of first order in both space and time. This is confirmed in our experiments. For an improved approximation, particularly of the moving fronts, numerical schemes with a higher order of consistency would be desirable. In principle, such schemes could be constructed along the same lines, for example, by replacing the implicit Euler method by a Runge–Kutta method in time, and the piecewise constant ansatz space $\mathcal{A}_{\mathcal{G}}$ by finite elements with functions of higher global regularity in space. However, it is unclear if a similar degree of structure preservation can be achieved for these schemes, and their analysis would be very different from the one presented here.

1.6 Structure of the Paper

This work is organized as follows. In Sect. 2, we present an overview of previous results in gradient flows pertaining our work. Section 3 is devoted to the introduction of the linear set of Lagrangian maps and the derivation of the numerical scheme. Section 4 shows the compactness of the approximated sequences of discretizations and we give conditions leading to the eventual convergence of the scheme towards (1.1). Section 5 deals with the consistency

of the scheme in two dimensions, while Sect. 6 gives several numerical tests showing the performance of this scheme.

2 Gradient Flow Structures

2.1 Notations from Probability Theory

$\mathcal{P}(X)$ is the space of probability measures on a given base set X . We say that a sequence (μ_n) of measures in $\mathcal{P}(X)$ converges narrowly to a limit μ in that space if

$$\int_X f(x) \, d\mu_n(x) \rightarrow \int_X f(x) \, d\mu(x)$$

for all bounded and continuous functions $f \in C_b^0(X)$. The push-forward $T_{\#}\mu$ of a measure $\mu \in \mathcal{P}(X)$ under a measurable map $T: X \rightarrow Y$ is the uniquely determined measure $\nu \in \mathcal{P}(Y)$ such that, for all $g \in C_b^0(Y)$,

$$\int_X g \circ T(x) \, d\mu(x) = \int_Y g(y) \, d\nu(y).$$

With a slight abuse of notation — identifying absolutely continuous measures with their densities—we denote the space of probability densities on \mathbb{R}^d of finite second moment by

$$\mathcal{P}_2^{\text{ac}}(\mathbb{R}^d) = \left\{ \rho \in L^1(\mathbb{R}^d); \rho \geq 0, \int_{\mathbb{R}^d} \rho(x) \, dx = 1, \int_{\mathbb{R}^d} \|x\|^2 \rho(x) \, dx < \infty \right\}.$$

Clearly, the reference density $\bar{\rho}$, which is supported on the compact set $K \subset \mathbb{R}^d$, belongs to $\mathcal{P}_2^{\text{ac}}(\mathbb{R}^d)$. If $G: K \rightarrow \mathbb{R}^d$ is a diffeomorphism onto its image (which is again compact), then the push-forward of $\bar{\rho}$'s measure produces again a density $G_{\#}\bar{\rho} \in \mathcal{P}_2^{\text{ac}}(\mathbb{R}^d)$, given by

$$G_{\#}\bar{\rho} = \frac{\bar{\rho}}{\det DG} \circ G^{-1}. \tag{2.1}$$

2.2 Gradient Flow in the Wasserstein Metric

Below, some basic facts about the Wasserstein metric and the formulation of (1.1) as gradient flow in that metric are briefly reviewed. For more detailed information, we refer the reader to the monographs of Ambrosio et al. [1] and Villani [39].

One of the many equivalent ways to define the L^2 -Wasserstein distance between $\rho_0, \rho_1 \in \mathcal{P}_2^{\text{ac}}(\mathbb{R}^d)$ is as follows:

$$W_2(\rho_0, \rho_1) := \inf \left\{ \int_{\mathbb{R}^d} \|T(x) - x\|^2 \rho_0(x) \, dx; T: \mathbb{R}^d \rightarrow \mathbb{R}^d \text{ measurable, } T_{\#}\rho_0 = \rho_1 \right\}^{\frac{1}{2}}. \tag{2.2}$$

The infimum above is in fact a minimum, and the — essentially unique — optimal map T^* is characterized by Brenier's criterion; see, e.g., Villani [39, Section 2.1]. A trivial but

essential observation is that if $\bar{\rho} \in \mathcal{P}_2^{\text{ac}}(\mathbb{R}^d)$ is a reference density with support $K \subset \mathbb{R}^d$, and $\rho_0 = (G_0)_\# \bar{\rho}$ with a measurable $G_0: K \rightarrow \mathbb{R}^d$, then (2.2) can be re-written as follows:

$$W_2(\rho_0, \rho_1) = \inf \left\{ \int_K \|G(\omega) - G_0(\omega)\|^2 \bar{\rho}(\omega) \, d\omega; G: K \rightarrow \mathbb{R}^d \text{ measurable, } G_\# \bar{\rho} = \rho_1 \right\}^{\frac{1}{2}}, \tag{2.3}$$

and the essentially unique minimizer G^* in (2.3) is related to the optimal map T^* in (2.2) via $G^* = T^* \circ G_0$.

W_2 is a metric on $\mathcal{P}_2^{\text{ac}}(\mathbb{R}^d)$; convergence in W_2 is equivalent to weak- \star convergence in $L^1(\mathbb{R}^d)$ and convergence of the second moment. Since P and hence also h are of super-linear growth at infinity, each sublevel set \mathcal{E} is weak- \star closed and thus complete with respect to W_2 .

As already mentioned above, solutions ρ to (1.1) constitute a gradient flow for the functional \mathcal{E} from (1.8) in the metric space $(\mathcal{P}_2^{\text{ac}}(\mathbb{R}^d); W_2)$. In fact, assuming that the potential V is λ -convex (i.e., $\nabla^2 V \geq \lambda \mathbb{1}$), the flow is even λ -contractive as a semi-group, thanks to the λ -uniform displacement convexity of \mathcal{E} (see McCann [30], or Daneri and Savaré [17]), which is a strengthened form of λ -uniform convexity along geodesics. The λ -contractivity of the flow implies various properties (see Ambrosio et al. [1, Section 11.2]) like global existence, uniqueness and regularity of the flow, monotonicity of \mathcal{E} and its sub-differential, uniform exponential estimates on the convergence (if $\lambda > 0$) or divergence (if $\lambda \leq 0$) of trajectories, quantified exponential rates for the approach to equilibrium (if $\lambda > 0$) and the like.

An important further consequence is that the unique flow can be obtained as the limit for $\tau \searrow 0$ of the time-discrete *minimizing movement scheme* (see Ambrosio et al. [1] and Jordan, Kinderlehrer and Otto [24]):

$$\rho_\tau^n := \operatorname{argmin}_{\rho \in \mathcal{P}_2^{\text{ac}}(\mathbb{R}^d)} \mathcal{E}_\tau(\rho; \rho_\tau^{n-1}), \quad \mathcal{E}_\tau(\rho, \hat{\rho}) := \frac{1}{2\tau} W_2(\rho, \hat{\rho})^2 + \mathcal{E}(\rho). \tag{2.4}$$

This time discretization is well-adapted to approximate λ -contractive gradient flows. All of the properties of mentioned above are already reflected on the level of these time-discrete solutions.

2.3 Gradient Flow in L^2

Equation (1.7) is the gradient flow of \mathbf{E} on the space $L^2(K \rightarrow \mathbb{R}^d; \bar{\rho})$ of square integrable (with respect to $\bar{\rho}$) maps $G: K \rightarrow \mathbb{R}^d$ (see Evans et al. [20] or Jordan et al. [25]). However, the variational structure behind this gradient flow is much weaker than above: most notably, \mathbf{E} is only poly-convex, but *not λ -uniformly convex*. Therefore, the abstract machinery for λ -contractive gradient flows in Ambrosio et al. [1] does not apply here. Clearly, by equivalence of (1.1) and (1.7) at least for sufficiently smooth solutions, certain properties of the primary gradient flow are necessarily inherited by this secondary flow, but for instance λ -contractivity of the flow in the L^2 -norm seems to fail.

Nevertheless, it can be proven (see Ambrosio, Lisini and Savaré [2]) that the gradient flow is globally well-defined, and it can again be approximated by the minimizing movement scheme:

$$G_\tau^n := \operatorname{argmin}_{G \in L^2(K \rightarrow \mathbb{R}^d; \bar{\rho})} \mathbf{E}_\tau(G; G_\tau^{n-1}), \quad \mathbf{E}_\tau(G; \hat{G}) = \frac{1}{2\tau} \int_K \|G - \hat{G}\|^2 \, d\bar{\rho} + \mathbf{E}(G|\bar{\rho}). \tag{2.5}$$

In fact, there is an equivalence between (2.5) and (2.4): simply substitute $(G_\tau^{n-1})_\# \bar{\rho}$ for ρ_τ^{n-1} and $G_\# \bar{\rho}$ for ρ in (2.4); notice that any $\rho \in \mathcal{P}_2^{\text{ac}}(\mathbb{R}^d)$ can be written as $G_\# \bar{\rho}$ with a suitable (highly non-unique) choice of $G \in L^2(K \rightarrow \mathbb{R}^d; \bar{\rho})$. This equivalence was already exploited in Carrillo et al. [13,15]. Thanks to the equality (2.3), the minimization with respect to $\rho = G_\# \bar{\rho}$ can be relaxed to a minimization with respect to G . Consequently, if $(G_\tau^0)_\# \bar{\rho} = \rho_\tau^0$, then $(G_\tau^n)_\# \bar{\rho} = \rho_\tau^n$ at all discrete times $n = 1, 2, \dots$. However, while the functional $\mathcal{E}_\tau(\cdot; \rho_\tau^{n-1})$ in (2.4) is $(\lambda + \tau^{-1})$ -uniformly convex in ρ along geodesics in W_2 , the functional $\mathbf{E}_\tau(\cdot; G_\tau^{n-1})$ in (2.5) has apparently no useful convexity properties in G on $L^2(K \rightarrow \mathbb{R}^d; \bar{\rho})$.

3 Definition of the Numerical Scheme

Recall the Lagrangian formulation of (1.1) that has been given in Lemma 1.1. For definiteness, fix a reference density $\bar{\rho} \in \mathcal{P}_2^{\text{ac}}(\mathbb{R}^d)$, whose support $K \subset \mathbb{R}^d$ is a compact, convex polytope.

3.1 Discretization in Space

Our spatial discretization is performed using a finite subspace of linear maps for the Lagrangian maps G . More specifically: let \mathcal{T} be some (finite) simplicial decomposition of K with nodes ω_1 to ω_L and n -simplices Δ_1 to Δ_M . In the case $d = 2$, which is of primary interest here, \mathcal{T} is a triangulation, with triangles Δ_m . The reference density $\bar{\rho}$ is approximated by a density $\bar{\rho}_\mathcal{T} \in \mathcal{P}_2^{\text{ac}}(\mathbb{R}^d)$ that is piecewise constant on the simplices of \mathcal{T} , with respective values

$$\bar{\rho}_\mathcal{T}^m := \frac{\mu_\mathcal{T}^m}{|\Delta_m|} \quad \text{for the simplex masses} \quad \mu_\mathcal{T}^m := \int_{\Delta_m} \bar{\rho}(\omega) \, d\omega. \tag{3.1}$$

The finite dimensional ansatz space $\mathcal{A}_\mathcal{T}$ is now defined as the set of maps $G: K \rightarrow \mathbb{R}^d$ that are globally continous, affine on each of the simplices $\Delta_m \in \mathcal{T}$, and orientation preserving. That is, on each $\Delta_m \subset \mathcal{T}$, the map $G \in \mathcal{A}_\mathcal{T}$ can be written as follows:

$$G(\omega) = A_m \omega + b_m \quad \text{for all } \omega \in \Delta_m, \tag{3.2}$$

with a suitable matrix $A_m \in \mathbb{R}^{d \times d}$ of positive determinant and a vector $b_m \in \mathbb{R}^d$.

For the calculations that follow, we shall use a more geometric way to describe the maps $G \in \mathcal{A}_\mathcal{T}$, namely by the positions $G_\ell = G(\omega_\ell)$ of the images of each node ω_ℓ . Denote by $(\mathbb{R}^d)_\mathcal{T}^L \subset \mathbb{R}^{L \cdot d}$ the space of L -tuples $\mathbf{G} = (G_\ell)_{\ell=1}^L$ of points $G_\ell \in \mathbb{R}^d$ with the same simplicial combinatorics (including orientation) as the ω_ℓ in \mathcal{T} . Clearly, any $G \in \mathcal{A}_\mathcal{T}$ is uniquely characterized by the L -tuple \mathbf{G} of its values, and moreover, any $\mathbf{G} \in (\mathbb{R}^d)_\mathcal{T}^L$ defines a $G \in \mathcal{A}_\mathcal{T}$.

More explicitly, fix a $\Delta_m \in \mathcal{T}$, with nodes labelled $\omega_{m,0}$ to $\omega_{m,d}$ in some orientation preserving order, and respective image points $G_{m,0}$ to $G_{m,d}$. With the standard d -simplex given by

$$\Delta^d := \left\{ \xi = (\xi_1, \dots, \xi_d) \in \mathbb{R}_{\geq 0}^d; \sum_{j=1}^d \xi_j \leq 1 \right\},$$

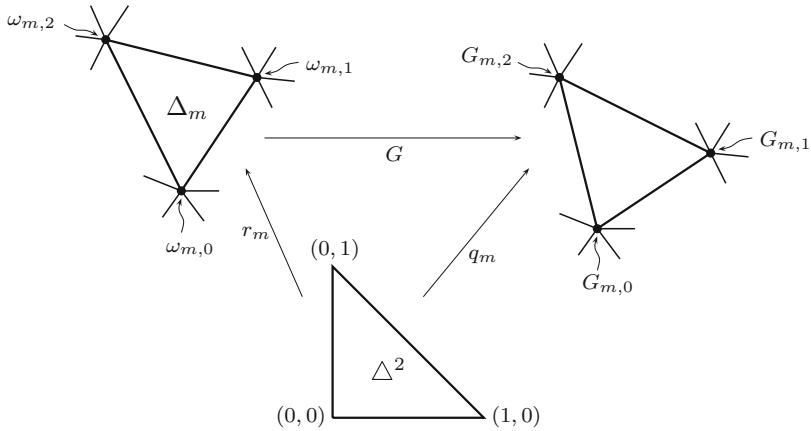


Fig. 1 A schematic representation of the spatial discretization for the case $d = 2$. Note that the upper-left triangle Δ_m is part of the reference triangulation \mathcal{T} , and is fixed in time. In contrast, the upper-right triangulation given by G_m will change with time; see Sect. 3.2

introduce the linear interpolation maps $r_m : \Delta^d \rightarrow K$ and $q_m : \Delta^d \rightarrow \mathbb{R}^d$ by

$$r_m(\xi) = \omega_{m,0} + \sum_{j=1}^d (\omega_{m,j} - \omega_{m,0}) \xi_j,$$

$$q_m(\xi) = G_{m,0} + \sum_{j=1}^d (G_{m,j} - G_{m,0}) \xi_j.$$

Then the affine map (3.2) equals to $q_m \circ r_m^{-1}$; this is shown schematically in Fig. 1 for the case $d = 2$. In particular, we obtain that

$$\det A_m = \frac{\det Dq_m}{\det Dr_m} = \frac{\det Q_{\mathcal{T}}^m[G]}{2|\Delta_m|} \quad \text{where} \quad Q_{\mathcal{T}}^m[G] := (G_{m,1} - G_{m,0} \mid \cdots \mid G_{m,d} - G_{m,0}). \tag{3.3}$$

For later reference, we give a more explicit representation for the transformed entropy \mathbf{E} for $G \in \mathcal{A}_{\mathcal{T}}$, and for the L^2 -distance between two maps $G, \hat{G} \in \mathcal{A}_{\mathcal{T}}$. Substitution of the special form (3.2) into (1.9) produces

$$\mathbf{E}(G|\bar{\rho}_{\mathcal{T}}) = \sum_{\Delta_m \in \mathcal{T}} \mu_{\mathcal{T}}^m [\mathbb{H}_{\mathcal{T}}^m(G) + \mathbb{V}_{\mathcal{T}}^m(G)] \tag{3.4}$$

with the internal energy [recall the definition of \tilde{h} from (1.9)]

$$\mathbb{H}_{\mathcal{T}}^m(G) := \tilde{h} \left(\frac{\det A_m}{\bar{\rho}_{\mathcal{T}}^m} \right) = \tilde{h} \left(\frac{\det Q_{\mathcal{T}}^m[G]}{2\mu_{\mathcal{T}}^m} \right)$$

and the potential energy

$$\mathbb{V}_{\mathcal{T}}^m(G) = \int_{\Delta_m} V(A_m \omega + b_m) \, d\omega = \int_{\Delta} V(r_m(\omega)) \, d\omega.$$

For the L^2 -difference of G and G^* , we have

$$\|G - G^*\|_{L^2(K; \bar{\rho}_{\mathcal{T}})}^2 = \int_K \|G - G^*\|^2 \bar{\rho}_{\mathcal{T}} \, d\omega = \sum_{\Delta_m \in \mathcal{T}} \mu_{\mathcal{T}}^m \mathbb{I}_{\mathcal{T}}^m(G, G^*). \tag{3.5}$$

Using Lemma B.1, we obtain on each simplex Δ_m :

$$\begin{aligned} \mathbb{I}_{\mathcal{T}}^m(G, G^*) &:= \int_{\Delta_m} \|G(\omega) - G^*(\omega)\|^2 \, d\omega \\ &= \int_{\Delta} \|r_m(\omega) - r_m^*(\omega)\|^2 \, d\omega \\ &= \frac{2}{(d+1)(d+2)} \sum_{0 \leq i < j \leq d} (G_{m,i} - G_{m,i}^*) \cdot (G_{m,j} - G_{m,j}^*). \end{aligned} \tag{3.6}$$

3.2 Discretization in Time

Let a time step $\tau > 0$ be given; in the following, we symbolize the spatio-temporal discretization by \boxplus , and we write $\boxplus \rightarrow 0$ for the joint limit of $\tau \rightarrow 0$ and vanishing mesh size in \mathcal{T} .

The discretization in time is performed in accordance with (2.5): we modify \mathbf{E}_{τ} from (2.5) by restriction to the ansatz space $\mathcal{A}_{\mathcal{T}}$. This leads to the minimization problem

$$\begin{aligned} G_{\boxplus}^n &:= \operatorname{argmin}_{G \in \mathcal{A}_{\mathcal{T}}} \mathbf{E}_{\boxplus}(G; G_{\boxplus}^{n-1}) \quad \text{where} \quad \mathbf{E}_{\boxplus}(G; G^*) \\ &= \frac{1}{2\tau} \|G - G^*\|_{L^2(K; \bar{\rho}_{\mathcal{T}})}^2 + \mathbf{E}(G | \bar{\rho}_{\mathcal{T}}). \end{aligned} \tag{3.7}$$

For a fixed discretization \boxplus , the fully discrete scheme is well-posed in the sense that for a given initial map $G_{\boxplus}^0 \in \mathcal{A}_{\mathcal{T}}$, an associated sequence $(G_{\boxplus}^n)_{n \geq 0}$ can be determined by successive solution of the minimization problems (3.7). One only needs to verify:

Lemma 3.1 *For each given $G^* \in \mathcal{A}_{\mathcal{T}}$, there exists at least one global minimizer $G \in \mathcal{A}_{\mathcal{T}}$ of $\mathbf{E}_{\boxplus}(\cdot; G^*)$.*

Remark 3.2 We do not claim uniqueness of the minimizers. Unfortunately, the minimization problem (3.7) inherits the lack of convexity from (2.5), whereas the correspondence between (2.5) and the convex problem (2.4) is lost under spatial discretization. A detailed discussion of \mathbf{E}_{\boxplus} 's (non-)convexity is provided in ‘‘Appendix C’’.

Proof of Lemma 3.1 We only sketch the main arguments. For definiteness, let us choose (just for this proof) one of the infinitely many equivalent norm-induced metrics on the dL -dimensional vector space $V_{\mathcal{T}}$ of all continuous maps $G : K \rightarrow \mathbb{R}^d$ that are piecewise affine with respect to the fixed simplicial decomposition \mathcal{T} : given $G, G' \in V_{\mathcal{T}}$ with their respective point locations $\mathbf{G}, \mathbf{G}' \in \mathbb{R}^{dL}$, i.e., $\mathbf{G} = (G_{\ell})_{\ell=1}^L$ for $G_{\ell} = G(\omega_{\ell})$, define the distance between these maps as the maximal \mathbb{R}^d -distance $\|G_{\ell} - G'_{\ell}\|$ of corresponding points $G_{\ell} \in \mathbf{G}, G'_{\ell} \in \mathbf{G}'$. Clearly, this metric makes $V_{\mathcal{T}}$ a complete space.

It is easily seen that the subset $\mathcal{A}_{\mathcal{T}}$ — which is singled out by requiring orientation preservation of the G 's—is an open subset of $V_{\mathcal{T}}$. It is further obvious that the map $G \mapsto \mathbf{E}_{\boxplus}(G; G^*)$ is continuous with respect to the metric. The claim of the lemma thus follows if we can show that the sub-level

$$S_c := \{G \in \mathcal{A}_{\mathcal{T}} ; \mathbf{E}_{\boxplus}(G; G^*) \leq c\} \quad \text{with} \quad c := \mathbf{E}(G^* | \bar{\rho}_{\mathcal{T}})$$

is a non-empty compact subset of $V_{\mathcal{T}}$. Clearly, $G^* \in S_c$, so it suffices to verify compactness.

S_c is bounded. We are going to show that there is a radius $R > 0$ such that no $G \in S_c$ has a distance larger than R to G^* . From non-negativity of \mathbf{E} , and from the representations (3.5) and (3.6), it follows that

$$\begin{aligned} c &\geq \frac{1}{2\tau} \|G - G^*\|_{L^2(K;\bar{\rho}_{\mathcal{T}})}^2 \geq \frac{\mu_{\mathcal{T}}}{2\tau} \sum_{\Delta_m \in \mathcal{T}} \mathbb{L}_{\mathcal{T}}^m(G, G^*) \\ &= \frac{\mu_{\mathcal{T}}}{(d+1)(d+2)\tau} \sum_{0 \leq i \leq j \leq d} (G_{m,i} - G_{m,i}^*) \cdot (G_{m,j} - G_{m,j}^*) \\ &\geq \frac{\mu_{\mathcal{T}}}{2(d+1)(d+2)\tau} \sum_{\ell=1}^L \|G_{\ell} - G_{\ell}^*\|^2, \end{aligned}$$

where $\mu_{\mathcal{T}} = \min_{\Delta_m} \mu_{\mathcal{T}}^m$. It is now easy to compute a suitable value for the radius R .

S_c is a closed subset of $V_{\mathcal{T}}$. It suffices to show that the limit $\bar{G} \in V_{\mathcal{T}}$ of any sequence $(G^{(k)})_{k=1}^{\infty}$ of maps $G^{(k)} \in S_c$ belongs to $\mathcal{A}_{\mathcal{T}}$. By definition of our metric on $V_{\mathcal{T}}$, global continuity and piecewise linearity of the $G^{(k)}$ trivially pass to the limit \bar{G} . We still need to verify that \bar{G} is orientation-preserving. Fix a simplex Δ_m and consider the corresponding matrices $A_m^{(k)}$ and \bar{A}_m from (3.2). Since the $G^{(k)}$ converge to \bar{G} in the metric, also $A_m^{(k)} \rightarrow \bar{A}_m$ entry-wise. Now, by non-negativity of \tilde{h} , we have for all k that

$$c \geq \mathbf{E}(G^{(k)} | \bar{\rho}_{\mathcal{T}}) \geq \mu_{\mathcal{T}}^m \tilde{h} \left(\frac{\det A_m^{(k)}}{\bar{\rho}_{\mathcal{T}}^m} \right),$$

and since $\tilde{h}(s) \rightarrow +\infty$ as $s \downarrow 0$, it follows that $\det A_m^{(k)} > 0$ is bounded away from zero, uniformly in k . But then also $\det \bar{A}_m > 0$, i.e., the m th linear map piece of the limit \bar{G} preserves orientation. \square

3.3 Fully Discrete Equations

We shall now derive the Euler–Lagrange equations associated to the minimization problem (3.7), i.e., for each given $G^* := G_{\boxplus}^{n-1} \in \mathcal{A}_{\mathcal{T}}$, we calculate the variations of $\mathbf{E}_{\boxplus}(G; G^*)$ with respect to the degrees of freedom of $G \in \mathcal{A}_{\mathcal{T}}$. Since that function is a weighted sum over the triangles $\Delta_m \in \mathcal{T}$, it suffices to perform the calculations for one fixed triangle Δ_m , with respective nodes $\omega_{m,0}$ to $\omega_{m,d}$, in positive orientation. The associated image points are $G_{m,0}$ to $G_{m,d}$. Since we may choose any vertex to be labelled $\omega_{m,0}$, it will suffice to perform the calculations at one fixed image point $G_{m,0}$.

- *mass term:*

$$\begin{aligned} \frac{\partial}{\partial G_{m,0}} \mathbb{L}_{\mathcal{T}}^m(G, G^*) &= \frac{2}{(d+1)(d+2)} \frac{\partial}{\partial G_{m,0}} \sum_{0 \leq i \leq j \leq d} (G_{m,i} - G_{m,i}^*) \cdot (G_{m,j} - G_{m,j}^*) \\ &= \frac{2}{(d+1)(d+2)} \left(2(G_{m,0} - G_{m,0}^*) + \sum_{j=1}^d (G_{m,j} - G_{m,j}^*) \right) \end{aligned}$$

- *internal energy:* observing that—recall (1.2)—

$$\tilde{h}'(s) = \frac{d}{ds} [sh(s^{-1})] = h(s^{-1}) - s^{-1}h'(s^{-1}) = -P(s^{-1}), \tag{3.8}$$

we obtain

$$\frac{\partial}{\partial G_{m,0}} \mathbb{H}_{\mathcal{T}}^m(G) = \frac{\partial}{\partial G_{m,0}} \tilde{h} \left(\frac{\det Q_{\mathcal{T}}^m[G]}{2\mu_{\mathcal{T}}^m} \right) = \frac{1}{2\mu_{\mathcal{T}}^m} P \left(\frac{2\mu_{\mathcal{T}}^m}{\det Q_{\mathcal{T}}^m[G]} \right) v_{\mathcal{T}}^m[G],$$

where

$$v_{\mathcal{T}}^m[G] := -\frac{\partial}{\partial G_{m,0}} \det Q_{\mathcal{T}}^m[G] = (\det Q_{\mathcal{T}}^m[G]) (Q_{\mathcal{T}}^m[G])^{-T} \sum_{j=1}^d e_j \tag{3.9}$$

is the uniquely determined vector in \mathbb{R}^d that is orthogonal to the $(d - 1)$ -simplex with corners $G_{m,1}$ to $G_{m,d}$ (pointing away from $G_{m,0}$) and whose length equals the $(d - 1)$ -volume of that simplex.

- *potential energy:*

$$\frac{\partial}{\partial G_{m,0}} \mathbb{V}_{\mathcal{T}}^m(G) = \frac{\partial}{\partial G_{m,0}} \int_{\Delta} V(r_m(\xi)) \, d\xi = \int_{\Delta} \nabla V(r_m(\xi)) (1 - \xi_1 - \dots - \xi_d) \, d\xi.$$

Now let ω_ℓ be a fixed vertex of \mathcal{T} . Summing over all simplices Δ_m that have ω_ℓ as a vertex, and choosing vertex labels in accordance with above, i.e., such that $\omega_{m,0} = \omega_\ell$ in Δ_m , produces the following Euler–Lagrange equation:

$$0 = \sum_{\omega_\ell \in \Delta_m} \mu_{\mathcal{T}}^m \left[\frac{1}{(d+1)(d+2)\tau} \left(2(G_{m,0} - G_{m,0}^*) + \sum_{j=1}^d (G_{m,j} - G_{m,j}^*) \right) + \frac{1}{2\mu_{\mathcal{T}}^m} P \left(\frac{2\mu_{\mathcal{T}}^m}{\det Q_{\mathcal{T}}^m[G]} \right) v_{\mathcal{T}}^m[G] + \int_{\Delta} \nabla V(r_m(\xi)) (1 - \xi_1 - \dots - \xi_d) \, d\xi \right]. \tag{3.10}$$

3.4 Approximation of the Initial Condition

For the approximation $\rho_{\boxplus}^0 = (G_{\boxplus}^0)_{\#} \bar{\rho}_{\mathcal{T}}$ of the initial datum $\rho^0 = G_{\#}^0 \bar{\rho}$, we require:

- ρ_{\boxplus}^0 converges to ρ^0 narrowly;
- $\mathcal{E}(\rho_{\boxplus}^0)$ is \boxplus -uniformly bounded, i.e.,

$$\bar{\mathcal{E}} := \sup \mathcal{E}(\rho_{\boxplus}^0) < \infty. \tag{3.11}$$

In our numerical experiments, we always choose $\bar{\rho} := \rho^0$, in which case $G^0: K \rightarrow \mathbb{R}^d$ can be taken as the identity on K , and we choose accordingly G_{\boxplus}^0 as the identity as well. Hence $\rho_{\boxplus}^0 = \bar{\rho}_{\mathcal{T}}$, which converges to $\rho^0 = \bar{\rho}$ even strongly in $L^1(K)$. Moreover, since h is convex, it easily follows from Jensen’s inequality that

$$\int_{\Delta_m} h(\bar{\rho}(x)) \, dx \geq |\Delta_m| h(\bar{\rho}_{\mathcal{T}}^m),$$

and therefore,

$$\mathcal{E}(\rho_{\boxplus}^0) \leq \mathcal{E}(\rho^0).$$

In more general situations, in which G^0 is not the identity, a sequence of approximations G_{\boxplus}^0 of G^0 is needed. Pointwise convergence $G_{\boxplus}^0 \rightarrow G^0$ is more than sufficient to guarantee narrow convergence of ρ_{\boxplus}^0 to ρ^0 , but the uniform bound (3.11) might require a well-adapted approximation, especially for non-smooth G^0 ’s.

4 Limit Trajectory

In this section, we assume that a sequence of vanishing discretizations $\boxplus \rightarrow 0$ is given, and we study the respective limit of the fully discrete solutions $(G_{\boxplus}^n)_{n \geq 0}$ that are produced by the inductive minimization procedure (3.7). For the analysis of that limit trajectory, it is more natural to work with the induced densities and velocities,

$$\rho_{\boxplus}^n := (G_{\boxplus}^n)_{\#} \bar{\rho}, \quad \mathbf{v}_{\boxplus}^n := \frac{\text{id} - G_{\boxplus}^{n-1} \circ (G_{\boxplus}^n)^{-1}}{\tau},$$

instead of the Lagrangian maps G_{\boxplus}^n themselves. Note that \mathbf{v}_{\boxplus}^n is only well-defined on the support of ρ_{\boxplus}^n —that is, on the image of G_{\boxplus}^n —and can be assigned arbitrary values outside. Let us introduce the piecewise constant in time interpolations $\tilde{\rho}_{\boxplus}: [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}_{\geq 0}$, and $\tilde{\mathbf{v}}_{\boxplus}: [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ as usual,

$$\tilde{\rho}_{\boxplus}(t) = \rho_{\boxplus}^n, \quad \tilde{\mathbf{v}}_{\boxplus}(t) = \mathbf{v}_{\boxplus}^n \quad \text{with } n \text{ such that } t \in ((n - 1)\tau, n\tau].$$

Note that $\tilde{\rho}(t, \cdot) \in \mathcal{P}_2^{\text{ac}}(\mathbb{R}^d)$ and $\tilde{\mathbf{v}}_{\boxplus}(t, \cdot) \in L^2(\mathbb{R}^d \rightarrow \mathbb{R}^d; \tilde{\rho}_{\boxplus}(t, \cdot))$ at each $t \geq 0$.

4.1 Energy Estimates

We start by proving the classical energy estimates on minimizing movements for our fully discrete scheme.

Lemma 4.1 *For each discretization \boxplus and for any indices $\bar{n} > \underline{n} \geq 0$, one has the a priori estimate*

$$\mathcal{E}(\rho_{\boxplus}^{\bar{n}}) + \frac{\tau}{2} \sum_{n=\underline{n}+1}^{\bar{n}} \left(\frac{W_2(\rho_{\boxplus}^n, \rho_{\boxplus}^{n-1})}{\tau} \right)^2 \leq \mathcal{E}(\rho_{\boxplus}^{\underline{n}}). \tag{4.1}$$

Consequently:

- (1) \mathbf{E} is monotonically decreasing, i.e., $\mathcal{E}(\tilde{\rho}_{\boxplus}(t)) \leq \mathcal{E}(\tilde{\rho}_{\boxplus}(s))$ for all $t \geq s \geq 0$;
- (2) $\tilde{\rho}_{\boxplus}$ is Hölder- $\frac{1}{2}$ -continuous in W_2 , up to an error τ ,

$$W_2(\tilde{\rho}_{\boxplus}(t), \tilde{\rho}_{\boxplus}(s)) \leq \sqrt{2\mathcal{E}(\rho_{\boxplus}^0)}(|t - s|^{\frac{1}{2}} + \tau^{\frac{1}{2}}) \quad \text{for all } t \geq s \geq 0. \tag{4.2}$$

- (3) $\tilde{\mathbf{v}}_{\boxplus}$ is square integrable with respect to $\tilde{\rho}_{\boxplus}$,

$$\int_0^T \int_{\mathbb{R}^d} \|\tilde{\mathbf{v}}_{\boxplus}\|^2 \tilde{\rho}_{\boxplus} \, dx \, dt \leq 2\mathcal{E}(\rho_{\boxplus}^0). \tag{4.3}$$

Proof By the definition of G_{\boxplus}^n as a minimizer, we know that $\mathbf{E}_{\boxplus}(G_{\boxplus}^n; G_{\boxplus}^{n-1}) \leq \mathbf{E}_{\boxplus}(G; G_{\boxplus}^{n-1})$ for any $G \in \mathcal{A}_{\mathcal{F}}$, and in particular for the choice $G := G_{\boxplus}^{n-1}$, which yields:

$$\frac{1}{2\tau} \int_K \|G_{\boxplus}^n - G_{\boxplus}^{n-1}\|^2 \bar{\rho}_{\mathcal{F}} \, d\omega + \mathbf{E}(G_{\boxplus}^n | \bar{\rho}_{\mathcal{F}}) \leq \mathbf{E}(G_{\boxplus}^{n-1} | \bar{\rho}_{\mathcal{F}}). \tag{4.4}$$

Summing these inequalities for $n = \underline{n} + 1, \dots, \bar{n}$, recalling that $\mathcal{E}(\rho_{\boxplus}^{\underline{n}}) = \mathbf{E}(G_{\boxplus}^{\underline{n}} | \bar{\rho}_{\mathcal{F}})$ by (1.9) and that $W_2(\rho_{\boxplus}^n, \rho_{\boxplus}^{n-1})^2 \leq \int_K |G_{\boxplus}^n - G_{\boxplus}^{n-1}|^2 \bar{\rho} \, d\omega$ by (2.3), produces (4.1).

Monotonicity of \mathcal{E} in time is obvious.

To prove (4.2), choose $\underline{n} \leq \bar{n}$ such that $s \in ((\underline{n} - 1)\tau, \underline{n}\tau]$ and $t \in ((\bar{n} - 1)\tau, \bar{n}\tau]$. Notice that $\tau(\bar{n} - \underline{n}) \leq t - s + \tau$. If $\underline{n} = \bar{n}$, the claim (4.2) is obviously true; let $\underline{n} < \bar{n}$ in the following. Combining the triangle inequality for the metric W_2 , estimate (4.1) above and Hölder’s inequality for sums, we arrive at

$$\begin{aligned}
 W_2(\tilde{\rho}_{\boxplus}(t), \tilde{\rho}_{\boxplus}(s)) &= W_2(\rho_{\boxplus}^{\bar{n}}, \rho_{\boxplus}^{\underline{n}}) \leq \sum_{n=\underline{n}+1}^{\bar{n}} W_2(\rho_{\boxplus}^n, \rho_{\boxplus}^{n-1}) \\
 &\leq \left[\sum_{n=\underline{n}+1}^{\bar{n}} \tau \right]^{\frac{1}{2}} \left[\sum_{n=\underline{n}+1}^{\bar{n}} \frac{W_2(\rho_{\boxplus}^n, \rho_{\boxplus}^{n-1})^2}{\tau} \right]^{\frac{1}{2}} \\
 &= [\tau(\bar{n} - \underline{n})]^{\frac{1}{2}} \left[\tau \sum_{n=\underline{n}+1}^{\bar{n}} \left(\frac{W_2(\rho_{\boxplus}^n, \rho_{\boxplus}^{n-1})}{\tau} \right)^2 \right]^{\frac{1}{2}} \\
 &\leq [t - s + \tau]^{\frac{1}{2}} [2(\mathcal{E}(\rho_{\boxplus}^{\bar{n}}) - \mathcal{E}(\rho_{\boxplus}^{\underline{n}}))]^{\frac{1}{2}} \\
 &\leq [|t - s|^{\frac{1}{2}} + \tau^{\frac{1}{2}}] \mathcal{E}(\rho_{\boxplus}^0)^{\frac{1}{2}}.
 \end{aligned}$$

Finally, changing variables using $x = G_{\boxplus}^n(\omega)$ in (4.4) yields

$$\frac{\tau}{2} \int_{\mathbb{R}^d} \|\mathbf{v}_{\boxplus}^n\|^2 \rho_{\boxplus}^n dx + \mathbf{E}(G_{\boxplus}^n) \leq \mathbf{E}(G_{\boxplus}^{n-1}),$$

and summing these inequalities from $n = 1$ to $n = N_\tau$ yields (4.3). □

4.2 Compactness of the Trajectories and Weak Formulation

Our main result on the weak limit of $\tilde{\rho}_{\boxplus}$ is the following.

Theorem 4.2 *Along a suitable sequence $\boxplus \rightarrow 0$, the curves $\tilde{\rho}_{\boxplus}: \mathbb{R}_{\geq 0} \rightarrow \mathcal{P}_2^{ac}(\mathbb{R}^d)$ convergence pointwise in time, i.e., $\tilde{\rho}_{\boxplus}(t) \rightarrow \rho_*(t)$ narrowly for each $t > 0$, towards a Hölder- $\frac{1}{2}$ -continuous limit trajectory $\rho_*: \mathbb{R}_{\geq 0} \rightarrow \mathcal{P}_2^{ac}(\mathbb{R}^d)$.*

Moreover, the discrete velocities $\tilde{\mathbf{v}}_{\boxplus}$ possess a limit $\mathbf{v}_ \in L^2(\mathbb{R}_{\geq 0} \times \mathbb{R}^d; \rho_*)$ such that $\tilde{\mathbf{v}}_{\boxplus} \tilde{\rho}_{\boxplus} \xrightarrow{*} \mathbf{v}_* \rho_*$ in $L^1(\mathbb{R}_{\geq 0} \times \mathbb{R}^d)$, and the continuity equation*

$$\partial_t \rho_* + \nabla \cdot (\rho_* \mathbf{v}_*) = 0 \tag{4.5}$$

holds in the sense of distributions.

Remark 4.3 The Hölder continuity of ρ_* implies that ρ_* satisfies the initial condition (1.1b) in the sense that $\rho_*(t) \rightarrow \rho^0$ narrowly as $t \downarrow 0$.

Proof of Theorem 4.2 We closely follow an argument that is part of the general convergence proof for the minimizing movement scheme as given in Ambrosio et al. [1, Section 11.1.3]. Below, convergence is shown for some arbitrary but fixed time horizon $T > 0$; a standard diagonal argument implies convergence at arbitrary times.

First observe that by estimate (4.2)—applied with $0 = s \leq t \leq T$ —it follows that $W_2(\tilde{\rho}_{\boxplus}(t), \rho_{\boxplus}^0)$ is bounded, uniformly in $t \in [0, T]$ and in \boxplus . Since further ρ_{\boxplus}^0 converges narrowly to ρ^0 by our hypotheses on the initial approximation, we conclude that all densities $\tilde{\rho}_{\boxplus}(t)$ belong to a sequentially compact subset for the narrow convergence. The second observation is that the term on the right hand side of (4.2) simplifies to $(2\bar{\mathcal{E}})^{\frac{1}{2}} |t - s|^{\frac{1}{2}}$ in the limit $\boxplus \rightarrow 0$. A straightforward application of the “refined version” of the Ascoli-Arzelà theorem (Proposition 3.3.1 in Ambrosio et al. [1]) yields the first part of the claim, namely the pointwise narrow convergence of $\tilde{\rho}_{\boxplus}$ towards a Hölder continuous limit curve ρ_* .

It remains to pass to the limit with the velocity $\tilde{\mathbf{v}}_{\boxplus}$. Towards that end, we define a probability measure $\tilde{\gamma}_{\boxplus} \in \mathcal{P}(Z_T)$ on the set $Z_T := [0, T] \times \mathbb{R}^d \times \mathbb{R}^d$ as follows:

$$\int_{Z_T} \varphi(t, x, v) \, d\tilde{\gamma}_{\boxplus}(t, x, v) = \int_0^T \int_{\mathbb{R}^d} \varphi(t, x, \tilde{\mathbf{v}}_{\boxplus}(t, x)) \tilde{\rho}_{\boxplus}(t, x) \, dx \frac{dt}{T},$$

for every bounded and continuous function $\varphi \in C_b^0(Z_T)$. For brevity, let $\tilde{M}_{\boxplus} \in \mathcal{P}([0, T] \times \mathbb{R}^d)$ be the (t, x) -marginals of $\tilde{\gamma}_{\boxplus}$, that have respective Lebesgue densities $\frac{\rho_{\boxplus}(t, x)}{T}$ on $[0, T] \times \mathbb{R}^d$. Thanks to the result from the first part of the proof, \tilde{M}_{\boxplus} converges narrowly to a limit M_* , which has density $\frac{\rho_*(t, x)}{T}$. On the other hand, the estimate (4.3) implies that

$$\int_{Z_T} |v|^2 \, d\tilde{\gamma}_{\boxplus}(t, x, v) = \int_{[0, T] \times \mathbb{R}^d} |\tilde{\mathbf{v}}_{\boxplus}(t, x)|^2 \, d\tilde{M}_{\boxplus}(t, x) \leq 2\bar{\epsilon}.$$

We are thus in the position to apply Theorem 5.4.4 in Ambrosio et al. [1], which yields the narrow convergence of $\tilde{\gamma}_{\boxplus}$ towards a limit γ_* . Clearly, the (t, x) -marginal of γ_* is M_* . Accordingly, we introduce the disintegration $\gamma_{(t, x)}$ of γ_* with respect to M_* , which is well-defined M_* -a.e.. Below, it will turn out that γ_* 's v -barycenter,

$$\mathbf{v}_*(t, x) := \int_{\mathbb{R}^d} v \, d\gamma_{(t, x)}(v), \tag{4.6}$$

is the sought-for weak limit of $\tilde{\mathbf{v}}_{\boxplus}$. The convergence $\tilde{\mathbf{v}}_{\boxplus} \tilde{\rho}_{\boxplus} \xrightarrow{*} \mathbf{v}_* \rho_*$ and the inheritance of the uniform L^2 -bound (4.3) to the limit \mathbf{v}_* are further direct consequences of Theorem 5.4.4 in Ambrosio et al. [1].

The key step to establish the continuity equation for the just-defined \mathbf{v}_* is to evaluate the limit as $\boxplus \rightarrow 0$ of

$$J_{\boxplus}[\phi] := \frac{1}{\tau} \left[\int_0^T \int_{\mathbb{R}^d} \phi(t, x) \tilde{\rho}_{\boxplus}(t, x) \, dx \, dt - \int_0^T \int_{\mathbb{R}^d} \phi(t, x) \tilde{\rho}_{\boxplus}(t - \tau, x) \, dx \, dt \right]$$

for any given test function $\phi \in C_c^\infty((0, T) \times \mathbb{R}^d)$ in two different ways. First, we change variables $t \mapsto t + \tau$ in the second integral, which gives

$$\begin{aligned} J_{\boxplus}[\phi] &= \int_0^T \int_{\mathbb{R}^d} \frac{\phi(t, x) - \phi(t + \tau, x)}{\tau} \tilde{\rho}_{\boxplus}(t, x) \, dx \, dt \xrightarrow{\boxplus \rightarrow 0} \\ &\quad - \int_0^T \int_{\mathbb{R}^d} \partial_t \phi(t, x) \rho_*(t, x) \, dx \, dt. \end{aligned}$$

For the second evaluation, we write

$$\rho_{\boxplus}^{n-1} = (G_{\boxplus}^{n-1} \circ (G_{\boxplus}^n)^{-1})_{\#} \rho_{\boxplus}^n = (\text{id} - \tau \mathbf{v}_{\boxplus}^n)_{\#} \rho_{\boxplus}^n,$$

and substitute accordingly $x \mapsto x - \tau \tilde{\mathbf{v}}_{\boxplus}(t, x)$ in the second integral, leading to

$$\begin{aligned} J_{\boxplus}[\phi] &= \int_0^T \int_{\mathbb{R}^d} \frac{\phi(t, x) - \phi(t, x - \tau \tilde{\mathbf{v}}_{\boxplus}(t, x))}{\tau} \tilde{\rho}_{\boxplus}(t, x) \, dx \, dt \\ &= \int_0^T \int_{\mathbb{R}^d} \nabla \phi(t, x) \cdot \tilde{\mathbf{v}}_{\boxplus}(t, x) \tilde{\rho}_{\boxplus}(t, x) \, dx \, dt + \epsilon_{\boxplus}[\phi] \\ &= \int_{Z_T} \nabla \phi(t, x) \cdot v \, d\tilde{\gamma}_{\boxplus}(t, x, v) + \epsilon_{\boxplus}[\phi] \\ &\xrightarrow{\boxplus \rightarrow 0} \int_{Z_T} \nabla \phi(t, x) \cdot v \, d\gamma_*(t, x, v) \end{aligned}$$

$$\begin{aligned}
 &= \int_{[0,T] \times \mathbb{R}^d} \nabla \phi(t, x) \cdot \left[\int_{\mathbb{R}^d} v \, d\gamma_{(t,x)}(v) \right] \, dM_*(t, x) \\
 &= \int_0^T \int_{\mathbb{R}^d} \nabla \phi(t, x) \cdot \mathbf{v}_*(t, x) \rho_*(t, x) \, dx \, dt.
 \end{aligned}$$

The error term $\epsilon_{\boxplus}[\phi]$ above is controlled via Taylor expansion of ϕ and by using (4.3),

$$|\epsilon_{\boxplus}[\phi]| \leq \int_0^T \int_{\mathbb{R}^d} \frac{\tau}{2} \|\phi\|_{C^2} \|\tilde{\mathbf{v}}_{\boxplus}(t, x)\|^2 \tilde{\rho}_{\boxplus}(t, x) \, dx \, dt \leq \bar{\mathcal{E}} \|\phi\|_{C^2} T \tau.$$

Equality of the limits for both evaluations of $J_{\boxplus}[\phi]$ for arbitrary test functions ϕ shows the continuity Eq. (4.5). □

Unfortunately, the convergence provided by Theorem 4.2 is generally not sufficient to conclude that ρ_* is a weak solution to (1.1), since we are not able to identify \mathbf{v}_* as $\mathbf{v}[\rho_*]$ from (1.4b). The problem is two-fold: first, weak- \star convergence of $\tilde{\rho}_{\boxplus}$ is insufficient to pass to the limit inside the nonlinear function P . Second, even if we would know that, for instance, $P(\tilde{\rho}_{\boxplus}) \xrightarrow{*} P(\rho_*)$, we would still need a \boxplus -independent a priori control on the regularity (e.g., maximal diameter of triangles) of the meshes generated by the G_{\boxplus}^n to justify the passage to limit in the weak formulation below.

The main difficulty in the weak formulation that we derive now is that we can only use “test functions” that are piecewise affine with respect to the changing meshes generated by the G_{\boxplus}^n . For definiteness, we introduce the space

$$\mathcal{D}(\mathcal{T}) := \left\{ \Gamma : K \rightarrow \mathbb{R}^d ; \Gamma \text{ is globally continuous, and is piecewise affine w.r.t. } \Delta_m \right\}.$$

Lemma 4.4 *Assume $S : \mathbb{R}^d \rightarrow \mathbb{R}^d$ is such that $S \circ G_{\boxplus}^n \in \mathcal{D}(\mathcal{T})$. Then:*

$$\int_{\mathbb{R}^d} P(\rho_{\boxplus}^n) \nabla \cdot S \, dx - \int_{\mathbb{R}^d} \nabla V \cdot S \rho_{\boxplus}^n \, dx = \int_{\mathbb{R}^d} S \cdot \mathbf{v}_{\boxplus}^n \rho_{\boxplus}^n \, dx. \tag{4.7}$$

Proof For all sufficiently small $\varepsilon > 0$, let $G_\varepsilon = (\text{id} + S) \circ G_{\boxplus}^n$. By definition of G_{\boxplus}^n as a minimizer, we have that $\mathbf{E}_{\boxplus}(G_\varepsilon; G_{\boxplus}^{n-1}) \geq \mathbf{E}_{\boxplus}(G_{\boxplus}^n; G_{\boxplus}^{n-1})$. This implies that

$$\begin{aligned}
 0 &\leq \frac{1}{\varepsilon} \int_K \left(\frac{1}{2\tau} [\|G_\varepsilon - G_{\boxplus}^{n-1}\|^2 - \|G_{\boxplus}^n - G_{\boxplus}^{n-1}\|^2] \right. \\
 &\quad \left. + \left[\tilde{h} \left(\frac{\det DG_\varepsilon}{\bar{\rho}_{\mathcal{T}}} \right) - \tilde{h} \left(\frac{\det DG_{\boxplus}^n}{\bar{\rho}_{\mathcal{T}}} \right) \right] + [V \circ G_\varepsilon - V] \right) \bar{\rho}_{\mathcal{T}} \, d\omega. \tag{4.8}
 \end{aligned}$$

We discuss limits of the three terms under the integral for $\varepsilon \searrow 0$. For the metric term:

$$\begin{aligned}
 \frac{1}{2\tau\varepsilon} \left[\|G_\varepsilon - G_{\boxplus}^{n-1}\|^2 - \|G_{\boxplus}^n - G_{\boxplus}^{n-1}\|^2 \right] &= \frac{G_{\boxplus}^n - G_{\boxplus}^{n-1}}{\tau} \cdot \frac{G_\varepsilon - G_{\boxplus}^n}{\varepsilon} + \frac{1}{2\tau\varepsilon} \|G_\varepsilon - G_{\boxplus}^n\|^2 \\
 &= \left[\left(\frac{\text{id} - T_{\boxplus}^n}{\tau} \right) \cdot S \right] \circ G_{\boxplus}^n + \frac{\varepsilon}{2\tau} \|S\|^2 \circ G_{\boxplus}^n,
 \end{aligned}$$

and since S is bounded, the last term vanishes uniformly on K for $\varepsilon \searrow 0$. For the internal energy, since $DG_\varepsilon = D(\text{id} + \varepsilon S) \circ G_{\boxplus}^n \cdot DG_{\boxplus}^n$, and recalling (3.8),

$$\begin{aligned}
 & \frac{1}{\varepsilon} \left[\tilde{h} \left(\frac{\det DG_\varepsilon}{\bar{\rho}_{\mathcal{T}}} \right) - \tilde{h} \left(\frac{\det DG_{\boxplus}^n}{\bar{\rho}_{\mathcal{T}}} \right) \right] \\
 &= \frac{1}{\varepsilon} \left[\tilde{h} \left(\frac{\det DG_{\boxplus}^n}{\bar{\rho}_{\mathcal{T}}} \det(\mathbb{1} + \varepsilon DS) \circ G_{\boxplus}^n \right) - \tilde{h} \left(\frac{\det DG_{\boxplus}^n}{\bar{\rho}_{\mathcal{T}}} \right) \right] \\
 &\xrightarrow{\varepsilon \searrow 0} \frac{\det DG_{\boxplus}^n}{\bar{\rho}_{\mathcal{T}}} \tilde{h}' \left(\frac{\det DG_{\boxplus}^n}{\bar{\rho}_{\mathcal{T}}} \right) \left(\lim_{\varepsilon \searrow 0} \frac{\det(\mathbb{1} + \varepsilon DS)}{\varepsilon} \right) \circ G_{\boxplus}^n \\
 &= -\frac{\det DG_{\boxplus}^n}{\bar{\rho}_{\mathcal{T}}} P \left(\frac{\bar{\rho}_{\mathcal{T}}}{\det DG_{\boxplus}^n} \right) \text{tr}[DS] \circ G_{\boxplus}^n \\
 &= -\frac{\det DG_{\boxplus}^n}{\bar{\rho}_{\mathcal{T}}} [P(\rho^n) \nabla \cdot S] \circ G_{\boxplus}^n.
 \end{aligned}$$

Since the piecewise constant function $\det DG_{\boxplus}^n$ has a positive lower bound, the convergence as $\varepsilon \searrow 0$ is uniform on K . Finally, for the potential energy,

$$\frac{1}{\varepsilon} [V \circ (\text{id} + \varepsilon S) \circ G_{\boxplus}^n - V \circ G_{\boxplus}^n] \xrightarrow{\varepsilon \searrow 0} [\nabla V \cdot S] \circ G_{\boxplus}^n.$$

Again, the convergence is uniform on K . Passing to the limit in the integral (4.8) yields

$$\begin{aligned}
 0 \leq & \int_K \left[\left(\frac{\text{id} - T_{\boxplus}^n}{\tau} \right) \cdot S \right] \circ G_{\boxplus}^n \bar{\rho}_{\mathcal{T}} \, d\omega \\
 & - \int_K [P(\rho^n) \nabla \cdot S] \circ G_{\boxplus}^n \det DG_{\boxplus}^n \, d\omega + \int_K [\nabla V \cdot S] \circ G_{\boxplus}^n \bar{\rho}_{\mathcal{T}} \, d\omega.
 \end{aligned}$$

The same inequality is true with $-S$ in place of S , hence this inequality is actually an equality. Since $\rho_{\boxplus}^n = (G_{\boxplus}^n)_{\#} \bar{\rho}_{\mathcal{T}}$, a change of variables $x = S_{\boxplus}^n(\omega)$ produces (4.7). \square

Corollary 4.5 *In addition to the hypotheses of Theorem 4.2, assume that*

- (1) $P(\tilde{\rho}_{\boxplus}^n) \xrightarrow{*} p_*$ in $L^1([0, T] \times \Omega)$;
- (2) each G_{\boxplus}^n is injective;
- (3) as $\boxplus \rightarrow 0$, all simplices in the images of \mathcal{T} under G_{\boxplus}^n have non-degenerate interior angles and tend to zero in diameter, uniformly w.r.t. n .

Then ρ_* satisfies the PDE

$$\partial_t \rho_* = \Delta p_* + \nabla \cdot (\rho_* \nabla V) \tag{4.9}$$

in the sense of distributions.

Proof Let a smooth test function $\zeta \in C_c^\infty(\mathbb{R}^d \rightarrow \mathbb{R}^d)$ be given. For each \boxplus and each n , a $\zeta_{\boxplus}^n : \mathbb{R}^d \rightarrow \mathbb{R}^d$ with $\zeta_{\boxplus}^n \circ G_{\boxplus}^n \in \mathcal{D}(\mathcal{T})$ can be constructed in such a way that

$$\zeta_{\boxplus}^n \rightarrow \zeta, \quad \nabla \cdot \zeta_{\boxplus}^n \rightarrow \nabla \cdot \zeta \tag{4.10}$$

uniformly on \mathbb{R}^d , and uniformly in n as $\boxplus \rightarrow 0$. This follows from our hypotheses on the \boxplus -uniform regularity of the Lagrangian meshes: inside the image of G_{\boxplus}^n , one can simply choose ζ_{\boxplus}^n as the affine interpolation of the values of ζ at the points $G_{\boxplus}^n(\omega_\ell)$. Outside, one can take an arbitrary approximation of ζ that is compatible with the piecewise-affine approximation on the boundary of G_{\boxplus}^n 's image; one may even choose $\zeta_{\boxplus}^n \equiv \zeta$ at sufficient distance to that boundary. The uniform convergences (4.10) then follow by standard finite element analysis.

Further, let $\eta \in C_c^\infty(0, T)$ be given. For each $t \in ((n - 1)\tau, n\tau]$, substitute $S(t, x) := \eta(t)\zeta_{\boxplus}^n(x)$ into (4.7). Integration of these equalities with respect to $t \in (0, T)$ yields

$$\int_0^T \int_{\mathbb{R}^d} P(\tilde{\rho}_{\boxplus}) \nabla \cdot S \, dx \, dt - \int_0^T \int_{\mathbb{R}^d} \nabla V \cdot S \, dx \, dt = \int_0^T \int_{\mathbb{R}^d} S \cdot \tilde{\mathbf{v}}_{\boxplus} \tilde{\rho}_{\boxplus} \, dx \, dt.$$

We pass to the limit $\boxplus \rightarrow 0$ in these integrals. For the first, we use that $P(\tilde{\rho}) \xrightarrow{*} p_*$ by hypothesis, for the last, we use Theorem 4.2 above. Since any test function $S \in C_c^\infty((0, T) \times \Omega)$ can be approximated in C^1 by linear combinations of products $\eta(t)\zeta(x)$ as above, we thus obtain the weak formulation of

$$\rho_* \mathbf{v}_* = \nabla p_* + \rho_* \nabla V.$$

In combination with the continuity Eq. (4.5), we arrive at (4.9). □

Remark 4.6 In principle, our discretization can also be applied to the linear Fokker–Planck equation with $P(r) = r$ and $h(r) = r \log r$. In that case, one automatically has $P(\tilde{\rho}) \xrightarrow{*} p_* \equiv P(\rho_*)$ thanks to Theorem 4.2. Corollary 4.5 above then provides an *a posteriori* criterion for convergence: if the Lagrangian mesh does not deform too wildly under the dynamics as the discretization is refined, then the discrete solutions converge to the genuine solution.

5 Consistency in 2D

In this section, we prove consistency of our discretization in the following sense. Under certain conditions on the spatial discretization \mathcal{T} , any smooth and positive solution ρ to the initial value problem (1.1) projects to a discrete solution that satisfies the Euler–Lagrange equations up to a controlled error. We restrict ourselves to $d = 2$ dimensions.

5.1 Smooth Lagrangian Evolution

First, we derive an alternative form of the velocity field \mathbf{v} from (1.4b) in terms of G .

Lemma 5.1 *For $\rho = G_{\#}\bar{\rho}$ with a smooth diffeomorphism $G : K \rightarrow \mathbb{R}^d$, we have*

$$\mathbf{v}[\rho] \circ G = \mathbf{V}[G] := P' \left(\frac{\bar{\rho}}{\det DG} \right) (DG)^{-T} \left(\text{tr}_{12} [(DG)^{-1} D^2 G]^T - \frac{\nabla \bar{\rho}}{\bar{\rho}} \right) - \nabla V \circ G. \tag{5.1}$$

Consequently, the Lagrangian map G —relative to the reference density $\bar{\rho}$ —for a smooth solution ρ to (1.1) satisfies

$$\partial_t G = \mathbf{V}[G]. \tag{5.2}$$

Proof On the one hand,

$$D[h'(\rho) \circ G] = [Dh'(\rho)] \circ G DG,$$

and on the other hand, by definition of the push forward,

$$\begin{aligned} D[h'(\rho) \circ G] &= Dh' \left(\frac{\bar{\rho}}{\det DG} \right) \\ &= h'' \left(\frac{\bar{\rho}}{\det DG} \right) \left(\frac{\bar{\rho}}{\det DG} \right) \left(\frac{D\bar{\rho}}{\bar{\rho}} - \text{tr}_{12} [(DG)^{-1}D^2G] \right) \\ &= [\rho h''(\rho)] \circ G \left(\frac{D\bar{\rho}}{\bar{\rho}} - \text{tr}_{12} [(DG)^{-1}D^2G] \right). \end{aligned}$$

Hence

$$\nabla h'(\rho) \circ G = [\rho h''(\rho)] \circ G (DG)^{-T} \left(\frac{\nabla \bar{\rho}}{\bar{\rho}} - \text{tr}_{12} [(DG)^{-1}D^2G]^T \right).$$

Observing that (1.2) implies that $rh''(r) = P'(r)$, we conclude (5.2) directly from (1.4b). □

5.2 Discrete Euler–Lagrange Equations in Dimension $d = 2$

In the planar case $d = 2$, the Euler–Lagrange equation (3.10) above can be rewritten in a more convenient way.

In the following, fix some vertex ω_\times of the triangulation, which is incident to precisely six triangles. For convenience, we assume that these are labelled Δ_0 to Δ_5 in counter-clockwise order. Similarly, the six neighboring vertices are labeled ω_0 to ω_5 in counter-clockwise order, so that Δ_k has vertices ω_k and ω_{k+1} , where we set $\omega_6 := \omega_0$.

Using these conventions and recalling Lemma B.2, the expression for the vector v in (3.9) simplifies to

$$v_{\mathcal{T}}^k = -\mathbb{J}(G_{k+1} - G_k), \quad \text{where } \mathbb{J} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Summing the Euler–Lagrange equation (3.10) over Δ_0 to Δ_5 , we obtain

$$\mathbf{p}_\times = \mathbf{J}_\times, \tag{5.3}$$

where the momentum term \mathbf{p}_\times and the impulse \mathbf{J}_\times , respectively, are given by

$$\mathbf{p}_\times = \frac{1}{12} \sum_{k=0}^5 \mu_{\mathcal{T}}^k \left[2 \left(\frac{G_\times - G_\times^*}{\tau} \right) + \left(\frac{G_k - G_k^*}{\tau} \right) + \left(\frac{G_{k+1} - G_{k+1}^*}{\tau} \right) \right] \tag{5.4}$$

$$\mathbf{J}_\times = \sum_{k=0}^5 \mu_{\mathcal{T}}^k \left[\frac{1}{2\mu_{\mathcal{T}}^k} P \left(\frac{2\mu_{\mathcal{T}}^k}{\det(G_k - G_\times | G_{k+1} - G_\times)} \right) \mathbb{J}(G_{k+1} - G_k) \right. \tag{5.5}$$

$$\left. - \int_{\Delta} \nabla V((1 - \xi_1 - \xi_2)G_\times + \xi_1 G_k + \xi_2 G_{k+1}) (1 - \xi_1 - \xi_2) d\xi \right]. \tag{5.6}$$

We shall now prove our main result on consistency. The setup is the following: a sequence of triangulations \mathcal{T}_ε on K , parametrized by $\varepsilon > 0$, and a sequence of time steps $\tau_\varepsilon = \mathcal{O}(\varepsilon)$ are given. We assume that there is an ε -independent region $K' \subset K$ on which the \mathcal{T}_ε are *almost hexagonal* in the following sense: each node $\omega_\times \in K'$ of \mathcal{T}_ε has precisely six neighbors—labelled ω_0 to ω_5 in counter-clockwise order—and there exists a rotation $R \in \text{SO}(2)$ such that

$$R(\omega_k - \omega_\times) = \varepsilon \sigma_k + \mathcal{O}(\varepsilon^2) \quad \text{with } \sigma_k = \begin{pmatrix} \cos \frac{\pi}{3} k \\ \sin \frac{\pi}{3} k \end{pmatrix} \tag{5.7}$$

for $k = 0, 1, \dots, 5$.

Now, let $G : [0, T] \times K \rightarrow \mathbb{R}^d$ be a given smooth solution to the Lagrangian evolution Eq. (5.2), and fix a time $t \in (0, T)$. For all sufficiently small $\varepsilon > 0$, we define maps $G_\varepsilon, G_\varepsilon^* \in \mathcal{A}_{\mathcal{T}_\varepsilon}$ by linear interpolation of the values of $G(t; \cdot)$ and $G(t - \tau; \cdot)$, respectively, on \mathcal{T}_ε . That is, $G_\varepsilon(\omega_\ell) = G(t; \omega_\ell)$ and $G_\varepsilon^*(\omega_\ell) = G(t - \tau; \omega_\ell)$, at all nodes ω_ℓ in \mathcal{T}_ε . Theorem 5.2 below states that the pair $G_\varepsilon, G_\varepsilon^*$ is an approximate solution to the discrete Euler–Lagrange equations (5.3) at all nodes ω_\times of the respective triangulation \mathcal{T}_ε that lie in K' .

The hexagonality hypothesis on the \mathcal{T}_ε is strong, but some very strong restriction of $\mathcal{A}_{\mathcal{T}_\varepsilon}$'s geometry is apparently necessary. See Remark 5.4 following the proof for further discussion.

Theorem 5.2 *Under the hypotheses and with the notations introduced above, the Euler–Lagrange equation (5.3) admits the following asymptotic expansion:*

$$\mathbf{p}_\times = \frac{\sqrt{3}}{2} \varepsilon^2 \bar{\rho}(\omega_\times) \partial_t G(t; \omega_\times) + \mathcal{O}(\varepsilon^3), \tag{5.8a}$$

$$\mathbf{J}_\times = \frac{\sqrt{3}}{2} \varepsilon^2 \bar{\rho}(\omega_\times) \mathbf{V}[G](t; \omega_\times) + \mathcal{O}(\varepsilon^3), \tag{5.8b}$$

as $\varepsilon \rightarrow 0$, uniformly at the nodes $\omega_\times \in K'$ of the respective \mathcal{T}_ε .

Remark 5.3 Up to an error $\mathcal{O}(\varepsilon^3)$, the geometric pre-factor $\frac{\sqrt{3}}{2} \varepsilon^2$ equals to one third of the total area of the hexagon with vertices ω_0 to ω_5 , and is thus equal to the integral of the piecewise affine hat function with peak at ω_\times .

Proof of Theorem 5.2 Throughout the proof, let $\varepsilon > 0$ be fixed; we shall omit the ε -index for \mathcal{T}_ε and τ_ε . First, we fix a node ω_\times of $\mathcal{T} \cap K'$. Thanks to the equivariance of both (5.2) and (5.3) under rigid motions of the domain, we may assume that R in (5.7) is the identity, and that $\omega_\times = 0$.

We collect some relations that are helpful for the calculations that follow. Trivially,

$$\sum_{k=0}^5 \sigma_k = 0, \quad \sum_{k=0}^5 \omega_k = \mathcal{O}(\varepsilon^2). \tag{5.9}$$

Moreover, we have that

$$|\Delta_k| = \det(\omega_k | \omega_{k+1}) = \varepsilon^2 \det(\sigma_k | \sigma_{k+1}) + \mathcal{O}(\varepsilon^3) = \frac{\sqrt{3}}{4} \varepsilon^2 + \mathcal{O}(\varepsilon^3). \tag{5.10}$$

On the other hand, by definition of $\mu_{\mathcal{T}}^k$ in (3.1), it follows that

$$\begin{aligned} \mu_{\mathcal{T}}^k &= |\Delta_k| \int_{\Delta_k} \bar{\rho} \, d\omega = \frac{1}{2} \det(\omega_k | \omega_{k+1}) \left[\bar{\rho} \left(\frac{\omega_k + \omega_{k+1}}{3} \right) + \mathcal{O}(\varepsilon) \right] \\ &= \frac{1}{2} \det(\omega_k | \omega_{k+1}) \left[\bar{\rho}_\times + \varepsilon \nabla \bar{\rho}_\times \cdot \frac{\sigma_k + \sigma_{k+1}}{3} + \mathcal{O}(\varepsilon^2) \right]. \end{aligned} \tag{5.11}$$

Combining (5.10) and (5.11) yields

$$\mu_{\mathcal{T}}^k = \varepsilon^2 \left(\frac{\sqrt{3}}{4} \bar{\rho}_\times + \mathcal{O}(\varepsilon) \right). \tag{5.12}$$

In accordance with the definition of G_ε and G_ε^* from G detailed above, let $G_\times := G(t, \omega_\times)$ and $G_\times^* = G(t - \tau, \omega_\times)$, and define G_k, G_k^* for $k = 0, \dots, 5$ in the analogous way. Further, we introduce $DG_\times = DG(t, \omega_\times)$, $D^2G_\times = D^2G(t, \omega_\times)$, $\partial_t G_\times = \partial_t G(t, \omega_\times)$.

To perform an expansion in the *momentum term*, first observe that

$$G(t - \tau; \omega_k) = G(t; \omega_k) - \tau \partial_t G(t; \omega_k) + \mathcal{O}(\tau^2),$$

for each $k = 0, 1, \dots, 5$, and so, using that $\tau = \mathcal{O}(\varepsilon)$ by hypothesis,

$$\frac{G_k - G_k^*}{\tau} = \partial_t G(t; \omega_k) + \mathcal{O}(\tau) = \partial_t G_\times + \mathcal{O}(\varepsilon) + \mathcal{O}(\tau) = \partial_t G_\times + \mathcal{O}(\varepsilon).$$

Using (5.12) and then (5.9) yields

$$\begin{aligned} \mathbf{p}_\times &= \frac{1}{12\tau} \sum_{k=0}^5 \varepsilon^2 \left(\frac{\sqrt{3}}{4} \bar{\rho}_\times + \mathcal{O}(\varepsilon) \right) [4\partial_t G_\times + \mathcal{O}(\varepsilon)] \\ &= \frac{\sqrt{3}}{2} \varepsilon^2 \bar{\rho}_\times \partial_t G_\times + \mathcal{O}(\varepsilon^3). \end{aligned}$$

This is (5.8a).

For the *impulse term*, we start with a Taylor expansion to second order in space:

$$G_k = G_\times + DG_\times \omega_k + \frac{1}{2} D^2 G_\times : [\omega_k]^2 + \mathcal{O}(\varepsilon^3).$$

We combine this with the observation that $(\omega_k | \omega_{k+1})^{-1} = \mathcal{O}(\varepsilon^{-1})$ to obtain:

$$\begin{aligned} & \frac{\mu_{\mathcal{F}}^k}{\det(G_k - G_\times | G_{k+1} - G_\times)} \\ &= \frac{\det(\omega_k | \omega_{k+1})}{\det DG_\times} \frac{\bar{\rho}_\times + \varepsilon \nabla \bar{\rho}_\times \cdot \frac{\sigma_k + \sigma_{k+1}}{3} + \mathcal{O}(\varepsilon^2)}{\det \left[(\omega_k | \omega_{k+1}) + \frac{1}{2} (DG_\times)^{-1} (D^2 G_\times : [\omega_k]^2 | D^2 G_\times : [\omega_{k+1}]^2) + \mathcal{O}(\varepsilon^3) \right]} \\ &= \frac{\bar{\rho}_\times}{\det DG_\times} \frac{1 + \varepsilon \frac{\nabla \bar{\rho}_\times \cdot \sigma_k + \sigma_{k+1}}{\bar{\rho}_\times} + \mathcal{O}(\varepsilon^2)}{\det \left[\mathbb{1} + \frac{1}{2} (DG_\times)^{-1} (D^2 G_\times : [\omega_k]^2 | D^2 G_\times : [\omega_{k+1}]^2) (\omega_k | \omega_{k+1})^{-1} + \mathcal{O}(\varepsilon^2) \right]} \\ &= \frac{\bar{\rho}_\times}{\det DG_\times} \left(1 + \varepsilon \left\{ \chi_k - \frac{1}{2} \vartheta_k \right\} + \mathcal{O}(\varepsilon^2) \right), \end{aligned}$$

where

$$\begin{aligned} \chi_k &= \frac{\nabla \bar{\rho}_\times \cdot \sigma_k + \sigma_{k+1}}{\bar{\rho}_\times} \cdot \frac{1}{3}, \\ \vartheta_k &= \text{tr} \left[((DG_\times)^{-1} D^2 G_\times : [\sigma_k]^2 | (DG_\times)^{-1} D^2 G_\times : [\sigma_{k+1}]^2) (\sigma_k | \sigma_{k+1})^{-1} \right]. \end{aligned}$$

Plugging this in leads to

$$\begin{aligned} & \sum_{k=0}^5 \left\{ \frac{1}{2} P \left(\frac{\bar{\rho}_\times}{\det DG_\times} \right) + \frac{\varepsilon}{2} P' \left(\frac{\bar{\rho}_\times}{\det DG_\times} \right) \left\{ \chi_k - \frac{1}{2} \vartheta_k \right\} + \mathcal{O}(\varepsilon^2) \right\} \mathbb{J} DG_\times (\omega_{k+1} - \omega_k) \\ &= \frac{1}{2} P \left(\frac{\bar{\rho}_0}{\det DG_\times} \right) \mathbb{J} DG_\times \left(\sum_{k=0}^5 (\omega_{k+1} - \omega_k) \right) \\ & \quad + \frac{\varepsilon^2}{4} P' \left(\frac{\bar{\rho}_\times}{\det DG_\times} \right) \mathbb{J} DG_\times \mathbb{J}^T \left(\sum_{k=0}^5 \{2\chi_k - \vartheta_k\} \mathbb{J} (\sigma_{k+1} - \sigma_k) \right) + \mathcal{O}(\varepsilon^3) \\ &= 0 + \frac{\sqrt{3}}{2} \varepsilon^2 P' \left(\frac{\bar{\rho}_\times}{\det DG_\times} \right) (DG_\times)^{-T} \left\{ \text{tr}_{12} [(DG_\times)^{-1} D^2 G_\times]^T - \frac{\nabla \rho_\times}{\rho_\times} \right\} + \mathcal{O}(\varepsilon^3), \end{aligned}$$

where we have use the auxiliary algebraic results from Lemmas B.2, B.3 and B.4.

For the remaining part of the impulse term, a very rough approximation is sufficient:

$$\nabla V(g) = \nabla V(G_\times) + \mathcal{O}(\varepsilon)$$

holds for any g that is a convex combination of $G_\times, G_0, \dots, G_5$, where the implicit constant is controlled in terms of the supremum of D^2V and DG on K' . With that, we simply have, using again (5.12):

$$\begin{aligned} & \sum_{k=0}^5 \mu_{\mathcal{T}}^k \int_{\Delta} \nabla V((1 - \xi_1 - \xi_2)G_\times + \xi_1 G_k + \xi_2 G_{k+1}) (1 - \xi_1 - \xi_2) d\xi \\ &= 6\varepsilon^2 \left(\frac{\sqrt{3}}{4} \bar{\rho}_\times + \mathcal{O}(\varepsilon) \right) (\nabla V(G_\times) + \mathcal{O}(\varepsilon)) = \frac{\sqrt{3}}{2} \varepsilon^2 \bar{\rho}_\times \nabla V(G_\times) + \mathcal{O}(\varepsilon^3). \end{aligned}$$

Together, this yields (5.8b). □

Remark 5.4 The hypotheses of Theorem (5.2) require that the \mathcal{T}_ε are almost hexagonal on K' . This seems like a technical hypothesis that simplifies calculations, but apparently, *some* strong symmetry property of the \mathcal{T}_ε is necessary for the validity of the result.

To illustrate the failure of consistency—at least in the specific form considered here—assume that $V \equiv 0$ and $\bar{\rho} \equiv 1$, and consider a sequence of triangulations \mathcal{T}_ε for which there is a node ω_\times such that (5.7) holds with the σ_k being replaced by a different six-tuple of vectors σ'_k . Repeating the steps of the proof above, it is easily seen that $\mathbf{p}_\times = a\varepsilon^2 \vartheta_t G(t; \omega_\times) + \mathcal{O}(\varepsilon^3)$, with an ε -independent constant $a > 0$ in place of $\sqrt{3}/2$, and that

$$\mathbf{J}_\times = -\frac{\varepsilon^2}{4} P' \left(\frac{1}{\det DG_\times} \right) (DG_\times)^{-T} \sum_{k=0}^5 \vartheta'_k \mathbb{J}(\sigma'_{k+1} - \sigma'_k) + \mathcal{O}(\varepsilon^3),$$

with

$$\vartheta'_k = \text{tr} \left[((DG_\times)^{-1} D^2 G_\times : [\sigma'_k]^2 | (DG_\times)^{-1} D^2 G_\times : [\sigma'_{k+1}]^2) (\sigma'_k | \sigma'_{k+1})^{-1} \right].$$

If a result of the form (5.8b)—with $\sqrt{3}/2$ replaced by a —was true, then this implies in particular that

$$\sum_{k=0}^5 \vartheta'_k \mathbb{J}(\sigma'_{k+1} - \sigma'_k) = a' \text{tr}_{12} [(DG_\times)^{-1} D^2 G_\times] \tag{5.13}$$

holds with some constant $a' > 0$ for arbitrary matrices $DG_\times \in \mathbb{R}^{2 \times 2}$ of positive determinant and tensors $D^2 G_\times \in \mathbb{R}^{2 \times 2 \times 2}$ that are symmetric in the second and third component. A specific example for which (5.13) is not true is given by

$$\sigma'_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = -\sigma'_3, \quad \sigma'_1 = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix} = -\sigma'_4, \quad \sigma'_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} = -\sigma'_5, \tag{5.14}$$

in combination with $DG_\times = \mathbb{1}$, and a $D^2 G_\times$ that is zero except for two ones, at the positions (1, 2, 2) and (2, 1, 1). In Lemma B.5, we show that the left-hand side in (5.13) equals to $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$; on the other hand, the right-hand side is clearly zero.

Note that this counter-example is significant, insofar as the skew (in fact, degenerate) hexagon described by the σ'_k in (5.14) corresponds to a popular method for triangulation of the plane.

6 Numerical Simulations in $d = 2$

6.1 Implementation

The Euler–Lagrange equations for the $d = 2$ -dimensional case have been derived in (5.3). We perform a small modification in the potential term in order to simplify calculations with presumably minimal loss in accuracy:

$$\begin{aligned} \mathbf{Z}_\times[G; G^*] = & \sum_{k=0}^5 \frac{\mu_{\mathcal{J}}^k}{12} \left[2 \left(\frac{G_\times - G_\times^*}{\tau} \right) + \left(\frac{G_k - G_k^*}{\tau} \right) + \left(\frac{G_{k+1} - G_{k+1}^*}{\tau} \right) \right] \\ & + \sum_{k=0}^5 \left[\frac{1}{2} \tilde{h}' \left(\frac{\det(G_k - G_\times | G_{k+1} - G_\times)}{2\mu_{\mathcal{J}}^k} \right) \mathbb{J}(G_{k+1} - G_k) \right. \\ & \left. + \frac{\mu_{\mathcal{J}}^k}{6} \nabla V(G_{k+\frac{1}{2}}) \right], \end{aligned}$$

with the short-hand notation

$$G_{k+\frac{1}{2}} = \frac{1}{3}(G_\times + G_k + G_{k+1}).$$

On the main diagonal, the Hessian amounts to

$$\begin{aligned} \mathbf{H}_{\times \times}[G] = & \left(\sum_{k=0}^5 \frac{\mu_{\mathcal{J}}^k}{6\tau} \right) \mathbb{1}_2 \\ & + \sum_{k=0}^5 \frac{1}{4\mu_{\mathcal{J}}^k} \tilde{h}'' \left(\frac{\det(G_k - G_\times | G_{k+1} - G_\times)}{2\mu_{\mathcal{J}}^k} \right) [\mathbb{J}(G_{k+1} - G_k)][\mathbb{J}(G_{k+1} - G_k)]^\top \\ & + \sum_{k=0}^5 \frac{\mu_{\mathcal{J}}^k}{18} \nabla^2 V(G_{k+\frac{1}{2}}) \end{aligned}$$

Off the main diagonal, the entries of the Hessian are given by

$$\begin{aligned} \mathbf{H}_{\times k}[G] = & \frac{\mu_{\mathcal{J}}^k + \mu_{\mathcal{J}}^{k-1}}{12\tau} \mathbb{1}_2 \\ & + \frac{1}{4\mu_{\mathcal{J}}^k} \tilde{h}'' \left(\frac{\det(G_k - G_\times | G_{k+1} - G_\times)}{2\mu_{\mathcal{J}}^k} \right) [\mathbb{J}(G_{k+1} - G_k)][\mathbb{J}(G_{k+1} - G_\times)]^\top \\ & - \frac{1}{4\mu_{\mathcal{J}}^{k-1}} \tilde{h}'' \left(\frac{\det(G_{k-1} - G_\times | G_k - G_\times)}{2\mu_{\mathcal{J}}^{k-1}} \right) [\mathbb{J}(G_k - G_{k-1})][\mathbb{J}(G_{k-1} - G_\times)]^\top \\ & + \frac{\mu_{\mathcal{J}}^k}{18} \nabla^2 V(G_{k+\frac{1}{2}}) + \frac{\mu_{\mathcal{J}}^{k-1}}{18} \nabla^2 V(G_{k-\frac{1}{2}}). \end{aligned}$$

The scheme consists of an inner (Newton) and an outer (time stepping) iteration. We start from a given initial density ρ_0 and define the solution at the next time step inductively by applying Newton’s method in the inner iteration. To this end we initialise $G^{(0)} := G^n$ with G^n , the solution at the n th time step, and define inductively

$$G^{(s+1)} := G^{(s)} + \delta G^{(s+1)},$$

where the update $\delta G^{(s+1)}$ is the solution to the linear system

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

The effort of each inner iteration step is essentially determined by the effort to invert the sparse matrix $\mathbf{H}[G^{(s)}]$. As soon as the norm of $\delta G^{(s+1)}$ drops below a given stopping threshold, define $G^{n+1} := G^{(s+1)}$ as approximate solution in the $n + 1$ st time step.

In all experiments the stopping criterion in the Newton iteration is set to 10^{-9} .

6.2 Numerical Experiments

In this section we present results of our numerical experiments for (1.1) with a cubic porous-medium nonlinearity $P(r) = r^3$ and different choices for the external potential V ,

$$\partial_t \rho = \Delta(u^3) + \nabla \cdot (u \nabla V). \tag{6.1}$$

Numerical experiment 1: unconfined evolution of Barenblatt profile

As a first example, we consider the “free” cubic porous medium equation, that is (6.1) with $V \equiv 0$. It is well-known (see, e.g., Vazquez [38]) that in the long-time limit $t \rightarrow \infty$, arbitrary solutions approach a self-similar one,

$$\rho^*(t, x) = t^{-d\alpha} \mathcal{B}_3(t^{-\alpha} x) \quad \text{with} \quad \alpha = \frac{1}{6}, \tag{6.2}$$

where \mathcal{B}_3 is the associated Barenblatt profile

$$\mathcal{B}_3(z) = \left(C_3 - \frac{1}{3} \|z\|^2 \right)_+^{\frac{1}{2}}, \tag{6.3}$$

where $C_3 = (2\pi)^{-\frac{2}{3}} \approx 0.29$ is chosen to normalize \mathcal{B}_3 ’s mass to unity.

In this experiment, we are only interested in the quality of the numerical approximation for the self-similar solution (6.2). To reduce numerical effort, we impose a four-fold symmetry of the approximation: we use the quarter circle as computational domain K , and interpret the discrete function thereon as one of four symmetric pieces of the full discrete solution. To preserve reflection symmetry over time, homogeneous Neumann conditions are imposed on the artificial boundaries. This is implemented by reducing the degrees of freedom of the nodes along the x - and y -axes to tangential motion. We initialize our simulation with a piecewise constant approximation of the profile of ρ^* from (6.3) at time $t = 0.01$. We choose a time step $\tau = 0.001$ and the final time $T = 2$. In Fig. 2, we have collected snapshots of the approximated density at different instances of time. The Barenblatt profile of the solution is very well pertained over time.

Remark 6.1 It takes less than 2 min to complete this simulation on standard laptop (MATLAB code on a mid-2013 MacBook Air 11” with 1.7 GHz Intel Core i7 processor).

Figure 3 shows surface plots of the discrete solution at different times in comparison with the Barenblatt profile at the respective time. By construction of the scheme, the initial mass is exactly conserved in time as the discrete solution propagates. The left plot in Fig. 4 shows the decay in the energy and gives quantitative information about the difference of the discrete solution to the analytical Barenblatt solution. The numerical solution shows good agreement with the analytical energy decay rate $c = 2/3$.

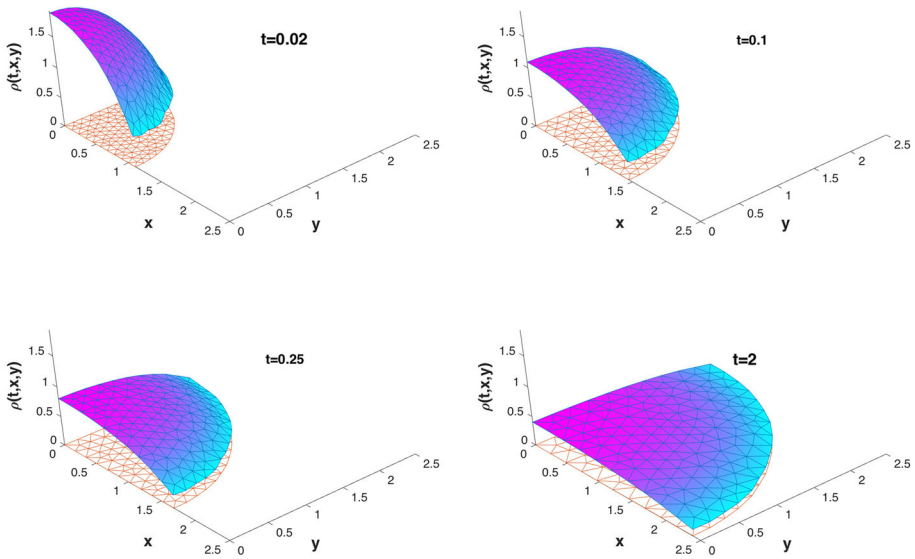


Fig. 2 Numerical experiment 1: fully discrete evolution of our approximation for the self-similar solution to the free porous medium equation. Snapshots are taken at times $t = 0.02$, $t = 0.1$, $t = 0.25$, and $t = 2.0$

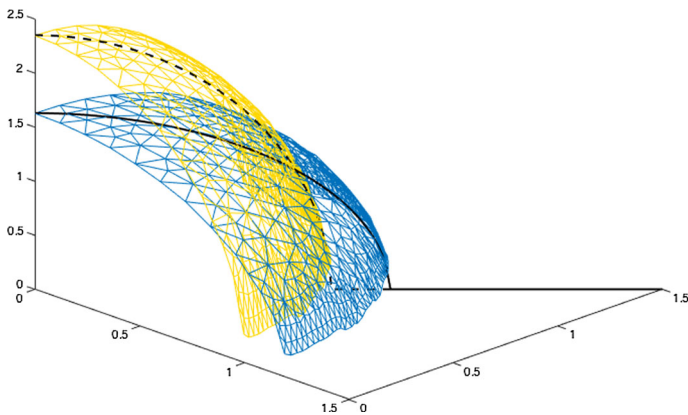


Fig. 3 Numerical experiment 1: comparison of the discrete solution (interpolated surface plots with triangulation) with the Barenblatt profile (solid and dashed black lines along the identity) at different times

We also compute the l_1 -error of the discrete solution to the exact Barenblatt profile and observe that it remains within the order of the fineness of the triangulation. The mass of the discrete solution is perfectly conserved, as guaranteed by the construction of our method.

To estimate the convergence order of our method, we run several experiments with the above initial data on different meshes. We fix the ratio $\tau/h_{\max}^2 = 0.4$ and compute the l_1 -error at time $T = 0.2$ on triangulations with $h_{\max} = 0.2, 0.1, 0.05, 0.025$. We expect the error to decay as a power of h_{\max} . The double logarithmic plot should reveal a line with its slope indicating the numerical convergence order. The right plot in Fig. 4 shows the result, the estimated numerical convergence order which is obtained from a least-squares fitted line

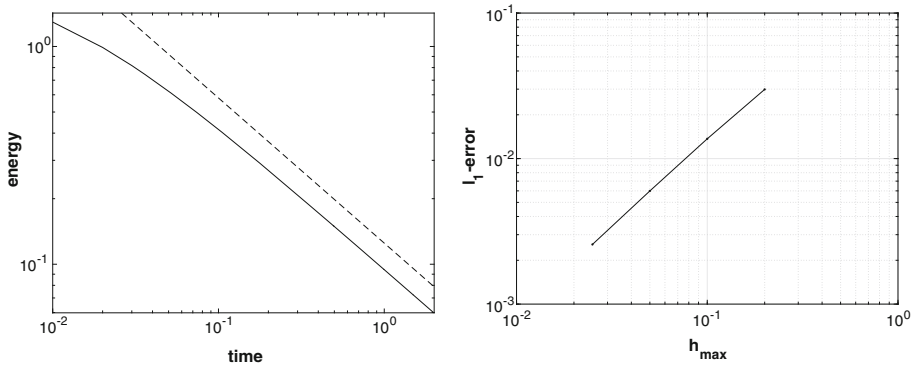


Fig. 4 Numerical experiment 1: decay of the energy of the discrete solution in comparison with the analytical decay $t^{-2/3}$ of the Barenblatt solution (left). Numerical convergence for fixed ratio $\tau/h_{\max}^2 = 0.4$ (right)

through the points is equal to 1.18. This indicates first order convergence of the scheme with respect to the spatial discretisation parameter h_{\max} .

Numerical experiment 2: Asymptotic self-similarity

In our second example, we are still concerned with the free cubic porous medium Eq. (6.1) with $V \equiv 0$. This time, we wish to give an indication that the discrete approximation of the self-similar solution from (6.2) from the previous experiment might inherit the global attractivity of its continuous counterpart. More specifically, we track the discrete evolution for the initial datum

$$\rho_0(x, y) = 3000(x^2 + y^2) \exp[-5(|x| + |y|)] + 0.1 \tag{6.4}$$

until time $T = 0.1$ and observe that it appears to approach the self-similar solution from above. Snapshots of the simulation are collected in Fig. 5.

Numerical experiment 3: two peaks merging into one under the influence of a confining potential

In this example we consider as initial condition two peaks, connected by a thin layer of mass, given by

$$\rho_0(x, y) = \exp[-20((x - 0.35)^2 + (y - 0.35)^2)] + \exp[-20((x + 0.35)^2 + (y + 0.35)^2)] + 0.001. \tag{6.5}$$

We choose a triangulation of the square $[-1.5, 1.5]^2$ and initialise the discrete solution piecewise constant in each triangle, with a value corresponding to (6.5), evaluated in the centre of mass of each triangle. We solve the porous medium equation with a confining potential, i.e. (1.1) with $P(r) = r^m$ and $V(x, y) = 5(x^2 + y^2)/2$. The time step is $\tau = 0.001$ and the final time is $T = 0.2$.

Figure 6 shows the evolution from the initial density. As time increases the peaks smoothly merge into each other. As the thin layer around the peaks is also subject to the potential the triangulated domain shrinks in time. Even if we do not know how to prevent theoretically the intersection of the images of the discrete Lagrangian maps, this seems not to be a problem in

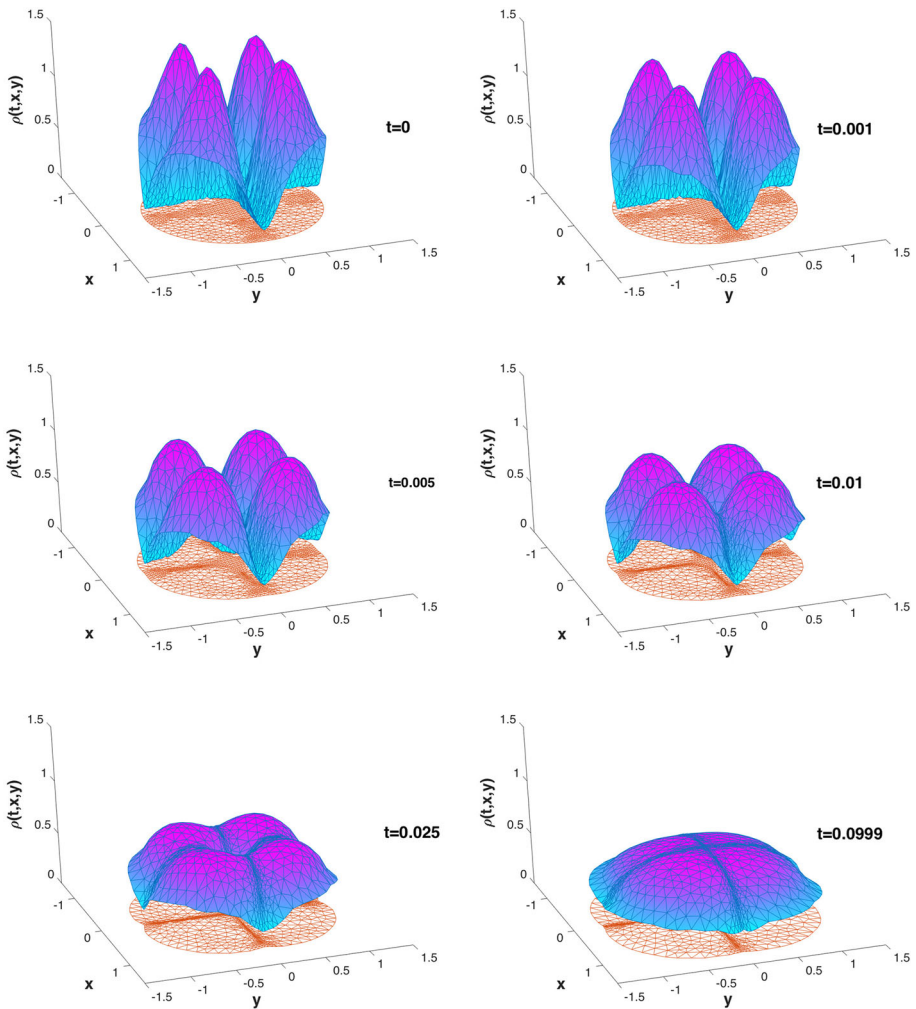


Fig. 5 Numerical experiment 2: fully discrete evolution for the initial density from (6.4) under the free porous medium equation. Snapshots are taken at times $t = 0.001, t = 0.005, t = 0.01, t = 0.025,$ and $t = 0.1$

practice. As time evolves, the discrete solution approaches the steady state Barenblatt profile given by

$$B(z) = \left(C - \frac{5}{3} ||z||^2 \right)_+^{\frac{1}{2}}, \tag{6.6}$$

where C is chosen as the mass of the density. The plot in Fig. 7 shows the exponential decay of the l_1 -distance of the discrete solution to the steady state Barenblatt profile (6.6). We observe that the decay agrees very well with the analytically predicted decay $\exp(-5t)$ until $t = 0.08$. For larger times, one would monitor triangle quality numerically, and re-mesh, locally coarsening the triangulation where necessary.

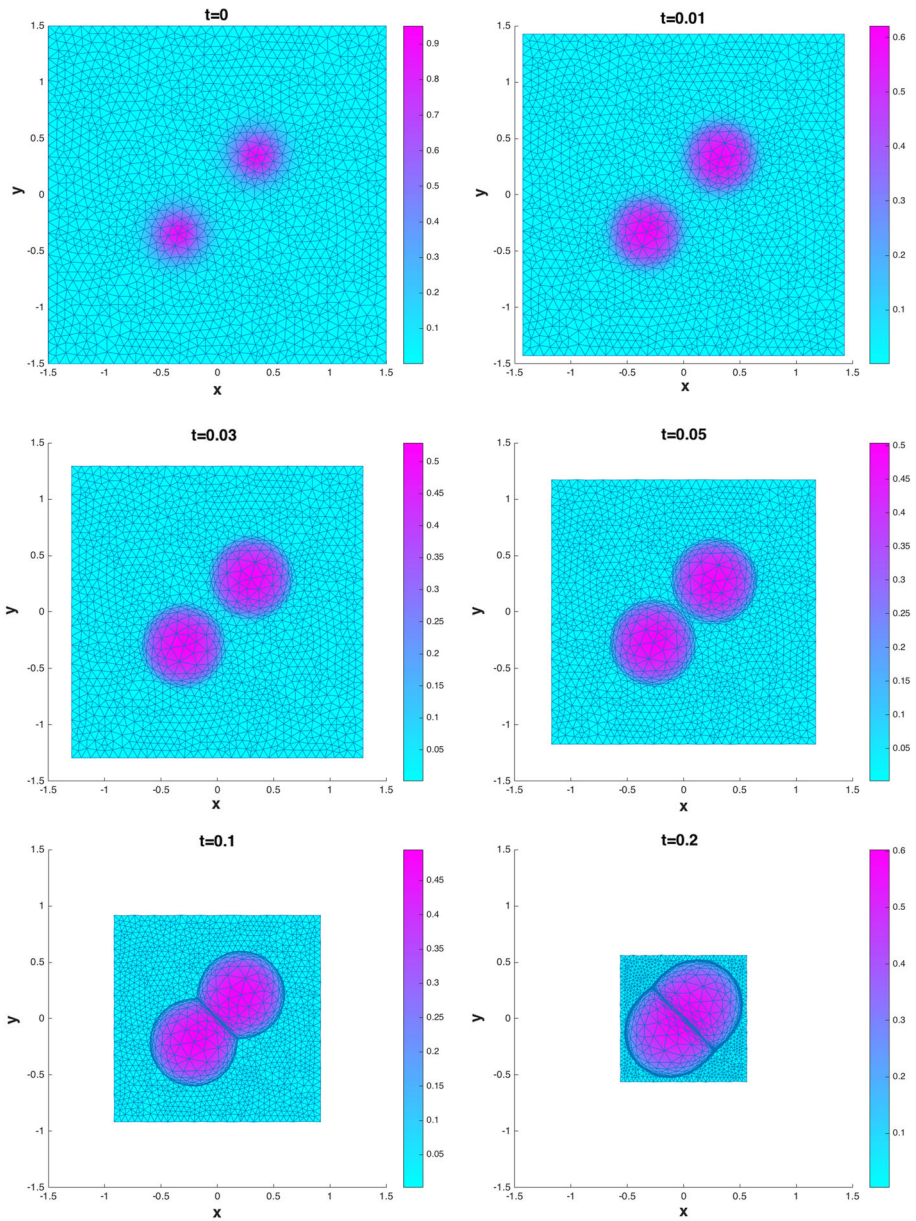


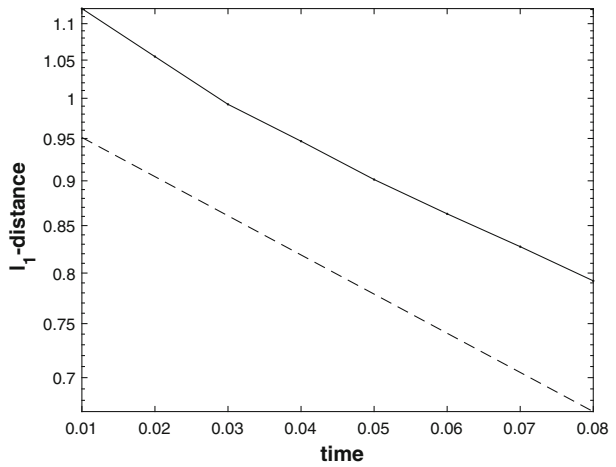
Fig. 6 Numerical experiment 3: evolution of two peaks merging under the porous medium equation with a confining potential

Numerical experiment 4: one peak splitting under the influence of a quartic potential

We consider as the initial condition

$$\rho_0(x, y) = 1 - (x^2 + y^2). \tag{6.7}$$

Fig. 7 Numerical experiment 3: two merging peaks: plot of the l_1 -distance of the discrete solution to the steady state Barenblatt profile in comparison with the analytical decay $c \exp(-5t)$



We choose a triangulation of the unit circle and initialise the discrete solution piecewise constant in each triangle, with a value corresponding to (6.7), evaluated in the centre of mass of each triangle. We solve the porous medium equation with a quartic potential, i.e. (1.1) with $P(r) = r^m$ and $V(x) = 5(x^2 + (1 - y^2)^2)/2$. The time step is $\tau = 0.005$ and the final time is $T = 0.02$.

Figure 8 shows the evolution of the initial density. As time increases the initial density is progressively split, until two new maxima emerge which are connected by a thin layer. For larger times, when certain triangles become excessively distorted, one would monitor triangle quality numerically, and re-mesh, locally refining the triangulation where necessary.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

Appendix A: Proof of the Lagrangian representation

Proof of Lemma 1.1 We verify that the density function given by $(G_t^{-1})_{\#}\rho_t$ on $K \subset \mathbb{R}^d$ is constant with respect to time t ; the identity (1.6) then follows since

$$\rho_t = (G_t \circ G_t^{-1})_{\#}\rho_t = (G_t)_{\#}[(G_t^{-1})_{\#}\rho_t] = (G_t)_{\#}[(G_0^{-1})_{\#}\rho^0] = (G_t)_{\#}\bar{\rho}.$$

Firstly, from the definition of the inverse,

$$G_t^{-1} \circ G_t = \text{id}$$

for all t , differentiating with respect to time yields

$$D(G_t^{-1}) \circ G_t \partial_t G_t + \partial_t(G_t^{-1}) \circ G_t = 0,$$

and so, using (1.5) and (1.4b),

$$\partial_t(G_t^{-1}) = -D(G_t^{-1})(\partial_t G_t \circ G_t^{-1}) = -D(G_t^{-1})\mathbf{v}[\rho_t]. \tag{A.1}$$

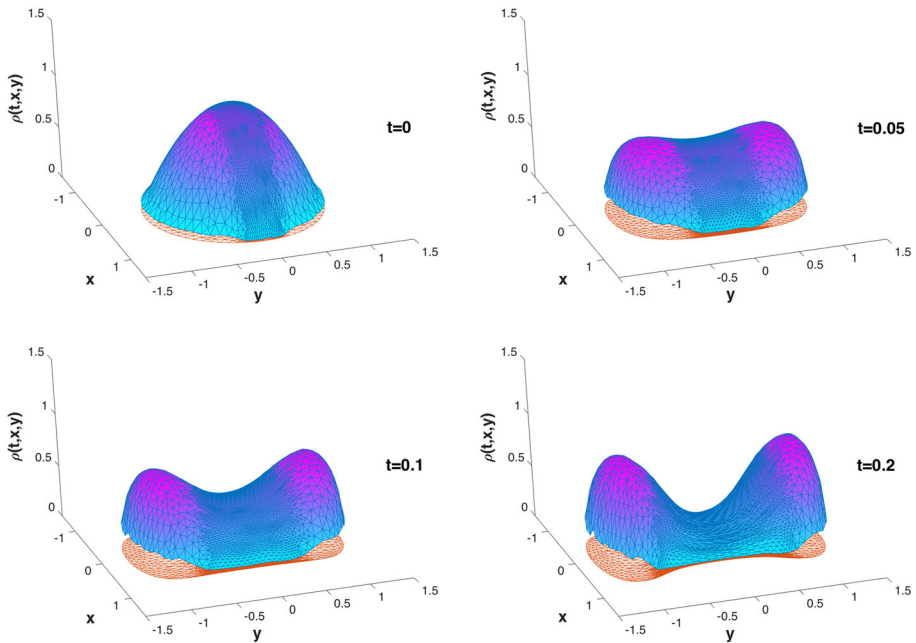


Fig. 8 Numerical experiment 4: evolution of the initial density under the porous medium equation with a quartic potential

Now, let φ be a smooth test function, and consider

$$\begin{aligned}
 \frac{d}{dt} \int \varphi (G_t^{-1})_{\#} \rho_t &= \frac{d}{dt} \int (\varphi \circ G_t^{-1}) \rho_t \\
 &= \int (\varphi \circ G_t^{-1}) \partial_t \rho_t + \int D\varphi \circ G_t^{-1} \partial_t (G_t^{-1}) \rho_t \\
 &= - \int (\varphi \circ G_t^{-1}) [\nabla \cdot (\rho_t v(\rho_t))] \\
 &\quad - \int (D\varphi \circ G_t^{-1}) D(G_t^{-1}) v(\rho_t) \rho_t && \text{by (1.1) and ((1.5))} \\
 &= \int (D\varphi \circ G_t^{-1}) D(G_t^{-1}) [v(\rho_t) - v(\rho_t)] \rho_t && \text{integrating by parts} \\
 &= 0.
 \end{aligned}$$

As φ was arbitrary, $(G_t^{-1})_{\#} \rho_t$ is constant with respect to time. □

Appendix B: Technical lemmas

Lemma B.1 Given $g_0, g_1, \dots, g_d \in \mathbb{R}^d$, then

$$\int_{\Delta^d} \left\| g_0 + \sum_{j=1}^d \omega_j (g_j - g_0) \right\|^2 d\omega = \frac{2}{(d+1)(d+2)} \sum_{0 \leq i < j \leq d} g_i \cdot g_j. \tag{B.1}$$

Proof Thanks to the symmetry of the integral with respect to the exchange of the components ω_j , the left-hand side of (B.1) equals to

$$\begin{aligned} & \|g_0\|^2 + 2 \left(\int_{\Delta} \omega_d \, d\omega \right) \sum_{1 \leq j \leq d} g_0 \cdot (g_j - g_0) \\ & + \left(\int_{\Delta} \omega_d^2 \, d\omega \right) \sum_{1 \leq j \leq d} \|g_j - g_0\|^2 + 2 \left(\int_{\Delta} \omega_{d-1} \omega_d \, d\omega \right) \sum_{1 \leq i < j \leq d} (g_i - g_0) \cdot (g_j - g_0). \end{aligned} \tag{B.2}$$

We calculate the integrals, using Fubini’s theorem. First integral:

$$\begin{aligned} \int_{\Delta} \omega_d \, d\omega &= \frac{1}{|\Delta^d|} \int_0^1 \omega_d (1 - \omega_d)^{d-1} |\Delta^{d-1}| \, d\omega_d \\ &= \frac{|\Delta^{d-1}|}{|\Delta^d|} \int_0^1 (1 - z) z^{d-1} \, dz = d \left(\frac{1}{d} - \frac{1}{d+1} \right) = \frac{1}{d+1}. \end{aligned}$$

Second integral:

$$\begin{aligned} \int_{\Delta} \omega_d^2 \, d\omega &= \frac{1}{|\Delta^d|} \int_0^1 \omega_d^2 (1 - \omega_d)^{d-1} |\Delta^{d-1}| \, d\omega_d \\ &= \frac{|\Delta^{d-1}|}{|\Delta^d|} \int_0^1 (1 - z)^2 z^{d-1} \, dz = d \left(\frac{1}{d} - \frac{2}{d+1} + \frac{1}{d+2} \right) = \frac{2}{(d+1)(d+2)}. \end{aligned}$$

Third integral:

$$\begin{aligned} \int_{\Delta} \omega_{d-1} \omega_d \, d\omega &= \frac{1}{|\Delta^d|} \int_0^1 \left[\int_0^{1-\omega_d} \omega_{d-1} \omega_d (1 - \omega_{d-1} - \omega_d)^{d-2} |\Delta^{d-2}| \, d\omega_{d-1} \right] \, d\omega_d \\ &= \frac{|\Delta^{d-2}|}{|\Delta^d|} \int_0^1 \left[\int_0^z (1 - z)(z - y) y^{d-2} \, dy \right] \, dz \\ &= d(d-1) \int_0^1 \left[\frac{1}{d-1} - \frac{1}{d} \right] (1 - z) z^d \, dz \\ &= \frac{1}{d+1} - \frac{1}{d+2} = \frac{1}{(d+1)(d+2)}. \end{aligned}$$

Substitute this into (B.2):

$$\begin{aligned} & \left(1 - \frac{2}{d+1} + \frac{d^2 + d}{(d+1)(d+2)} \right) \|g_0\|^2 + \left(\frac{2}{d+1} - \frac{2d+2}{(d+1)(d+2)} \right) \sum_{1 \leq j \leq d} g_0 \cdot g_j \\ & + \frac{2}{(d+1)(d+2)} \sum_{1 \leq j \leq d} \|g_j\|^2 + \frac{2}{(d+1)(d+2)} \sum_{1 \leq i < j \leq d} g_i \cdot g_j \\ & = \frac{2}{(d+1)(d+2)} \left(\|g_0\|^2 + \sum_{1 \leq j \leq d} g_0 \cdot g_j + \sum_{1 \leq j \leq d} \|g_j\|^2 + \sum_{1 \leq i < j \leq d} g_i \cdot g_j \right). \end{aligned}$$

Collecting terms yields the right-hand side of (B.1). □

Lemma B.2 For each $A \in \mathbb{R}^{2 \times 2}$, we have $\mathbb{J}A\mathbb{J}^T = (\det A) A^{-T}$.

Proof This is verified by direct calculation:

$$\mathbb{J}A\mathbb{J}^T = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} a_{22} & -a_{21} \\ -a_{12} & a_{11} \end{pmatrix} = (\det A) A^{-T}.$$

□

Lemma B.3 With $\sigma_k \in \mathbb{R}^2$ defined as in (5.7), we have that

$$\sum_{k=0}^5 \mathbb{J}(\sigma_k - \sigma_{k+1}) \left(\frac{\sigma_k + \sigma_{k+1}}{3} \right)^T = \sqrt{3} \mathbb{1}.$$

Proof With the abbreviations $\phi_x = \frac{\pi}{3}x$ and $\psi = \frac{\pi}{3}$:

$$\begin{aligned} \sum_{k=0}^5 \mathbb{J}(\sigma_k - \sigma_{k+1}) \left(\frac{\sigma_k + \sigma_{k+1}}{3} \right)^T &= \frac{1}{3} \sum_{k=0}^5 \begin{pmatrix} \sin \phi_{k+1} - \sin \phi_k \\ \cos \phi_k - \cos \phi_{k+1} \end{pmatrix} \begin{pmatrix} \cos \phi_k + \cos \phi_{k+1} \\ \sin \phi_k + \sin \phi_{k+1} \end{pmatrix}^T \\ &= \frac{1}{3} \sum_{k=0}^5 \begin{pmatrix} 2 \sin \frac{\psi}{2} \\ \sin \phi_{k+\frac{1}{2}} \end{pmatrix} \begin{pmatrix} \cos \phi_{k+\frac{1}{2}} \\ 2 \cos \frac{\psi}{2} \end{pmatrix} \begin{pmatrix} \cos \phi_{k+\frac{1}{2}} \\ \sin \phi_{k+\frac{1}{2}} \end{pmatrix}^T \\ &= \frac{\sin \psi}{3} \sum_{k=0}^5 \begin{pmatrix} 2 \cos^2 \phi_{k+\frac{1}{2}} & 2 \cos \phi_{k+\frac{1}{2}} \sin \phi_{k+\frac{1}{2}} \\ 2 \cos \phi_{k+\frac{1}{2}} \sin \phi_{k+\frac{1}{2}} & 2 \sin^2 \phi_{k+\frac{1}{2}} \end{pmatrix} \\ &= \frac{\sqrt{3}}{6} \sum_{k=0}^5 \left[\mathbb{1} + \begin{pmatrix} \cos \phi_{2k+1} & \sin \phi_{2k+1} \\ \sin \phi_{2k+1} & -\cos \phi_{2k+1} \end{pmatrix} \right] = \sqrt{3} \mathbb{1}. \end{aligned}$$

□

Lemma B.4 Let the scheme $B := (b_{pqr})_{p,q,r \in \{1,2\}} \in \mathbb{R}^{2 \times 2 \times 2}$ of eight numbers $b_{pqr} \in \mathbb{R}$ be symmetric in the last two indices, $b_{pqr} = b_{prq}$. With $\sigma_k \in \mathbb{R}^2$ defined as in (5.7), we have that

$$\sum_{k=0}^5 \text{tr} [(\sigma_k | \sigma_{k+1})^{-1} (B : [\sigma_k]^2 | B : [\sigma_{k+1}]^2)] \mathbb{J}(\sigma_k - \sigma_{k+1}) = 2\sqrt{3} \text{tr}_{12}[B]^T. \tag{B.3}$$

Proof In principle, this lemma can be verified by a direct calculation, by writing out the six terms in the sum explicitly and using trigonometric identities. Below, we give a slightly more conceptual proof, in which we use symmetry arguments to reduce the number of expressions significantly.

For the matrix involving B , we obtain

$$\begin{aligned} &(B : [\sigma_k]^2 | B : [\sigma_{k+1}]^2) \\ &= \left(b_{111}\sigma_{k,1}^2 + b_{122}\sigma_{k,2}^2 + 2b_{112}\sigma_{k,1}\sigma_{k,2} \quad b_{111}\sigma_{k+1,1}^2 + b_{122}\sigma_{k+1,2}^2 + 2b_{112}\sigma_{k+1,1}\sigma_{k+1,2} \right), \\ &\quad \left(b_{211}\sigma_{k,1}^2 + b_{222}\sigma_{k,2}^2 + 2b_{212}\sigma_{k,1}\sigma_{k,2} \quad b_{211}\sigma_{k+1,1}^2 + b_{222}\sigma_{k+1,2}^2 + 2b_{212}\sigma_{k+1,1}\sigma_{k+1,2} \right), \end{aligned}$$

while clearly

$$(\sigma_k | \sigma_{k+1})^{-1} = \frac{2}{\sqrt{3}} \begin{pmatrix} \sigma_{k+1,2} & -\sigma_{k+1,1} \\ -\sigma_{k,2} & \sigma_{k,1} \end{pmatrix}.$$

The sum of the diagonal entries of the matrix product are easily calculated,

$$T_k := \text{tr}[(\sigma_k | \sigma_{k+1})^{-1} (B : [\sigma_k]^2 | B : [\sigma_{k+1}]^2)] = \frac{2}{\sqrt{3}} \sum_{p,q,r=1}^2 b_{pqr} \gamma_{pqr,k},$$

with the trigonometric expressions

$$\begin{aligned} \gamma_{111,k} &= \sigma_{k,1}^2 \sigma_{k+1,2} - \sigma_{k+1,1}^2 \sigma_{k,2}, & \gamma_{122,k} &= \sigma_{k,2}^2 \sigma_{k+1,1} - \sigma_{k+1,2}^2 \sigma_{k,1}, \\ \gamma_{112,k} &= \gamma_{121,k} = \sigma_{k,1} \sigma_{k,2} \sigma_{k+1,2} - \sigma_{k+1,1} \sigma_{k+1,2} \sigma_{k,2}, \\ \gamma_{211,k} &= \sigma_{k+1,1}^2 \sigma_{k,1} - \sigma_{k,1}^2 \sigma_{k+1,1}, & \gamma_{222,k} &= \sigma_{k+1,2}^2 \sigma_{k,1} - \sigma_{k,2}^2 \sigma_{k+1,1}, \\ \gamma_{212,k} &= \gamma_{221,k} = \sigma_{k+1,1} \sigma_{k+1,2} \sigma_{k,1} - \sigma_{k,1} \sigma_{k,2} \sigma_{k+1,1}. \end{aligned}$$

To key step is to calculate the sum over $k = 0, 1, \dots, 5$ of the products of T_k with the respective vector

$$\eta_k = \mathbb{J}(\sigma_k - \sigma_{k+1}) = \begin{pmatrix} \sigma_{k+1,2} - \sigma_{k,2} \\ \sigma_{k,1} - \sigma_{k+1,1} \end{pmatrix}.$$

Several simplifications of this sum can be performed, thanks to the particular form of the $\gamma_{pqr,k}$ and elementary trigonometric identities. First, observe that $\sigma_{k+3} = -\sigma_k$, and hence that $\gamma_{pqr,k+3} = -\gamma_{pqr,k}$. Since further $\eta_{k+3} = -\eta_k$, it follows that

$$\gamma_{pqr,k+3} \eta_{k+3} = \gamma_{pqr,k} \eta_k. \tag{B.4}$$

Second, η can be evaluated explicitly for $k = 1, 2, 3$:

$$\eta_0 = \frac{1}{2} \begin{pmatrix} \sqrt{3} \\ 1 \end{pmatrix}, \quad \eta_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \eta_2 = \frac{1}{2} \begin{pmatrix} -\sqrt{3} \\ 1 \end{pmatrix}. \tag{B.5}$$

Third, since $\sigma_{0,1} = -\sigma_{3,1}$ and $\sigma_{1,1} = -\sigma_{2,1}$, as well as $\sigma_{0,2} = \sigma_{3,2}$ and $\sigma_{1,2} = \sigma_{2,2}$, we obtain that

$$\gamma_{pqr,1} = 0 \text{ if } p + q + r \text{ is odd, and } \gamma_{pqr,2} = (-1)^{p+q+r} \gamma_{pqr,0}. \tag{B.6}$$

By putting this together, we arrive at

$$\begin{aligned} \sum_{k=0}^5 \gamma_{pqr,k} \eta_k &\stackrel{\text{(B.4)}}{=} 2 \sum_{k=0}^2 \gamma_{pqr,k} \eta_k \\ &\stackrel{\text{(B.5)}}{=} \begin{pmatrix} \sqrt{3}(\gamma_{pqr,0} - \gamma_{pqr,2}) \\ \gamma_{pqr,0} + 2\gamma_{pqr,1} + \gamma_{pqr,2} \end{pmatrix} \\ &\stackrel{\text{(B.6)}}{=} \begin{pmatrix} \sqrt{3}(1 - (-1)^{p+q+r})\gamma_{pqr,0} \\ (1 + (-1)^{p+q+r})(\gamma_{pqr,0} + \gamma_{pqr,1}) \end{pmatrix} = \begin{pmatrix} 2\sqrt{3} \gamma_{pqr,0} (1 - \epsilon_{pqr}) \\ 2(\gamma_{pqr,0} + \gamma_{pqr,1}) \epsilon_{pqr} \end{pmatrix}, \end{aligned}$$

where $\epsilon_{pqr} = 1$ if $p + q + r$ is even, and $\epsilon_{pqr} = 0$ if $p + q + r$ is odd. By elementary computations,

$$\begin{aligned} p + q + r \text{ odd, } k = 0 : & \gamma_{111,0} = \frac{\sqrt{3}}{2}, \gamma_{122,0} = 0, \gamma_{212,0} = \gamma_{221,0} = \frac{\sqrt{3}}{4}; \\ p + q + r \text{ even, } k = 0 : & \gamma_{211,0} = -\frac{1}{4}, \gamma_{222,0} = \frac{3}{4}, \gamma_{112,0} = \gamma_{121,0} = 0; \\ p + q + r \text{ even, } k = 1 : & \gamma_{211,1} = \frac{1}{4}, \gamma_{222,1} = \frac{3}{4}, \gamma_{112,1} = \gamma_{121,1} = \frac{3}{4}, \end{aligned}$$

and so the final result is:

$$\begin{aligned} & \sum_{k=0}^5 \operatorname{tr} [(\sigma_k | \sigma_{k+1})^{-1} (B : [\sigma_k]^2 | B : [\sigma_{k+1}]^2)] \mathbb{J}(\sigma_k - \sigma_{k+1}) \\ &= \sum_{k=0}^5 T_k \eta_k = \frac{2}{\sqrt{3}} \sum_{p,q,r=1}^2 \left(b_{pqr} \sum_{k=0}^5 \gamma_{pqr,k} \eta_k \right) = 2\sqrt{3} \begin{pmatrix} b_{111} + b_{212} \\ b_{222} + b_{112} \end{pmatrix}, \end{aligned}$$

which is (B.3). □

Lemma B.5 *With $\sigma'_k \in \mathbb{R}^2$ defined as in (5.14), and with $B = (b_{pqr})_{p,q,r \in \{1,2\}} \in \mathbb{R}^{2 \times 2 \times 2}$ such that $b_{pqr} = 0$ except for $b_{122} = b_{211} = 1$, we have that*

$$\sum_{k=0}^5 \operatorname{tr} [(\sigma'_k | \sigma'_{k+1})^{-1} (B : [\sigma'_k]^2 | B : [\sigma'_{k+1}]^2)] \mathbb{J}(\sigma_k - \sigma_{k+1}) = - \begin{pmatrix} 1 \\ 1 \end{pmatrix}. \tag{B.7}$$

Proof This is a slightly tedious, but straightforward calculation. First, by the choice of B ,

$$\beta_k := (B : [\sigma'_k]^2 | B : [\sigma'_{k+1}]^2) = \begin{pmatrix} (\sigma'_{k,2})^2 & (\sigma'_{k+1,2})^2 \\ (\sigma'_{k,1})^2 & (\sigma'_{k+1,1})^2 \end{pmatrix},$$

and so, by definition of the σ'_k in (5.14),

$$\beta_0 = \beta_3 = \begin{pmatrix} 0 & \frac{1}{4} \\ 1 & \frac{1}{4} \end{pmatrix}, \quad \beta_1 = \beta_2 = \begin{pmatrix} \frac{1}{4} & 1 \\ \frac{1}{4} & 0 \end{pmatrix}, \quad \beta_4 = \beta_5 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

For the inverse matrices $S_k := (\sigma'_k | \sigma'_{k+1})^{-1}$, we obtain

$$S_0 = \begin{pmatrix} 1 & -1 \\ 0 & 2 \end{pmatrix} = -S_3, \quad S_1 = \begin{pmatrix} 2 & 0 \\ -1 & 1 \end{pmatrix} = -S_4, \quad S_2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = -S_5.$$

For the traces $T_k := \operatorname{tr} [S_k \beta_k]$, we thus obtain the values:

$$T_0 = T_1 = -\frac{1}{2}, \quad T_3 = T_4 = \frac{1}{2}, \quad T_2 = T_5 = 0.$$

In conclusion,

$$\sum_{k=0}^5 T_k \mathbb{J}(\sigma_k - \sigma_{k+1}) = \mathbb{J} \left[-\frac{1}{2}(\sigma_0 - \sigma_2) + \frac{1}{2}(\sigma_3 - \sigma_5) \right] = \mathbb{J} \begin{pmatrix} -1 \\ 1 \end{pmatrix} = - \begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

which is (B.7). □

Appendix C: Lack of convexity

Below, we discuss why the minimization problem (3.7) is not convex. More precisely, we show that $G \mapsto \mathbf{E}_{\boxplus}(G; \hat{G})$ is not convex as a function of G on the affine ansatz space $\mathcal{A}_{\mathcal{J}}$. Since $\mathbf{E}_{\boxplus}(G; \hat{G})$ is a convex combination of the expressions $\mathbb{H}_m((A_m | b_m); (\hat{A}_m | \hat{b}_m))$, it clearly suffices to discuss the convexity of the latter.

We consider a curve $s \mapsto (A_m + s\alpha_m | b_m + s\beta_m)$ and evaluate the second derivatives of the components of the functional at $s = 0$. First,

$$\begin{aligned}
 \text{I} &:= \frac{d^2}{ds^2} \Big|_{s=0} \left(\frac{1}{2\tau} \int_{\Delta_m} |(A_m - \hat{A}_m + s\alpha_m)\omega + (b_m - \hat{b}_m) + s\beta_m|^2 d\omega \right) \\
 &= \frac{1}{\tau} \int_{\Delta_m} |\alpha_m\omega + \beta_m|^2 d\omega.
 \end{aligned}$$

Second,

$$\begin{aligned}
 \text{II} &:= \frac{d^2}{ds^2} \Big|_{s=0} \int_{\Delta_m} V((A_m + s\alpha_m)\omega + (b_m + s\beta_m)) d\omega \\
 &= \int_{\Delta_m} (\alpha_m\omega + \beta_m)^T \cdot \nabla^2 V(A_m\omega + b_m) \cdot (\alpha_m\omega + \beta_m) d\omega.
 \end{aligned}$$

If we assume that $\nabla^2 V \geq \lambda \mathbb{1}$, then we obtain for the sum of these two contributions that

$$\text{I} + \text{II} \geq \left(\frac{1}{\tau} + \lambda \right) \int_{\Delta_m} |\alpha_m\omega + \beta_m|^2 d\omega.$$

For the remaining term, however, we obtain — using the abbreviations $\tilde{g}(s) = s\tilde{h}'(s)$ and $\tilde{f}(s) = s\tilde{g}'(s)$ — that

$$\begin{aligned}
 &\frac{d^2}{ds^2} \Big|_{s=0} \tilde{h} \left(\frac{\det(A_m + s\alpha_m)}{\bar{\rho}_m} \right) \\
 &= \frac{d}{ds} \Big|_{s=0} \left\{ \tilde{g} \left(\frac{\det(A_m + s\alpha_m)}{\bar{\rho}_m} \right) \text{tr} [(A_m + s\alpha_m)^{-1}\alpha_m] \right\} \\
 &= \tilde{f} \left(\frac{\det A_m}{\bar{\rho}_m} \right) (\text{tr} [A_m^{-1}\alpha_m])^2 - \tilde{g} \left(\frac{\det A_m}{\bar{\rho}_m} \right) \text{tr} [(A_m^{-1}\alpha_m)^2].
 \end{aligned}$$

Now observe that $\tilde{f}(s) = P'(1/s) - sP(1/s)$ is a non-negative, and $\tilde{g}(s) = -sP(1/s)$ is a non-positive function. Thus, from the two terms in the final sum, the first one is generally non-negative whereas the second one is of indefinite sign. Choosing

$$\alpha_m := A_m \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \text{such that} \quad (\text{tr} [A_m^{-1}\alpha_m])^2 = 0, \text{tr} [(A_m^{-1}\alpha_m)^2] = 2,$$

the sum is obviously negative.

References

1. Ambrosio, L., Gigli, N., Savaré, G.: Gradient Flows in Metric Spaces and in the Space of Probability Measures. Lectures in Mathematics ETH Zürich, 2nd edn. Birkhäuser Verlag, Basel (2008)
2. Ambrosio, L., Lisini, S., Savaré, G.: Stability of flows associated to gradient vector fields and convergence of iterated transport maps. *Manuscr. Math.* **121**(1), 1–50 (2006)
3. Benamou, J.-D., Carlier, G., Mérigot, Q., Oudet, E.: Discretization of functionals involving the Monge–Ampère operator. *Numer. Math.* **134**(3), 611–636 (2015)
4. Bessemoulin-Chatard, M., Filbet, F.: A finite volume scheme for nonlinear degenerate parabolic equations. *SIAM J. Sci. Comput.* **34**(5), B559–B583 (2012)
5. Blanchet, A., Calvez, V., Carrillo, J.A.: Convergence of the mass-transport steepest descent scheme for the sub-critical Patlak-Keller-Segel model. *SIAM J. Numer. Anal.* **46**(2), 691–721 (2008)
6. Budd, C.J., Collins, G.J., Huang, W.Z., Russell, R.D.: 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**(1754), 1047–1077 (1999)
7. Calvez, V., Gallouët, T.O.: Particle approximation of the one dimensional Keller-Segel equation, stability and rigidity of the blow-up. *Discrete Contin. Dyn. Syst. Ser. A* **36**(3), 1175–1208 (2015)

8. Cancès, C., Guichard, C.: Numerical analysis of a robust free energy diminishing finite volume scheme for parabolic equations with gradient structure. *Found. Comput. Math.* (2016). <https://doi.org/10.1007/s10208-016-9328-6>
9. Carrillo, J., Craig, K., Patacchini, F.: A blob method for diffusion. arXiv preprint [arXiv:1709.09195](https://arxiv.org/abs/1709.09195) (2017)
10. Carrillo, J.A., Chertock, A., Huang, Y.: A finite-volume method for nonlinear nonlocal equations with a gradient flow structure. *Commun. Comput. Phys.* **17**(01), 233–258 (2015)
11. Carrillo, J.A., Huang, Y., Patacchini, F.S., Wolansky, G.: Numerical study of a particle method for gradient flows. *Kinet. Relat. Models* **10**(3), 613–641 (2017)
12. Carrillo, J.A., Lisini, S.: On the asymptotic behavior of the gradient flow of a polyconvex functional. In: Holden, H., Karlsen, K.H. (eds.) *Nonlinear Partial Differential Equations and Hyperbolic Wave Phenomena*, Volume 526 of *Contemporary Mathematics*, pp. 37–51. American Mathematical Society, Providence (2010)
13. Carrillo, J.A., Moll, J.S.: Numerical simulation of diffusive and aggregation phenomena in nonlinear continuity equations by evolving diffeomorphisms. *SIAM J. Sci. Comput.*, **31**(6), 4305–4329 (2009/10)
14. Carrillo, J.A., Patacchini, F.S., Sternberg, P., Wolansky, G.: Convergence of a particle method for diffusive gradient flows in one dimension. *SIAM J. Math. Anal.* **48**(6), 3708–3741 (2016)
15. Carrillo, J.A., Ranetbauer, H., Wolfram, M.-T.: Numerical simulation of nonlinear continuity equations by evolving diffeomorphisms. *J. Comput. Phys.* **327**, 186–202 (2016)
16. Cavalli, F., Naldi, G., Puppo, G., Semplice, M.: High-order relaxation schemes for nonlinear degenerate diffusion problems. *SIAM J. Numer. Anal.* **45**(5), 2098–2119 (2007)
17. Daneri, S., Savaré, G.: Eulerian calculus for the displacement convexity in the Wasserstein distance. *SIAM J. Math. Anal.* **40**(3), 1104–1122 (2008)
18. Degond, P., Mustieles, F.-J.: A deterministic approximation of diffusion equations using particles. *SIAM J. Sci. Statist. Comput.* **11**(2), 293–310 (1990)
19. Düring, B., Matthes, D., Milišić, J.P.: A gradient flow scheme for nonlinear fourth order equations. *Discrete Contin. Dyn. Syst. Ser. B* **14**(3), 935–959 (2010)
20. Evans, L., Savin, O., Gangbo, W.: Diffeomorphisms and nonlinear heat flows. *SIAM J. Math. Anal.* **37**(3), 737–751 (2005)
21. Gianazza, U., Savaré, G., Toscani, G.: The wasserstein gradient flow of the fisher information and the quantum drift-diffusion equation. *Arch. Ration. Mech. Anal.* **194**(1), 133–220 (2009)
22. Gosse, L., Toscani, G.: Identification of asymptotic decay to self-similarity for one-dimensional filtration equations. *SIAM J. Numer. Anal.* **43**(6), 2590–2606 (2006). (electronic)
23. Gosse, L., Toscani, G.: Lagrangian numerical approximations to one-dimensional convolution-diffusion equations. *SIAM J. Sci. Comput.* **28**(4), 1203–1227 (2006). (electronic)
24. Jordan, R., Kinderlehrer, D., Otto, F.: The variational formulation of the Fokker–Planck equation. *SIAM J. Math. Anal.* **29**(1), 1–17 (1998)
25. Junge, O., Matthes, D., Osberger, H.: A fully discrete variational scheme for solving nonlinear fokker-planck equations in higher space dimensions. arXiv preprint [arXiv:1509.07721](https://arxiv.org/abs/1509.07721) (2015)
26. Kurganov, A., Tadmor, E.: New high-resolution central schemes for nonlinear conservation laws and convection-diffusion equations. *J. Comput. Phys.* **160**(1), 241–282 (2000)
27. Lions, P.-L., Mas-Gallic, S.: Une méthode particulière déterministe pour des équations diffusives non linéaires. *C. R. Acad. Sci. Paris Sér. I Math.* **332**(4), 369–376 (2001)
28. Mas-Gallic, S.: The diffusion velocity method: a deterministic way of moving the nodes for solving diffusion equations. *Transp. Theory Stat. Phys.* **31**(4–6), 595–605 (2002)
29. Matthes, D., Osberger, H.: Convergence of a variational Lagrangian scheme for a nonlinear drift diffusion equation. *ESAIM Math. Model. Numer. Anal.* **48**(3), 697–726 (2014)
30. McCann, R.J.: A convexity principle for interacting gases. *Adv. Math.* **128**(1), 153–179 (1997)
31. Osberger, H., Matthes, D.: Convergence of a variational Lagrangian scheme for a nonlinear drift diffusion equation. *ESAIM Math. Model. Numer. Anal.* **48**(3), 697–726 (2014)
32. Osberger, H., Matthes, D.: Convergence of a fully discrete variational scheme for a thin-film equation. In: Bergounioux, M., et al. (eds.) *Topological Optimization and Optimal Transport in the Applied Sciences. Radon Series on Computational and Applied Mathematics 17*, pp. 356–399. De Gruyter (2017)
33. Osberger, H., Matthes, D.: A convergent Lagrangian discretization for a nonlinear fourth order equation. *Found. Comput. Math.* **17**, 1–54 (2015)
34. Otto, F.: Lubrication approximation with prescribed nonzero contact angle. *Commun. Partial Differ. Equ.* **23**(11–12), 2077–2164 (1998)
35. Otto, F.: The geometry of dissipative evolution equations: the porous medium equation. *Commun. Partial Differ. Equ.* **26**(1–2), 101–174 (2001)
36. Peyré, G.: Entropic approximation of Wasserstein gradient flows. *SIAM J. Imaging Sci.* **8**(4), 2323–2351 (2015)

37. Russo, G.: Deterministic diffusion of particles. *Commun. Pure Appl. Math.* **43**(6), 697–733 (1990)
38. Vázquez, J .L.: *The Porous Medium Equation: Mathematical Theory*. Oxford University Press, Oxford (2007)
39. Villani, C.: *Topics in Optimal Transportation*. Graduate Studies in Mathematics, vol. 58. American Mathematical Society, Providence (2003)
40. Westdickenberg, M., Wilkening, J.: Variational particle schemes for the porous medium equation and for the system of isentropic euler equations. *ESAIM. Math. Model. Numer. Anal.* **44**(1), 133–166 (2010)