ESAIM: M2AN 48 (2014) 697–726
DOI: 10.1051/m2an/2013126

ESAIM: Mathematical Modelling and Numerical Analysis
www.esaim-m2an.org

CONVERGENCE OF A VARIATIONAL LAGRANGIAN SCHEME FOR A NONLINEAR DRIFT DIFFUSION EQUATION*

Daniel Matthes[1] and Horst Osberger[1]

Abstract. We study a Lagrangian numerical scheme for solution of a nonlinear drift diffusion equation on an interval. The discretization is based on the equation's gradient flow structure with respect to the Wasserstein distance. The scheme inherits various properties from the continuous flow, like entropy monotonicity, mass preservation, metric contraction and minimum/ maximum principles. As the main result, we give a proof of convergence in the limit of vanishing mesh size under a CFL-type condition. We also present results from numerical experiments.

Mathematics Subject Classification. 65N12, 49Q20, 65N06.

Received January 29, 2013. Revised August 2, 2013
Published online February 7, 2014.

1. Introduction

In this paper, we propose and analyze a particular spatio-temporal discretization of the following nonlinear initial-boundary value problem on an interval $I = [a, b]$:

$$\partial_t u = \mathrm{P}(u)_{xx} + (V_x(x)u)_x, \quad u_x(t;a) = u_x(t;b) = 0, \quad u(0;x) = u^0(x) \geq 0. \tag{1.1}$$

We are interested in approximating non-negative weak solutions $u : [0,T] \times I \to \mathbb{R}_{\geq 0}$ to (1.1) on arbitrary time horizons $T > 0$. Our assumptions are that

- the nonlinearity $\mathrm{P} : \mathbb{R}_{\geq 0} \to \mathbb{R}$ is C^2-smooth on $\mathbb{R}_{>0}$, and satisfies

$$\mathrm{P}(0) = 0, \quad \mathrm{P}'(r) > 0, \quad \lim_{r \downarrow 0} \mathrm{P}'(r) < \infty, \quad \lim_{r \to \infty} \mathrm{P}'(r) = +\infty, \quad s \mapsto \mathrm{P}(1/s) \quad \text{is concave,} \tag{1.2}$$

  the prototypical example being the porous medium term $\mathrm{P}(r) = r^m$ with some $m > 1$;
- the potential $V : I \to \mathbb{R}$ is C^2-smooth with

$$V_x(a) = V_x(b) = 0. \tag{1.3}$$

Keywords and phrases. Lagrangian discretization, nonlinear Fokker-Planck equation, gradient flow, Wasserstein metric.

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

[1] Zentrum Mathematik, TU München, Boltzmannstr. 3, 85748 Garching, Germany. matthes@ma.tum.de; osberger@ma.tum.de

Article published by EDP Sciences

© EDP Sciences, SMAI 2014

https://doi.org/10.1051/m2an/2013126 Published online by Cambridge University Press

Under the given regularity assumptions, there are numerous possibilities to design efficient numerical schemes for solution of (1.1), *e.g.*, using finite differences. The particular discretization under consideration here is special insofar as it is based on the representation of (1.1) as a gradient flow, and it inherits certain qualitative features of that variational structure.

1.1. Gradient flow structure

We summarize some basic facts about the variational formulation of (1.1). The divergence form in combination with the no-flux boundary conditions and (1.3) implies the conservation of mass

$$\int_I u(t;x)\,\mathrm{d}x = M := \int_I u^0(x)\,\mathrm{d}x > 0 \quad \text{for all } t > 0, \tag{1.4}$$

and we shall consider M as some fixed quantity from now on. The space $\mathrm{D}^M(I) \subset L^1(I)$ of bounded and strictly positive densities u with total mass M can be endowed with the L^2-Wasserstein metric $\mathbf{W}_2$; the definition and elementary properties of this metric are reviewed in Section 2 below. Next, introduce the *energy functional* $\mathbf{E}$ for $u \in \mathrm{D}^M(I)$ by

$$\mathbf{E}(u) = \int_I \phi(u(x))\,\mathrm{d}x + \int_I u(x)V(x)\,\mathrm{d}x, \tag{1.5}$$

where the internal energy potential $\phi : \mathbb{R}_{\geq 0} \to \mathbb{R}$ is an arbitrary second anti-derivative of $r \mapsto P'(r)/r$; see (2.3). The link between the energy $\mathbf{E}$ and equation (1.1) – which has been rigorously established by Otto [18] – is that solutions to (1.1) form a gradient flow in the energy landscape of $\mathbf{E}$ with respect to the metric $\mathbf{W}_2$. Further, it has been observed by McCann [17] that the functional $\mathbf{E}$ is $(-\Lambda)$-convex along geodesics in $\mathbf{W}_2$, with

$$\Lambda = \max_{x\in I}\,(-V_{xx}(x)) \geq 0. \tag{1.6}$$

Consequently, the gradient flow is $(-\Lambda)$-contractive[2]. Some implications are:

(1) The energy $\mathbf{E}(u(t))$ is monotonically decreasing in t.
(2) Two solutions u, v diverge at most at an exponential rate of Λ w.r.t. $\mathbf{W}_2$, *i.e.*,

$$\mathbf{W}_2(u(t), v(t)) \leq \mathbf{W}_2(u^0, v^0)e^{\Lambda t} \quad \text{for all } t > 0. \tag{1.7}$$

(3) There is a unique non-negative and mass preserving global solution for measure valued initial conditions, *i.e.*, the density u^0 in (1.1) can be replaced by an arbitrary non-negative measure on I with mass M.

Below, we discuss in which sense these properties are inherited by our discretization.

1.2. Discretization

Semi-discretization in time of gradient flow equations like (1.1) has become a key tool in existence proofs and for the rigorous derivation of a priori estimates. The celebrated *minimizing movement* scheme [1] (also referred to as *JKO* [13] or simply *implicit Euler* scheme) works in the situation at hand as follows: given a time step $\tau > 0$, one defines inductively – starting from $u_\tau^0 = u^0$ – approximations u_τ^n of $u(n\tau)$ as minimizers in $\mathrm{D}^M(I)$ of "penalized energy functionals" $\mathbf{E}_\tau(\cdot, u_\tau^{n-1})$, given by

$$\mathbf{E}_\tau(u, u_\tau^{n-1}) = \frac{1}{2\tau}\mathbf{W}_2\left(u, u_\tau^{n-1}\right)^2 + \mathbf{E}(u). \tag{1.8}$$

Thanks to the $(-\Lambda)$-convexity of $\mathbf{E}$, it follows from the theory developed in [1] that the functions $\bar{u}_\tau : [0,\infty) \to \mathrm{D}^M(I)$ obtained by piecewise constant interpolation in time converge for $\tau \downarrow 0$ to the unique weak solution $u : [0,\infty) \to \mathrm{D}^M(I)$ of (1.1).

[2]Note that $(-\Lambda)$-contractivity with $\Lambda > 0$ is a weaker property than contractivity of the flow. Indeed, trajectories may diverge, but not faster than in (1.7).

https://doi.org/10.1051/m2an/2013126 Published online by Cambridge University Press

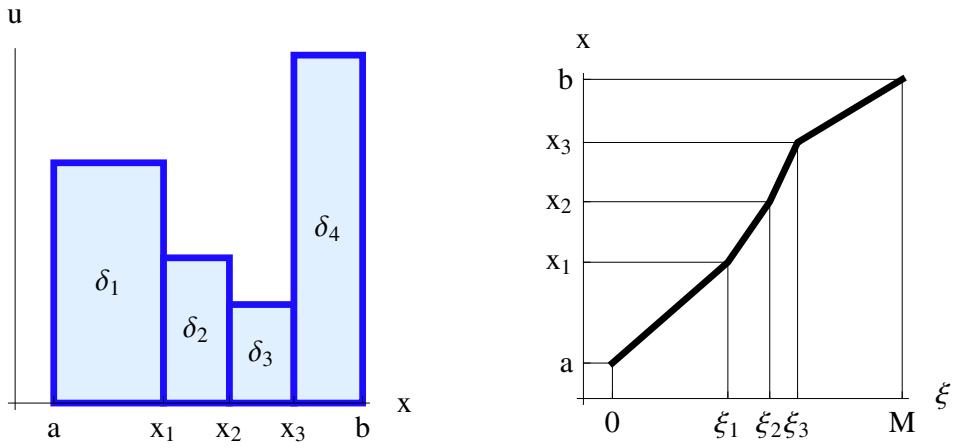


FIGURE 1. A typical density function $u \in \mathrm{D}^M_{\boldsymbol{\xi}}(I)$ (left) and inverse distribution function $\mathrm{X} \in \mathfrak{X}_{\boldsymbol{\xi}}$.

To obtain a *full* (spatio-temporal) discretization, we perform the minimization of $\mathbf{E}_\tau$ not over the entire set $\mathrm{D}^M(I)$, but over a submanifold $\mathrm{D}^M_{\boldsymbol{\xi}}(I)$ of finite dimension $(K-1) \in \mathbb{N}$. The discretization parameter $\boldsymbol{\xi} = (\xi_0, \xi_1, \ldots, \xi_K)$ is an increasing sequence of numbers $\xi_k \in [0, M]$, with $0 = \xi_0 < \xi_1 < \xi_2 < \cdots < \xi_K = M$. The corresponding submanifold $\mathrm{D}^M_{\boldsymbol{\xi}}(I)$ consists of piecewise constant density functions $u \in \mathrm{D}^M(I)$ of the form

$$u = \sum_{k=1}^{K} u_k \mathbf{1}_{(x_{k-1}, x_k)}, \tag{1.9}$$

where the end points x_k of the intervals are variable subject to the constraint

$$a = x_0 < x_1 < \cdots < x_K = b,$$

and the positive weights u_k are given in terms of the x_k by

$$(x_k - x_{k-1})u_k = \delta_k := \xi_k - \xi_{k-1} > 0. \tag{1.10}$$

One may think of δ_1 to δ_K as lumps of mass, each of which is uniformly distributed on its respective interval $(x_{k-1}, x_k]$, see Figure 1. Our *Lagrangian* scheme is now defined as follows.

Given a discretization $\Delta = (\tau; \boldsymbol{\xi})$ consisting of a time step $\tau > 0$ and a spatial mesh $\boldsymbol{\xi}$, and an initial condition $u^0_\Delta \in \mathrm{D}^M_{\boldsymbol{\xi}}(I)$, define a *discrete solution* $u_\Delta = (u^0_\Delta, u^1_\Delta, \ldots)$ inductively by

$$u^n_\Delta = \operatorname*{argmin}_{u \in \mathrm{D}^M_{\boldsymbol{\xi}}(I)} \mathbf{E}_\tau \left(u, u^{n-1}_\Delta\right) \quad \text{for } n = 1, 2, \ldots \tag{1.11}$$

The seemingly involved definition of the recursion (1.11) leads to a simple and practical numerical scheme, whose complexity is comparable to that of a standard discretization of (1.1) by finite differences. In addition, the Lagrangian nature of the scheme admits a geometric interpretation of the solution: the time-dependent end points $x_k(t)$ of the intervals in (1.9) for the transport of mass elements on I. Concerning structure preservation, we summarize some noteworthy features of this approach:

- The energy $\mathbf{E}(u^n_\Delta)$ is monotone in n, and all u^n_Δ are positive and have the same mass M.

https://doi.org/10.1051/m2an/2013126 Published online by Cambridge University Press

- Any two discrete solutions u_Δ and v_Δ on the same grid Δ satisfy the expansion estimate

$$\mathbf{W}_2\left(u_\Delta^n, v_\Delta^n\right) \le (1-2\Lambda\tau)^{-n/2}\mathbf{W}_2\left(u_\Delta^0, v_\Delta^0\right),$$

  which turns into (1.7) in the limit $\tau \downarrow 0$. See Section 5.1.
- The scheme is applicable to arbitrary initial data $u^0 \in L^1(I)$ with finite energy. In fact, weak convergence of u_Δ^0 to u^0 suffices to conclude strong convergence of the discrete solution u_Δ in $L^1([0,T]\times I)$ to the correct weak solution u of (1.1). See Theorem 1.1 below.
- Discrete solutions obey a minimum/maximum principle. See Section 5.2.

The choice of $\mathrm{D}_{\boldsymbol{\xi}}^M(I)$ originates from an alternative formulation of equation (1.1). Namely, u is a (positive and classical) solution to (1.1) iff its inverse distribution function X – see Section 2 for its definition – satisfies the initial-boundary value problem

$$\partial_t \mathrm{X} = \psi'(\mathrm{X}_\xi)_\xi - V_x \circ \mathrm{X}, \quad \mathrm{X}(t;0) = a,\ \mathrm{X}(t;M) = b, \quad \mathrm{X}(0;\xi) = \mathrm{X}^0(\xi), \tag{1.12}$$

where $\psi : \mathbb{R}_+ \to \mathbb{R}$ is defined by

$$\psi(s) = s\phi\left(s^{-1}\right) \quad \text{for all } s > 0. \tag{1.13}$$

The variational structure of (1.12) is quite apparent: solutions X to (1.12) are gradient flows of the functional

$$\mathbb{E}(\mathrm{X}) = \int_0^M \psi\left(\mathrm{X}_\xi(\xi)\right)\,\mathrm{d}\xi + \int_0^M V\circ\mathrm{X}(\xi)\,\mathrm{d}\xi \tag{1.14}$$

with respect to the usual scalar product on $L^2([0,M])$. In effect, we discretize the L^2-gradient flow (1.12) rather than the $\mathbf{W}_2$-gradient flow (1.1), representing X as a linear combination of piecewise linear ansatz functions with respect to the (time-independent) mesh $\boldsymbol{\xi}$.

1.3. Convergence result

Our main result is the following.

Theorem 1.1. *Let a non-negative initial condition $u^0 \in L^1(I)$ of mass M with $\mathbf{E}(u^0) < \infty$ be given, and fix a time horizon $T > 0$.*

Consider a sequence of discretizations $\Delta^{(j)} = (\tau^{(j)}; \boldsymbol{\xi}^{(j)})$, consisting of time steps $\tau^{(j)} \downarrow 0$ and spatial meshes $\boldsymbol{\xi}^{(j)}$, and an associated sequence of initial conditions $u^0_{\Delta^{(j)}} \in \mathrm{D}^M_{\boldsymbol{\xi}^{(j)}}(I)$. Assume that $u^0_{\Delta^{(j)}} \to u^0$ weakly in $L^1(I)$, assume that $\delta_k^{(j)}$ defined from $\boldsymbol{\xi}^{(j)}$ according to (1.10) satisfies $\max_k(\delta_k^{(j)}) \downarrow 0$, and further assume that $\max_k \delta_k^{(j)} / \min_\ell \delta_\ell^{(j)} \le \bar\alpha$ and that $\mathbf{E}(u^0_{\Delta^{(j)}}) \le \overline{\mathbf{E}}$, for appropriate j-independent constants $\bar\alpha, \overline{\mathbf{E}} > 0$. Finally, assume that the following inverse CFL condition *holds:*

$$\left(\max_k \delta_k^{(j)}\right)^2 < 6\psi''\left(\frac{6\bar\alpha e^{2\Lambda T}}{\min_x u^0_{\Delta^{(j)}}}\right)\tau^{(j)}, \tag{1.15}$$

with ψ defined in (1.13), and with $\Lambda \ge 0$ from (1.6).

The scheme (1.11) produces a sequence of discrete solutions $u_{\Delta^{(j)}}$. Denote by $\bar u_{\Delta^{(j)}} : [0,\infty) \to \mathrm{D}^M_{\boldsymbol{\xi}_j}(I)$ the respective interpolants that are piecewise constant in time, see (6.2). Then $\bar u_{\Delta^{(j)}}$ converges strongly in $L^1([0,T]\times I)$ to the unique weak solution u of (1.1).

A comment is due on condition (1.15). Since $\psi''(s) \to 0$ for $s \to \infty$, this condition implies that the non-negative initial datum u^0 needs to be approximated by strictly positive data $u^0_{\Delta^{(j)}}$, and the smaller one whishes to choose the minimal value of $u^0_{\Delta^{(j)}}$, the finer one needs to make the grid $\boldsymbol{\xi}^{(j)}$. Condition (1.15) thus quantifies the

https://doi.org/10.1051/m2an/2013126 Published online by Cambridge University Press

intuitive requirement that not only the mesh of ξ_k's in $[0, M]$, but also the induced mesh of x_k's in I should become arbitrarily fine in the limit, uniformly for all times $t \in [0, T]$. Consequently, our scheme does not allow to track propagating fronts – like spreading Barenblatt profiles – directly on the discrete level as in [5,19], but it is able to approximate these fronts arbitrarily well with strictly positive solutions if the mesh is sufficiently fine. Note that – thanks to the implicit nature of scheme – we do not need to impose a smallness condition on the time step. Such a (classical) CFL condition has been used *e.g.* for the analysis of the explicit Lagrangian scheme in [12].

1.4. Related results from the literature

Studies on Lagrangian schemes for (1.1) are widely scattered in the literature. Already MacCamy and Sokolovsky [16] present a discretization that is almost identical to ours, for (1.1) with $\mathrm{P}(u) = u^2$ and $V \equiv 0$. Another pioneering work in this direction is the paper by Russo [21], who compares several (semi-)Lagrangian discretizations in the linear case $\mathrm{P}(u) = u$; extensions to two spatial dimensions are also discussed. Later, Budd *et al.* [5] used a moving mesh to capture self-similar solutions of the porous medium equation on the whole line. The general theme was picked up recently by Carrillo and Moll [7], who define a Lagrangian discretization of aggregation equations in two space dimensions, based on the reformulation in terms of evolving diffeomorphisms [10].

The connection between Lagrangian schemes and the gradient flow structure of equation (1.1) was investigated by Kinderlehrer and Walkington [14] and in a series of unpublished theses [15, 19]. In a recent paper by Westdickenberg and Wilkening [24], a similar scheme for (1.1) is obtained as a by-product in the process of designing a structure preserving discretization for the Euler equations. Burger *et al.* [6] devise a numerical scheme for (1.1) in dimension two on basis of the gradient flow structure, using the hydrodynamical formulation of the Wasserstein distance [3] instead of the Lagrangian approach. The Lagrangian approach was adapted to fourth order equations, namely by Cavalli and Naldi [8] for the Hele-Shaw flow, and by Düring *et al.* [9] for the DLSS equation.

In the aforementioned works, numerical schemes are defined and used in experiments; qualitative properties and convergence are not studied analytically. Some analytical investigations have been carried out by Gosse and Toscani [12]: for a Lagrangian scheme with explicit time discretization, they prove comparison principles, and they rigorously discuss stability and consistency. Also, a full discretization of the Keller–Segel model has been analyzed by Blanchet *et al.* [4] in view of convergence to equilibrium. However, to the best of our knowledge, a proof for convergence of discrete to continuous (weak) solutions is not available in the literature.

Finally, a remark is due on an alternative way of proving convergence of the scheme. By use of stability results for gradient flows [1,2] and the machinery of Γ-convergence, it seems likely that Theorem 1.1 can be obtained by exploiting the variational structure more deeply than we do here. In particular, the theory on perturbed λ-contractive gradient flows developed by Serfaty [22] indicates an alternative route towards the same goal. We followed the elementary approach based on a priori estimates here, partly in order to avoid heavy machinery, but mainly with the aim to develop a "stable" concept of proof that generalizes more directly to gradient flows without convexity properties (like fourth order equations).

1.5. Outline of the paper

Section 2 below summarizes some basic results on inverse distribution functions and convexity in the Wasserstein metric. In Section 3, we describe in detail the spatial discretization and study the restrictions of the Wasserstein metric and energy to the $\mathrm{D}^M_{\boldsymbol{\xi}}(I)$. The discrete scheme (1.11) is studied in Section 4, and we derive the Euler–Lagrange equations. Section 5 provides a summary of some qualitative properties of the discretization, like metric contraction and the minimum/maximum principle. The proof of Theorem 1.1 is given in Section 6. The paper concludes with the results of various numerical experiments in Section 7, and with a calculation of the consistency order.

https://doi.org/10.1051/m2an/2013126 Published online by Cambridge University Press

2. Preliminaries and notations

For an introduction to the theory of optimal transportation, we refer to [23]. A comprehensive theory of gradient flows in the Wasserstein metric can be found in [1].

2.1. Inverse distribution functions

Throughout the paper, we shall denote by

$$\mathrm{D}^M(I) := \left\{ u \in L^1(I) \cap L^\infty(I) \,\middle|\, \operatorname*{ess\,inf}_{x\in I} u(x) > 0,\ \int_I u(x)\,\mathrm{d}x = M \right\}$$

the space of positive density functions of total mass M. For $u \in \mathrm{D}^M(I)$, define its distribution function $U : I \to [0, M]$ by

$$U(t;x) = \int_a^x u(t;y)\,\mathrm{d}y,$$

and introduce its inverse function $\mathrm{X} = U^{-1} : [0, M] \to I$. By our choice of $\mathrm{D}^M(I)$, the latter is well-defined and belongs to

$$\mathfrak{X} := \left\{ \mathrm{X} \in C^{0,1}([0,M]; I) \,\middle|\, \mathrm{X}(0) = a,\ \mathrm{X}(M) = b,\ \mathrm{X} \text{ strictly increasing} \right\}.$$

Thanks to the Lipschitz continuity of U and X, we can differentiate the identity $U \circ \mathrm{X}(\xi) = \xi$ at almost every $\xi \in [0, M]$ and obtain the relation

$$u(\mathrm{X}(\xi))\mathrm{X}_\xi(\xi) = 1 \qquad \text{for a.e. } \xi \in [0, M]. \tag{2.1}$$

The inverse distribution function allows for an explicit representation of the Wasserstein distance in one spatial dimension.

Lemma 2.1 (see *e.g.* [23])**.** *Let $u_0, u_1 \in \mathrm{D}^M(I)$ have inverse distribution functions $\mathrm{X}_0, \mathrm{X}_1 \in \mathfrak{X}$. Then their Wasserstein distance amounts to*

$$\mathbf{W}_2(u_0, u_1) = \left(\int_0^M [X_1(\xi) - X_0(\xi)]^2\,\mathrm{d}\xi \right)^{1/2}, \tag{2.2}$$

and a minimal geodesic $(u_s)_{0\le s\le 1}$ connecting u_0 to u_1 in $\mathbf{W}_2$ is given by $X_s = sX_1 + (1-s)X_0$.

2.2. Properties of the energy

Given $\mathrm{P} : \mathbb{R}_{\ge 0} \to \mathbb{R}$, let $\phi : \mathbb{R}_{\ge 0} \to \mathbb{R}$ be a second anti-derivative of $r \mapsto \mathrm{P}'(r)/r$ with $\phi(0) = 0$, and define $\psi : \mathbb{R}_+ \to \mathbb{R}$ by $\psi(s) = s\phi(1/s)$.

Proposition 2.2. *The function ϕ is strictly convex and satisfies*

$$\mathrm{P}(r) = r\phi'(r) - \phi(r) \quad \text{for all } r \ge 0. \tag{2.3}$$

Moreover, $\psi(s) \to \infty$ for $s \downarrow 0$, and ψ'' is a positive non-increasing function.

Proof. The claims about ϕ and ψ are direct consequences of the hypotheses in (1.2). By definition, $\phi''(r) = \mathrm{P}'(r)/r > 0$ for all $r > 0$, so ϕ is strictly convex. (2.3) follows by integration of both sides w.r.t. $r > 0$. It follows further that

$$\begin{aligned}
\psi'(s) &= \phi\left(s^{-1}\right) - s^{-1}\phi'\left(s^{-1}\right) = -\mathrm{P}\left(s^{-1}\right), \\
\psi''(s) &= s^{-2}\mathrm{P}'\left(s^{-1}\right) > 0, \\
\psi'''(s) &= -\,\mathrm{d}^2\mathrm{P}\left(s^{-1}\right)/\,\mathrm{d}s^2 \ge 0,
\end{aligned}$$

https://doi.org/10.1051/m2an/2013126 Published online by Cambridge University Press

so ψ'' is indeed positive and non-increasing. Finally,

$$\lim_{s\downarrow 0}\psi(s)=\psi(1)+\lim_{s\downarrow 0}\int_s^1 \mathrm{P}\left(\sigma^{-1}\right)\,\mathrm{d}\sigma=\psi(1)+\lim_{r\to\infty}\int_1^r\frac{\mathrm{P}(\rho)}{\rho^2}\,\mathrm{d}\rho=+\infty$$

since $\mathrm{P}'(\rho)\to\infty$ for $\rho\to\infty$, and hence also $\mathrm{P}(\rho)/\rho\to\infty$. □

Introduce the functionals $\mathbf{E}$ on $\mathrm{D}^M(I)$ and $\mathbb{E}$ on $\mathfrak{X}$, respectively, by (1.5) and (1.14), then the following result is an easy consequence of the change of variables $x=\mathrm{X}(\xi)$ under the integrals in the definition (1.5).

Lemma 2.3. *For every $u\in\mathrm{D}^M(I)$ with inverse distribution function $\mathrm{X}\in\mathfrak{X}$, one has*

$$\mathbf{E}(u)=\mathbb{E}(\mathrm{X}). \tag{2.4}$$

The convexity of the functional $\mathbf{E}$ with respect to the Wasserstein metric is most conveniently studied when the latter is considered as a functional of X instead of u. Indeed, by Lemma 2.1 above, geodesic interpolation between $u_0,u_1\in\mathrm{D}^M(I)$ corresponds to linear interpolation between $\mathrm{X}_0,\mathrm{X}_1\in\mathfrak{X}$.

Lemma 2.4. *The functional $\mathbb{E}$ is bounded from below,*

$$\mathbb{E}(\mathrm{X})\geq\underline{\mathbf{E}}:=(b-a)\phi\left(\frac{M}{b-a}\right)+M\min_{x\in I}V(x), \tag{2.5}$$

and it is $(-\Lambda)$-convex on $\mathfrak{X}$ with the Λ given in (1.6), i.e.,

$$\mathbb{E}\left((1-s)\mathrm{X}^0+s\mathrm{X}^1\right)\leq(1-s)\mathbb{E}\left(\mathrm{X}^0\right)+s\mathbb{E}\left(\mathrm{X}^1\right)+\frac{\Lambda s(1-s)}{2}\int_0^M\left[\mathrm{X}^0(\xi)-\mathrm{X}^1(\xi)\right]^2\,\mathrm{d}\xi \tag{2.6}$$

for all $\mathrm{X}^0,\mathrm{X}^1\in\mathfrak{X}$, and every $s\in[0,1]$.

Proof. Since ϕ is convex, the lower bound follows by Jensen's inequality:

$$\mathbb{E}(\mathrm{X})\geq M\psi\left(\int_0^M\mathrm{X}_\xi(\xi)\frac{\mathrm{d}\xi}{M}\right)+\int_0^M\min_{x\in I}V(x)\,\mathrm{d}\xi.$$

By definition of ψ, this yields (2.5). Next, let $\mathrm{X}^0,\mathrm{X}^1\in\mathfrak{X}$ and $s\in[0,1]$ be given. Since $\psi:\mathbb{R}_+\to\mathbb{R}$ is convex by hypothesis, it follows in particular that

$$\int_0^M\psi\left((1-s)\mathrm{X}^0_\xi(\xi)+s\mathrm{X}^1_\xi(\xi)\right)\,\mathrm{d}\xi\leq(1-s)\int_0^M\psi\left(\mathrm{X}^0_\xi(\xi)\right)\,\mathrm{d}\xi+s\int_0^M\psi\left(\mathrm{X}^1_\xi(\xi)\right)\,\mathrm{d}\xi.$$

Further, a Taylor expansion yields

$$V\big((1-s)y+sz\big)\leq(1-s)V(y)+sV(z)+\frac{\Lambda}{2}s(1-s)(y-z)^2$$

for arbitrary $y,z\in I$. In combination, this implies inequality (2.6). □

3. Spatial discretization

Inside the space $\mathfrak{X}$ of inverse distribution functions, we define the finite-dimensional subspace $\mathfrak{X}_{\boldsymbol{\xi}}$ of those functions, which are piecewise affine with respect to a given partition $\boldsymbol{\xi}$ of $[0,M]$ into sub-intervals. Correspondingly, there is a finite-dimensional submanifold $\mathrm{D}^M_{\boldsymbol{\xi}}(I)$ of $\mathrm{D}^M(I)$ consisting of those densities, whose inverse distribution functions belong to $\mathfrak{X}_{\boldsymbol{\xi}}$. Densities in $\mathrm{D}^M_{\boldsymbol{\xi}}(I)$ are piecewise constant. Since we shall work simultaneously in the spaces $\mathrm{D}^M_{\boldsymbol{\xi}}(I)$ and $\mathfrak{X}_{\boldsymbol{\xi}}$, we need to introduce various notations.

https://doi.org/10.1051/m2an/2013126 Published online by Cambridge University Press

3.1. Ansatz spaces

A vector $\boldsymbol{\xi} = (\xi_0, \xi_1, \dots, \xi_K)$ with entries ξ_j such that

$$0 = \xi_0 < \xi_1 < \cdots < \xi_K = M$$

defines a partition of $[0, M]$ into K sub-intervals. We denote the lengths of the intervals by

$$\delta_k = \xi_k - \xi_{k-1} \quad \text{for all } k = 1, \dots, K,$$

and introduce further

$$\underline{\delta}(\boldsymbol{\xi}) = \min_k \delta_k, \quad \overline{\delta}(\boldsymbol{\xi}) = \max_k \delta_k, \quad \alpha(\boldsymbol{\xi}) = \frac{\overline{\delta}(\boldsymbol{\xi})}{\underline{\delta}(\boldsymbol{\xi})}. \tag{3.1}$$

The associated $(K-1)$-dimensional reduction of the space $\mathfrak{X}$ is given by

$$\mathfrak{X}_{\boldsymbol{\xi}} := \left\{ \mathrm{X} \in \mathfrak{X} \,\middle|\, \mathrm{X} \text{ piecewise affine on each } [\xi_{k-1}, \xi_k], \text{ for } k = 1, \dots, K \right\}.$$

Functions in $\mathfrak{X}_{\boldsymbol{\xi}}$ are conveniently represented as linear combinations of the $K+1$ hat functions θ_0 to θ_K defined by

$$\theta_m(\xi) = \begin{cases} (\xi - \xi_{m-1})/\delta_m & \text{for } \xi_{m-1} \le \xi \le \xi_m \ (\text{if } m \ge 1), \\ (\xi_{m+1} - \xi)/\delta_{m+1} & \text{for } \xi_m \le \xi \le \xi_{m+1} \ (\text{if } m \le K-1), \\ 0 & \text{otherwise.} \end{cases}$$

More precisely, there is a one-to-one correspondence between functions $\mathrm{X} \in \mathfrak{X}_{\boldsymbol{\xi}}$ and vectors in

$$\mathfrak{x}_{\boldsymbol{\xi}} := \left\{ \vec{\mathrm{x}} = (x_1, \dots, x_{K-1}) \,\middle|\, a < x_1 < x_2 < \cdots < x_{K-1} < b \right\} \subset I^{K-1};$$

this correspondence is established by means of $\mathbf{X}_{\boldsymbol{\xi}} : \mathfrak{x}_{\boldsymbol{\xi}} \to \mathfrak{X}_{\boldsymbol{\xi}}$, with

$$\mathrm{X} = \mathbf{X}_{\boldsymbol{\xi}}[\vec{\mathrm{x}}] = \sum_{k=0}^{K} x_k \theta_k. \tag{3.2}$$

Remark 3.1. By definition, $\vec{\mathrm{x}} \in \mathfrak{x}_{\boldsymbol{\xi}}$ has components x_1 to x_{K-1}. In (3.2) above and for the rest of the paper, we shall always use the convention that

$$x_0 = a \quad \text{and} \quad x_K = b. \tag{3.3}$$

Also, we introduce in analogy to $\overline{\delta}(\boldsymbol{\xi})$ the mesh width

$$\overline{\delta}(\vec{\mathrm{x}}) = \max_k (x_k - x_{k-1}). \tag{3.4}$$

Occasionally, it will be more convenient to work with vectors $\vec{z} \in \mathbb{R}_+^K$ of difference quotients: define $\mathbf{z}_{\boldsymbol{\xi}} : \mathfrak{x}_{\boldsymbol{\xi}} \to \mathbb{R}_+^K$ by

$$\vec{\mathrm{z}} = (z_1, \dots, z_K) = \mathbf{z}_{\boldsymbol{\xi}}[\vec{\mathrm{x}}] \in \mathbb{R}_+^K, \quad \text{with} \quad z_k = \frac{x_k - x_{k-1}}{\delta_k}, \tag{3.5}$$

using again our convention (3.3).

https://doi.org/10.1051/m2an/2013126 Published online by Cambridge University Press

Finally, we introduce the associated $(K-1)$-dimensional submanifold $\mathrm{D}^M_{\boldsymbol{\xi}}(I) := \mathbf{u}_{\boldsymbol{\xi}}[\mathfrak{x}_{\boldsymbol{\xi}}] \subset \mathrm{D}^M(I)$ as the image of the injective map $\mathbf{u}_{\boldsymbol{\xi}} : \mathfrak{x}_{\boldsymbol{\xi}} \to \mathrm{D}^M_{\boldsymbol{\xi}}(I)$ with

$$\mathbf{u}_{\boldsymbol{\xi}}[\vec{\mathrm{x}}] = \sum_{k=1}^{K} u_k \mathbf{1}_{(x_{k-1},x_k]}, \quad \text{where} \quad u_k = \frac{\xi_k - \xi_{k-1}}{x_k - x_{k-1}}. \tag{3.6}$$

3.2. Representation of the Wasserstein distance

The Wasserstein distance between any two elements of $\mathrm{D}^M_{\boldsymbol{\xi}}(I)$ is easy to compute using (2.2).

Lemma 3.2. *Fix a discretization $\boldsymbol{\xi}$, and let $u^0, u^1 \in \mathrm{D}^M_{\boldsymbol{\xi}}(I)$ have representations $u^0 = \mathbf{u}_{\boldsymbol{\xi}}[\vec{\mathrm{x}}^0]$, $u^1 = \mathbf{u}_{\boldsymbol{\xi}}[\vec{\mathrm{x}}^1]$, with $\vec{\mathrm{x}}^0, \vec{\mathrm{x}}^1 \in \mathfrak{x}_{\boldsymbol{\xi}}$. Then*

$$\mathbf{W}_2\left(u^0, u^1\right)^2 = \left(\vec{\mathrm{x}}^0 - \vec{\mathrm{x}}^1\right)^T \mathrm{W} \left(\vec{\mathrm{x}}^0 - \vec{\mathrm{x}}^1\right) \tag{3.7}$$

with the symmetric tridiagonal matrix $\mathrm{W} = (\mathrm{W}_{m,k})_{m,k=1}^{K-1} \in \mathbb{R}^{(K-1)\times(K-1)}$ given by

$$\mathrm{W} = \frac{1}{6}\begin{pmatrix} 2(\delta_1+\delta_2) & \delta_2 & \cdots & 0 & 0 \\ \delta_2 & 2(\delta_2+\delta_3) & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 2(\delta_{K-2}+\delta_{K-1}) & \delta_{K-1} \\ 0 & 0 & \cdots & \delta_{K-1} & 2(\delta_{K-1}+\delta_K) \end{pmatrix}. \tag{3.8}$$

Moreover, for every $v \in \mathbb{R}^{K-1}$,

$$\frac{1}{6}\sum_{k=1}^{K-1}(\delta_k+\delta_{k+1})v_k^2 \le v^T \mathrm{W} v \le \frac{1}{2}\sum_{k=1}^{K-1}(\delta_k+\delta_{k+1})v_k^2. \tag{3.9}$$

Proof. The first statement of this lemma follows by straight-forward calculations. To prove (3.9), let $v \in \mathbb{R}^{K-1}$ be given and observe that

$$\begin{aligned} 3v^T \mathrm{W} v &= \sum_{m=1}^{K-1}(\delta_m+\delta_{m+1})v_m^2 + \sum_{m=2}^{K-1}\delta_m v_m v_{m-1} \\ &= \delta_1 v_1^2 + \delta_K v_{K-1}^2 + \sum_{m=2}^{K-1}\delta_m\left(v_m^2 + v_{m-1}^2\right) + \sum_{m=1}^{K}\delta_m v_m v_{m-1}. \end{aligned}$$

From here, (3.9) is immediately deduced using binomial formulas. □

3.3. Representation of the energy

The restriction of the energy $\mathbb{E}$ from (2.4) to the subspace $\mathfrak{X}_{\boldsymbol{\xi}}$ is naturally associated to the functional $\mathbb{E}_{\boldsymbol{\xi}} : \mathfrak{x}_{\boldsymbol{\xi}} \to \mathbb{R}$ with

$$\mathbb{E}_{\boldsymbol{\xi}}(\vec{\mathrm{x}}) := \mathbb{E}\left(\mathbf{X}_{\boldsymbol{\xi}}[\vec{\mathrm{x}}]\right) = \mathbf{E}\left(\mathbf{u}_{\boldsymbol{\xi}}[\vec{\mathrm{x}}]\right). \tag{3.10}$$

By straight-forward calculations, one obtains the following more explicit representation.

https://doi.org/10.1051/m2an/2013126 Published online by Cambridge University Press

Lemma 3.3. *For every $\vec{\mathrm{x}} \in \mathfrak{x}_{\boldsymbol{\xi}}$, we have*

$$\mathbb{E}_{\boldsymbol{\xi}}(\vec{\mathrm{x}}) = \sum_{k=1}^{K} \delta_k \psi\left(\frac{x_k - x_{k-1}}{\delta_k}\right) + \int_0^M V\left(\mathbf{X}_{\boldsymbol{\xi}}[\vec{\mathrm{x}}](\xi)\right) \mathrm{d}\xi. \tag{3.11}$$

Moreover, the (Euclidean) gradient vector $\partial_{\vec{\mathrm{x}}}\mathbb{E}_{\boldsymbol{\xi}}(\vec{\mathrm{x}}) = \left(\partial_{x_m}\mathbb{E}_{\boldsymbol{\xi}}(\vec{\mathrm{x}})\right)_{m=1}^{K-1} \in \mathbb{R}^{K-1}$ is given by

$$\left[\partial_{\vec{\mathrm{x}}}\mathbb{E}_{\boldsymbol{\xi}}(\vec{\mathrm{x}})\right]_m = -\psi'\left(\frac{x_{m+1} - x_m}{\delta_{m+1}}\right) + \psi'\left(\frac{x_m - x_{m-1}}{\delta_m}\right) + \int_0^M V_x\left(\mathbf{X}_{\boldsymbol{\xi}}[\vec{\mathrm{x}}](\xi)\right) \theta_m(\xi)\, \mathrm{d}\xi, \tag{3.12}$$

and the Hessian matrix $\partial^2_{\vec{\mathrm{x}}}\mathbb{E}_{\boldsymbol{\xi}}(\vec{\mathrm{x}}) = \left(\partial_{x_m x_k}\mathbb{E}_{\boldsymbol{\xi}}(\vec{\mathrm{x}})\right)_{m,k=1}^{K-1} \in \mathbb{R}^{(K-1)\times(K-1)}$ is symmetric with

$$\left[\partial^2_{\vec{\mathrm{x}}}\mathbb{E}_{\boldsymbol{\xi}}\left(\vec{\mathrm{x}}\right)\right]_{m,k} = \begin{cases} \dfrac{1}{\delta_{m+1}}\psi''\left(\dfrac{x_{m+1} - x_m}{\delta_{m+1}}\right) + \dfrac{1}{\delta_m}\psi''\left(\dfrac{x_m - x_{m-1}}{\delta_m}\right) \\ \qquad\qquad + \displaystyle\int_0^M V_{xx}\left(\mathbf{X}_{\boldsymbol{\xi}}\left[\vec{\mathrm{x}}\right]\right)\theta_m^2\, \mathrm{d}\xi \quad \text{if } m = k, \\ -\dfrac{1}{\delta_m}\psi''\left(\dfrac{x_m - x_{m-1}}{\delta_m}\right) + \displaystyle\int_0^M V_{xx}\left(\mathbf{X}_{\boldsymbol{\xi}}\left[\vec{\mathrm{x}}\right]\right)\theta_m\theta_{m-1}\, \mathrm{d}\xi \quad \text{if } k = m-1, \\ 0 \qquad\qquad \text{if } 1 \le k < m-1. \end{cases} \tag{3.13}$$

Further, the functional $\mathbb{E}_{\boldsymbol{\xi}}$ inherits boundedness and convexity from $\mathbb{E}$.

Lemma 3.4. *$\mathbb{E}_{\boldsymbol{\xi}}$ is bounded from below by $\underline{\mathbf{E}}$ defined in (2.5). Further, it is $(-\Lambda)$-convex with respect to the quadratic structure induced by* W, i.e., *$\nabla^2\mathbb{E}_{\boldsymbol{\xi}}(\vec{\mathrm{x}}) + \Lambda\mathrm{W}$ is positive semi-definite for arbitrary $\vec{\mathrm{x}}^0 \in \mathfrak{x}_{\boldsymbol{\xi}}$. Consequently,*

$$\left(\vec{\mathrm{x}}^1 - \vec{\mathrm{x}}^0\right)^T \left(\partial_{\vec{\mathrm{x}}}\mathbb{E}_{\boldsymbol{\xi}}\left(\vec{\mathrm{x}}^1\right) - \partial_{\vec{\mathrm{x}}}\mathbb{E}_{\boldsymbol{\xi}}\left(\vec{\mathrm{x}}^0\right)\right) \ge -\Lambda\left(\vec{\mathrm{x}}^1 - \vec{\mathrm{x}}^0\right)^T \mathrm{W}\left(\vec{\mathrm{x}}^1 - \vec{\mathrm{x}}^0\right) \tag{3.14}$$

holds for every $\vec{\mathrm{x}}^0, \vec{\mathrm{x}}^1 \in \mathfrak{x}_{\boldsymbol{\xi}}$.

Proof. Boundedness from below is a trivial consequence of (2.5) and the definition of $\mathbb{E}_{\boldsymbol{\xi}}$ by restriction of $\mathbb{E}$. Convexity is a direct consequence of the convexity (2.6) of $\mathbb{E}$, taking into account (3.7), and that $\mathbf{X}_{\boldsymbol{\xi}}$ is an affine map. The estimate (3.14) is obtained by Taylor expansion. □

4. Time-discrete evolution

Throughout this section, we fix a pair $\Delta = (\tau, \boldsymbol{\xi})$ of a time step with $\tau > 0$ and a spatial discretization $\boldsymbol{\xi} = (\xi_0, \dots, \xi_K)$.

4.1. Minimizing movements

To start with, we verify that the procedure (1.11) can be used to define inductively – starting from a prescribed initial datum $u^0_\Delta \in \mathrm{D}^M_{\boldsymbol{\xi}}(I)$ – a *discrete solution* $u_\Delta := (u^n_\Delta)_{n=0}^\infty$. By definition of $\mathrm{D}^M_{\boldsymbol{\xi}}(I)$ as the image of $\mathfrak{x}_{\boldsymbol{\xi}}$ under $\mathbf{u}_{\boldsymbol{\xi}}$, and by (2.2) and (3.10),

$$u^n_\Delta = \mathbf{u}_{\boldsymbol{\xi}}\left[\vec{\mathrm{x}}^n_\Delta\right] \tag{4.1}$$

establishes a one-to-one correspondence between discrete solutions u_Δ and sequences $(\vec{\mathrm{x}}^n_\Delta)_{n=0}^\infty$, where each $\vec{\mathrm{x}}^n_\Delta \in \mathfrak{x}_{\boldsymbol{\xi}}$ is a minimizer in

$$\mathbb{E}_\Delta\left(\vec{\mathrm{x}}, \vec{\mathrm{x}}^{n-1}_\Delta\right) := \frac{1}{2\tau}\left(\vec{\mathrm{x}} - \vec{\mathrm{x}}^{n-1}_\Delta\right)^T \mathrm{W}\left(\vec{\mathrm{x}} - \vec{\mathrm{x}}^{n-1}_\Delta\right) + \mathbb{E}_{\boldsymbol{\xi}}\left(\vec{\mathrm{x}}\right) \quad \to \quad \min. \tag{4.2}$$

https://doi.org/10.1051/m2an/2013126 Published online by Cambridge University Press

Proposition 4.1. *Assume that $\tau\Lambda < 1$, with $\Lambda \geq 0$ defined in (1.6), and let $u_\Delta^0 = \mathbf{u}_{\boldsymbol{\xi}}[\vec{\mathrm{x}}_\Delta^0] \in \mathrm{D}_{\boldsymbol{\xi}}^M(I)$ be given. Then there are sequences $(u_\Delta^n)_{n=0}^\infty$ and $(\vec{\mathrm{x}}_\Delta^n)_{n=0}^\infty$, related by (4.1), such that $u_\Delta^n \in \mathrm{D}_{\boldsymbol{\xi}}^M(I)$ is the unique minimizer of $\mathbf{E}_\tau(\cdot, u_\Delta^{n-1})$ on $\mathrm{D}_{\boldsymbol{\xi}}^M(I)$, and $\vec{\mathrm{x}}_\Delta^n$ is the unique minimizer of $\mathbb{E}_\Delta(\cdot, \vec{\mathrm{x}}_\Delta^{n-1})$, for every $n \in \mathbb{N}$. Moreover, $\vec{\mathrm{x}}_\Delta^n$ is the unique solution in $\mathfrak{x}_{\boldsymbol{\xi}}$ to the system of Euler–Lagrange equations*

$$\frac{1}{\tau}\mathrm{W}\left(\vec{\mathrm{x}} - \vec{\mathrm{x}}_\Delta^{n-1}\right) = -\partial_{\vec{\mathrm{x}}}\mathbb{E}_{\boldsymbol{\xi}}\left(\vec{\mathrm{x}}\right), \tag{4.3}$$

with $\partial_{\vec{\mathrm{x}}}\mathbb{E}_{\boldsymbol{\xi}}(\vec{\mathrm{x}})$ explicitly given in (3.12).

Proof. It suffices to prove unique solvability of (4.2) for $\vec{\mathrm{x}} \in \mathfrak{x}_{\boldsymbol{\xi}}$. To this end, observe that

$$\mathbb{E}_\Delta\left(\vec{\mathrm{x}}, \vec{\mathrm{x}}_\Delta^{n-1}\right) = \mathbb{E}_{\boldsymbol{\xi}}\left(\vec{\mathrm{x}}\right) + \frac{\Lambda}{2}\left(\vec{\mathrm{x}} - \vec{\mathrm{x}}_\Delta^{n-1}\right)^T \mathrm{W}\left(\vec{\mathrm{x}} - \vec{\mathrm{x}}_\Delta^{n-1}\right) + \frac{1}{2}\left(\tau^{-1} - \Lambda\right)\left(\vec{\mathrm{x}} - \vec{\mathrm{x}}_\Delta^{n-1}\right)^T \mathrm{W}\left(\vec{\mathrm{x}} - \vec{\mathrm{x}}_\Delta^{n-1}\right)$$

for every $\vec{\mathrm{x}} \in \mathfrak{x}_{\boldsymbol{\xi}}$. From Lemma 2.4, we know that the sum of the first two terms on the right-hand side constitutes a convex function in $\vec{\mathrm{x}} \in \mathfrak{x}_{\boldsymbol{\xi}}$. Since $\tau\Lambda < 1$, and since W is positive definite by Lemma 3.2, the last term is strictly convex. Thus, $\mathbb{E}_\Delta(\cdot, \vec{\mathrm{x}}_\Delta^{n-1})$ possesses at most one critical point in $\mathfrak{x}_{\boldsymbol{\xi}}$.

To show the existence of a minimizer, let $(\vec{\mathrm{x}}^{(j)})_{j\in\mathbb{N}}$ be a minimizing sequence for $\mathbb{E}_\Delta(\cdot, \vec{\mathrm{x}}^{n-1})$ in $\mathfrak{x}_{\boldsymbol{\xi}}$. Since each of the $K-1$ components $x_k^{(j)}$ belongs to the compact interval I, we may assume without loss of generality that $\vec{\mathrm{x}}^{(j)}$ converges to some $\vec{\mathrm{x}}^* \in I^{K-1}$. It remains to be proven that $\vec{\mathrm{x}}^* \in \mathfrak{x}_{\boldsymbol{\xi}}$. Since $(\vec{\mathrm{x}}^{(j)})_{j\in\mathbb{N}}$ is a minimizing sequence, $\mathbb{E}_\Delta(\vec{\mathrm{x}}^{(j)}, \vec{\mathrm{x}}^{n-1})$ is bounded, and so, for every $m \in \{1, \ldots, K\}$:

$$\begin{aligned} C &\geq \frac{1}{2\tau}\left(\vec{\mathrm{x}}^{(j)} - \vec{\mathrm{x}}_\Delta^{n-1}\right)^T \mathrm{W}\left(\vec{\mathrm{x}}^{(j)} - \vec{\mathrm{x}}_\Delta^{n-1}\right) + \sum_{k=1}^K \delta_k \psi\left(\frac{x_k^{(j)} - x_{k-1}^{(j)}}{\delta_k}\right) + \int_0^M V\left(\mathbf{X}_{\boldsymbol{\xi}}\left[\vec{\mathrm{x}}^{(j)}\right](\xi)\right) \mathrm{d}\xi \\ &\geq \delta_m \psi\left(\frac{x_m^{(j)} - x_{m-1}^{(j)}}{\delta_m}\right) + (M - \delta_m)\psi\left(\frac{b - a - (x_m^{(j)} - x_{m-1}^{(j)})}{M - \delta_m}\right) + M \min_{x\in I} V(x). \end{aligned}$$

Since $\psi(s) \to \infty$ for $s \downarrow 0$, this implies that $x_m^{(j)} - x_{m-1}^{(j)} \geq \epsilon\delta_m > 0$ with some $\epsilon > 0$ for all j, and thus also $x_m^* - x_{m-1}^* \geq \epsilon\delta_m > 0$, implying $\vec{\mathrm{x}}^* \in \mathfrak{x}_{\boldsymbol{\xi}}$. By continuity of $\mathbb{E}_\Delta(\cdot, \vec{\mathrm{x}}_\Delta^{n-1})$ on $\mathfrak{x}_{\boldsymbol{\xi}}$, it follows that $\vec{\mathrm{x}}^*$ is a minimizer.

Consequently, $\mathbb{E}_\Delta(\cdot, \vec{\mathrm{x}}_\Delta^{n-1})$ possesses a unique critical point in $\mathfrak{x}_{\boldsymbol{\xi}}$, thus the corresponding Euler–Lagrange equations (4.3) are uniquely solvable. □

4.2. Euler–Lagrange equations for the difference quotients

The analysis that follows will be based primarily on another representation of the system (4.3) of Euler–Lagrange equations, which is formulated in terms of the difference quotients $\vec{z}_\Delta^n = (z_1^n, \ldots, z_K^n)$ introduced in (3.5):

$$z_k^n = \frac{x_k^n - x_{k-1}^n}{\delta_k} = \frac{1}{u_k^n}. \tag{4.4}$$

Introduce quadratic analogues $\Theta_1, \ldots, \Theta_K : [0, M] \to \mathbb{R}$ of the piecewise linear hat functions $\theta_0, \ldots, \theta_K$ as follows. Let the numbers $\gamma_1, \ldots, \gamma_K \in (-1, 1)$ be defined by $\gamma_1 = \gamma_K = 0$, and

$$\gamma_k = \frac{\delta_{k+1} - \delta_{k-1}}{\delta_{k+1} + 2\delta_k + \delta_{k-1}} \quad \text{for } k = 2, \ldots, K-1. \tag{4.5}$$

https://doi.org/10.1051/m2an/2013126 Published online by Cambridge University Press

Then the Θ_k are given by

$$\Theta_k(\xi) = \begin{cases} \frac{1+\gamma_k}{2\delta_{k-1}}(\xi - \xi_{k-2})^2 & \text{if } \xi_{k-2} \le \xi \le \xi_{k-1}, \\ \frac{1-\gamma_k^2}{4}(\delta_{k+1} + \delta_k + \delta_{k-1}) - \frac{1}{4\delta_k}\left(2\xi - (\xi_k + \xi_{k-1}) - \gamma_k\delta_k\right)^2 & \text{if } \xi_{k-1} \le \xi \le \xi_k, \\ \frac{1-\gamma_k}{2\delta_{k+1}}(\xi_{k+1} - \xi)^2 & \text{if } \xi_k \le \xi \le \xi_{k+1}, \\ 0 & \text{otherwise} \end{cases}$$

for $k = 2, \dots, K-1$, and by

$$\Theta_1(\xi) = \begin{cases} \frac{1}{2}(\delta_1 + \delta_2) - \frac{1}{2\delta_1}\xi^2 & \text{if } 0 \le \xi \le \xi_1, \\ \frac{1}{2\delta_2}(\xi_2 - \xi)^2 & \text{if } \xi_1 \le \xi \le \xi_2, \\ 0 & \text{otherwise,} \end{cases}$$

$$\Theta_K(\xi) = \begin{cases} \frac{1}{2\delta_{K-1}}(\xi - \xi_{K-2})^2 & \text{if } \xi_{K-2} \le \xi \le \xi_{K-1}, \\ \frac{1}{2}(\delta_K + \delta_{K-1}) - \frac{1}{2\delta_K}(M - \xi)^2 & \text{if } \xi_{K-1} \le \xi \le M, \\ 0 & \text{otherwise.} \end{cases}$$

For each $k = 2, \dots, K-1$, the function Θ_k is supported on $[\xi_{k-2}, \xi_{k+1}]$, and it satisfies

$$-(\Theta_k)_\xi = (1-\gamma_k)\theta_k - (1+\gamma_k)\theta_{k-1}. \tag{4.6}$$

And $(\Theta_1)_\xi = -\theta_1$, $(\Theta_K)_\xi = \theta_{K-1}$. Next, we define the matrix $\widetilde{\mathrm{W}} = (\widetilde{\mathrm{W}}_{m,k})_{m,k=1}^K \in \mathbb{R}^{K\times K}$ by

$$\widetilde{\mathrm{W}}_{m,k} = \int_{\xi_{k-1}}^{\xi_k} \Theta_m(\xi)\,\mathrm{d}\xi. \tag{4.7}$$

The matrix $\widetilde{\mathrm{W}}$ essentially plays the same role for the $\vec{\mathrm{z}}_\Delta^n$ as W for the $\vec{\mathrm{x}}_\Delta^n$.

Lemma 4.2. *The matrix $\widetilde{\mathrm{W}}$ is tri-diagonal and has entries*

$$\widetilde{\mathrm{W}}_{m,k} = \begin{cases} \frac{1}{6}\delta_m^2 + \frac{1-\gamma_m}{4}\delta_m\delta_{m+1} + \frac{1+\gamma_m}{4}\delta_m\delta_{m-1} & \text{if } 2 \le k = m \le K-1, \\ \frac{1}{3}\delta_1^2 + \frac{1}{2}\delta_2\delta_1 & \text{if } k = m = 1, \\ \frac{1}{3}\delta_K^2 + \frac{1}{2}\delta_{K-1}\delta_K & \text{if } k = m = K, \\ \frac{1+\gamma_m}{6}\delta_{m-1}^2 & \text{if } k = m-1, \\ \frac{1-\gamma_m}{6}\delta_{m+1}^2 & \text{if } k = m+1, \\ 0 & \text{otherwise.} \end{cases} \tag{4.8}$$

Proof. The explicit representation of $\widetilde{\mathrm{W}}$ is obtained by a tedious, but straight-forward computation that requires nothing but integration of quadratic polynomials and the use of (4.5). □

Lemma 4.3. *For the solution $(\vec{\mathrm{x}}_\Delta^n)_{n=0}^\infty$ obtained in Proposition 4.1, let $(\vec{\mathrm{z}}_\Delta^n)_{n=0}^\infty$ be the associated sequence from (4.4). Then each $\vec{\mathrm{z}}_\Delta^n$ satisfies the following system of Euler–Lagrange equations:*

$$\begin{aligned} \frac{1}{\tau}\left[\widetilde{\mathrm{W}}\left(\vec{\mathrm{z}}_\Delta^n - \vec{\mathrm{z}}_\Delta^{n-1}\right)\right]_m &= (1-\gamma_m)\psi'\left(z_{m+1}^n\right) - 2\psi'\left(z_m^n\right) + (1+\gamma_m)\psi'\left(z_{m-1}^n\right) \\ &\quad - \sum_{k=m-1}^{m+1} z_k^n \int_{\xi_{k-1}}^{\xi_k} V_{xx}\left(\mathbf{X}_{\boldsymbol{\xi}}\left[\vec{\mathrm{x}}^n\right]\right)\Theta_m\,\mathrm{d}\xi \end{aligned} \tag{4.9}$$

https://doi.org/10.1051/m2an/2013126 Published online by Cambridge University Press

for every $m = 2, \dots, K-1$, *and*

$$\frac{1}{\tau}\left[\widetilde{\mathrm{W}}\left(\vec{\mathrm{z}}_\Delta^n - \vec{\mathrm{z}}_\Delta^{n-1}\right)\right]_1 = \psi'\left(z_2^n\right) - \psi'\left(z_1^n\right) - \sum_{k=1}^{2} z_k^n \int_{\xi_{k-1}}^{\xi_k} V_{xx}\left(\mathbf{X}_{\boldsymbol{\xi}}\left[\vec{\mathrm{x}}^n\right]\right) \Theta_1 \,\mathrm{d}\xi, \tag{4.10}$$

$$\frac{1}{\tau}\left[\widetilde{\mathrm{W}}\left(\vec{\mathrm{z}}_\Delta^n - \vec{\mathrm{z}}_\Delta^{n-1}\right)\right]_K = \psi'\left(z_{K-1}^n\right) - \psi'\left(z_K^n\right) - \sum_{k=K-1}^{K} z_k^n \int_{\xi_{k-1}}^{\xi_k} V_{xx}\left(\mathbf{X}_{\boldsymbol{\xi}}\left[\vec{\mathrm{x}}^n\right]\right) \Theta_K \,\mathrm{d}\xi. \tag{4.11}$$

Proof. Fix an index $m \in \{2, \dots, K-1\}$. With γ_m given by (4.5), multiply the mth and the $(m-1)$th component of the Euler–Lagrange system (4.3) by $1-\gamma_m$ and $1+\gamma_m$, respectively, and substract the latter from the first. This yields

$$\begin{aligned}
&\frac{1}{\tau}\left(\frac{1-\gamma_m}{6}\delta_{m+1}x_{m+1}^n + \left[\frac{1-\gamma_m}{3}(\delta_m+\delta_{m+1}) - \frac{1+\gamma_m}{6}\delta_m\right]x_m^n\right.\\
&\quad\left. + \left[\frac{1-\gamma_m}{6}\delta_{m-1} - \frac{1+\gamma_m}{3}(\delta_m+\delta_{m-1})\right]x_{m-1}^n - \frac{1+\gamma_m}{6}\delta_{m-1}x_{m-2}^n\right)\\
&= (1-\gamma_m)\psi'\left(z_{m+1}^n\right) - 2\psi'\left(z_m^n\right) + (1+\gamma_m)\psi'\left(z_{m-1}^n\right)\\
&\quad - \int_0^M V_x\left(\mathbf{X}_{\boldsymbol{\xi}}\left[\vec{\mathrm{x}}^n\right]\right)\left[(1-\gamma_m)\theta_m - (1+\gamma_m)\theta_{m-1}\right]\,\mathrm{d}\xi.
\end{aligned}$$

The expression on the left-hand side can be rewritten as

$$\begin{aligned}
&\frac{1-\gamma_m}{6}\delta_{m+1}\left(x_{m+1}^n - x_m^n\right) + \left[\frac{1-\gamma_m}{2}\delta_{m+1} + \frac{1-3\gamma_m}{6}\delta_m\right]x_m^n\\
&\quad - \left[\frac{1+\gamma_m}{2}\delta_{m-1} + \frac{1+3\gamma_m}{6}\delta_m\right]x_{m-1}^n + \frac{1+\gamma_m}{6}\delta_{m-1}\left(x_{m-1}^n - x_{m-1}^n\right)\\
&= \widetilde{\mathrm{W}}_{m,m+1}z_{m+1}^n + \widetilde{\mathrm{W}}_{m,m}z_m^n + \widetilde{\mathrm{W}}_{m,m-1}z_{m-1}^n,
\end{aligned}$$

where we have used the relation

$$\frac{1-\gamma_m}{2}\delta_{m+1} + \frac{1-3\gamma_m}{6}\delta_m = \frac{1+\gamma_m}{2}\delta_{m-1} + \frac{1+3\gamma_m}{6}\delta_m = \frac{1}{\delta_m}\widetilde{\mathrm{W}}_{m,m},$$

which is a consequence of our definition of γ_m in (4.5). Thus, we obtain

$$\begin{aligned}
\frac{1}{\tau}\left[\widetilde{\mathrm{W}}\left(\vec{\mathrm{z}}_\Delta^n - \vec{\mathrm{z}}_\Delta^{n-1}\right)\right]_m &= (1-\gamma_m)\psi'\left(z_{m+1}^n\right) - 2\psi'\left(z_m^n\right) + (1+\gamma_m)\psi'\left(z_{m-1}^n\right)\\
&\quad - \int_0^M V_x\left(\mathbf{X}_{\boldsymbol{\xi}}\left[\vec{\mathrm{x}}_\Delta^n\right]\right)\left[(1-\gamma_m)\theta_m - (1+\gamma_m)\theta_{m-1}\right]\,\mathrm{d}\xi.
\end{aligned} \tag{4.12}$$

Using property (4.6) of the function Θ_m, and integrating by parts, we arrive at

$$\begin{aligned}
-\int_0^M V_x\left(\mathbf{X}_{\boldsymbol{\xi}}\left[\vec{\mathrm{x}}_\Delta^n\right]\right)\left[(1-\gamma_m)\theta_m - (1+\gamma_m)\theta_{m-1}\right]\,\mathrm{d}\xi &= \int_0^M V_x\left(\mathbf{X}_{\boldsymbol{\xi}}\left[\vec{\mathrm{x}}_\Delta^n\right]\right)\left(\Theta_m\right)_\xi \,\mathrm{d}\xi = V_x\left(\mathbf{X}_{\boldsymbol{\xi}}\left[\vec{\mathrm{x}}_\Delta^n\right]\right)\left(\Theta_m\right)_\xi\Big|_{\xi=0}^{\xi=M}\\
&\quad - \int_0^M \left(\mathbf{X}_{\boldsymbol{\xi}}\left[\vec{\mathrm{x}}_\Delta^n\right]\right)_\xi V_{xx}\left(\mathbf{X}_{\boldsymbol{\xi}}\left[\vec{\mathrm{x}}_\Delta^n\right]\right)\Theta_m \,\mathrm{d}\xi.
\end{aligned} \tag{4.13}$$

The boundary terms in the second line vanish, since $\Theta_m(0) = \Theta_m(M) = 0$. Finally, observe that $\left(\mathbf{X}_{\boldsymbol{\xi}}[\vec{\mathrm{x}}_\Delta^n]\right)_\xi(\xi) = z_k$ for all $\xi_{k-1} < \xi < \xi_k$. Insert the result into (4.12) to arrive at (4.9).

Equations (4.10) and (4.11) are directly obtained from (4.3) for $m = 1$ and for $m = K-1$, respectively, after an integration by parts like in (4.13). The boundary term vanishes because of hypothesis (1.3). □

https://doi.org/10.1051/m2an/2013126 Published online by Cambridge University Press

4.3. Energy dissipation

The estimates derived below are at the core of our convergence proof in Section 6. We start with two energy-type estimates, which are classical in the theory of gradient flows. We recall that $u_\Delta^n = \mathbf{u}_{\boldsymbol{\xi}}[\vec{\mathrm{x}}_\Delta^n]$, and that the gradient $\partial_{\vec{\mathrm{x}}}\mathbb{E}_{\boldsymbol{\xi}}$ is explicitly given in (3.12).

Lemma 4.4. *For every $N \in \mathbb{N}$,*

$$\frac{1}{2\tau}\sum_{n=1}^{N}\mathbf{W}_2\left(u_\Delta^n, u_\Delta^{n-1}\right)^2 \le \mathbf{E}_{\boldsymbol{\xi}}\left(u_\Delta^0\right) - \mathbf{E}_{\boldsymbol{\xi}}\left(u_\Delta^N\right), \tag{4.14}$$

$$\frac{\tau}{2}\sum_{n=1}^{N}\left[\partial_{\vec{\mathrm{x}}}\mathbb{E}_{\boldsymbol{\xi}}\left(\vec{\mathrm{x}}_\Delta^n\right)\right]^T \mathrm{W}^{-1}\left[\partial_{\vec{\mathrm{x}}}\mathbb{E}_{\boldsymbol{\xi}}\left(\vec{\mathrm{x}}_\Delta^n\right)\right] \le \mathbf{E}_{\boldsymbol{\xi}}\left(u_\Delta^0\right) - \mathbf{E}_{\boldsymbol{\xi}}\left(u_\Delta^N\right). \tag{4.15}$$

Proof. Since u_Δ^n minimizes $\mathbf{E}_\Delta(\cdot, u_\Delta^{n-1})$, we have in particular $\mathbf{E}_\Delta(u_\Delta^n, u_\Delta^{n-1}) \le \mathbf{E}_\Delta(u_\Delta^{n-1}, u_\Delta^{n-1})$, which implies that

$$\frac{1}{2\tau}\mathbf{W}_2\left(u_\Delta^n, u_\Delta^{n-1}\right)^2 \le \mathbf{E}\left(u_\Delta^{n-1}\right) - \mathbf{E}\left(u_\Delta^n\right).$$

Evaluation of the telescopic sum yields (4.14). For proving (4.15), multiply the system (4.3) of Euler–Lagrange equations first by $(\tau/2)\mathrm{W}^{-1}$, and then by the transpose of (4.3) itself from the left, and sum over $n = 1$ to $n = N$:

$$\frac{\tau}{2}\sum_{n=1}^{N}\left[\partial_{\vec{\mathrm{x}}}\mathbb{E}_{\boldsymbol{\xi}}\left(\vec{\mathrm{x}}_\Delta^n\right)\right]^T \mathrm{W}^{-1}\left[\partial_{\vec{\mathrm{x}}}\mathbb{E}_{\boldsymbol{\xi}}\left(\vec{\mathrm{x}}_\Delta^n\right)\right] = \sum_{n=1}^{N}\left(\frac{1}{2\tau}\left(\vec{\mathrm{x}}_\Delta^n - \vec{\mathrm{x}}_\Delta^{n-1}\right)^T \mathrm{W}\left(\vec{\mathrm{x}}_\Delta^n - \vec{\mathrm{x}}_\Delta^{n-1}\right)\right).$$

Insert this in (4.14) to obtain (4.15). □

5. Qualitative properties of the discretization

Throughout this section, we fix a space-time discretization $\Delta = (\tau, \boldsymbol{\xi})$ and consider a given discrete solution $u_\Delta = (u_\Delta^n)_{n=0}^\infty$.

5.1. Metric contraction

One of the fundamental properties of our solution scheme is the preservation of the bound (1.7) on the divergence of solutions.

Proposition 5.1. *If $v_\Delta = (v_\Delta^n)_{n=0}^\infty$ is any other discrete solution, then*

$$\mathbf{W}_2\left(u_\Delta^n, v_\Delta^n\right)^2 \le (1 - 2\Lambda\tau)^{-n}\mathbf{W}_2\left(u_\Delta^0, v_\Delta^0\right)^2 \tag{5.1}$$

for all $n \in \mathbb{N}$.

Remark 5.2. Since $(1 - 2\Lambda\tau)^n < \exp(-2\Lambda n\tau)$ for every $n \in \mathbb{N}$, estimate (5.1) is slightly worse for every $\tau > 0$ than the limiting estimate (1.7).

Proof. For $\vec{\mathrm{x}}_\Delta$, $\vec{\mathrm{y}}_\Delta$ such that $u_\Delta^n = \mathbf{u}_{\boldsymbol{\xi}}[\vec{\mathrm{x}}_\Delta^n]$, $v_\Delta^n = \mathbf{u}_{\boldsymbol{\xi}}[\vec{\mathrm{y}}_\Delta^n]$ we know by Proposition 4.1 that

$$\mathrm{W}\left(\vec{\mathrm{x}}_\Delta^n - \vec{\mathrm{x}}_\Delta^{n-1}\right) = -\tau\partial_{\vec{\mathrm{x}}}\mathbb{E}_{\boldsymbol{\xi}}\left(\vec{\mathrm{x}}_\Delta^n\right) \quad \text{and} \quad \mathrm{W}\left(\vec{\mathrm{y}}_\Delta^n - \vec{\mathrm{y}}_\Delta^{n-1}\right) = -\tau\partial_{\vec{\mathrm{x}}}\mathbb{E}_{\boldsymbol{\xi}}\left(\vec{\mathrm{y}}_\Delta^n\right).$$

Substracting these equations, we obtain

$$\mathrm{W}\left(\vec{\mathrm{x}}_\Delta^n - \vec{\mathrm{y}}_\Delta^n\right) + \tau\left(\partial_{\vec{\mathrm{x}}}\mathbb{E}_{\boldsymbol{\xi}}\left(\vec{\mathrm{x}}_\Delta^n\right) - \partial_{\vec{\mathrm{x}}}\mathbb{E}_{\boldsymbol{\xi}}\left(\vec{\mathrm{y}}_\Delta^n\right)\right) = \mathrm{W}\left(\vec{\mathrm{x}}_\Delta^{n-1} - \vec{\mathrm{y}}_\Delta^{n-1}\right). \tag{5.2}$$

https://doi.org/10.1051/m2an/2013126 Published online by Cambridge University Press

Multiply (5.2) first by W^{-1} and then by the transpose of (5.2) itself from the left, and use the positivity of W^{-1} – see Lemma 3.2 – to obtain

$$\begin{aligned}(\vec{\mathrm{x}}_\Delta^n - \vec{\mathrm{y}}_\Delta^n)^T \,\mathrm{W}\,(\vec{\mathrm{x}}_\Delta^n - \vec{\mathrm{y}}_\Delta^n) + 2\tau\,(\vec{\mathrm{x}}_\Delta^n - \vec{\mathrm{y}}_\Delta^n)^T\,(\partial_{\vec{\mathrm{x}}}\mathbb{E}_{\boldsymbol{\xi}}\,(\vec{\mathrm{x}}_\Delta^n) - \partial_{\vec{\mathrm{x}}}\mathbb{E}_{\boldsymbol{\xi}}\,(\vec{\mathrm{y}}_\Delta^n))\\ \leq \left(\vec{\mathrm{x}}_\Delta^{n-1} - \vec{\mathrm{y}}_\Delta^{n-1}\right)^T \mathrm{W}\left(\vec{\mathrm{x}}_\Delta^{n-1} - \vec{\mathrm{y}}_\Delta^{n-1}\right).\end{aligned}$$

Combining this with the convexity property (3.14), we arrive at the recursive relation

$$(1-2\Lambda\tau)\,(\vec{\mathrm{x}}_\Delta^n - \vec{\mathrm{y}}_\Delta^n)^T\,\mathrm{W}\,(\vec{\mathrm{x}}_\Delta^n - \vec{\mathrm{y}}_\Delta^n) \leq \left(\vec{\mathrm{x}}_\Delta^{n-1} - \vec{\mathrm{y}}_\Delta^{n-1}\right)^T \mathrm{W}\left(\vec{\mathrm{x}}_\Delta^{n-1} - \vec{\mathrm{y}}_\Delta^{n-1}\right).$$

Iteration of this estimate and application of (3.7) yields (5.1). □

5.2. The maximum and minimum principles

Recall that ψ'' is a positive and non-increasing function by Lemma 2.3, and recall the definition of $\alpha(\boldsymbol{\xi})$ from (3.1).

Proposition 5.3 (Minimum principle)**.** *Assume that $\Lambda\tau < 1/2$, and assume further that $\boldsymbol{\xi}$ and τ are related by the* inverse CFL condition

$$\overline{\delta}(\boldsymbol{\xi})^2 \leq 6\psi''\,(Z_T^*)\,\tau, \quad \textit{where} \quad \underline{Z}_T^* := \frac{6\alpha(\boldsymbol{\xi})e^{2T\Lambda}}{\min_x u_\Delta^0}. \tag{5.3}$$

Then, for every n with $n\tau \leq T$,

$$\min_x u_\Delta^n \geq \mathrm{e}^{-2n\tau\Lambda} \min_x u_\Delta^0. \tag{5.4}$$

Proof. The minimum principle (5.4) for u_Δ is equivalent to a maximum principle for $\vec{\mathrm{z}}_\Delta$. By induction on n, we prove

$$Z^{(n)} := \max_k z_k^n \leq (1-\Lambda\tau)^{-n} \max_k z_k^0, \tag{5.5}$$

which is a slightly sharper estimate than (5.4) since $(1-\tau\Lambda) > \mathrm{e}^{-2\tau\Lambda}$ under our assumption $\tau\Lambda < 1/2$. So fix n with $n\tau \leq T$ and assume that $z_k^{n-1} \leq Z^{(n-1)}$ for all k. Suppose that $Z^{(n)} = z_m^n$ for some $m \in \{2,\dots,K-1\}$. We estimate the integral term in the Euler–Lagrange equation (4.9), using the positivity of the z_k^n and that $\Theta_m \geq 0$:

$$\begin{aligned}\left[\widetilde{\mathrm{W}}\vec{\mathrm{z}}^n\right]_m \leq \left[\widetilde{\mathrm{W}}\vec{\mathrm{z}}^{n-1}\right]_m + \tau\left((1-\gamma_m)\psi'\left(z_{m+1}^n\right) - 2\psi'\left(z_m^n\right) + (1+\gamma_m)\psi'\left(z_{m-1}^n\right)\right)\\ + \Lambda\tau \sum_{k=m-1}^{m+1} z_k^n \int_{\xi_{k-1}}^{\xi_k} \Theta_m \,\mathrm{d}\xi.\end{aligned}$$

Recalling (4.7), we conclude further that

$$\begin{aligned}(1-\Lambda\tau)\left[\widetilde{\mathrm{W}}\vec{\mathrm{z}}^n\right]_m \leq \left[\widetilde{\mathrm{W}}\vec{\mathrm{z}}^{n-1}\right]_m\\ + (1-\gamma_m)\tau\left(\psi'\left(z_{m+1}^n\right) - \psi'\left(z_m^n\right)\right) + (1+\gamma_m)\tau\left(\psi'\left(z_{m-1}^n\right) - \psi'\left(z_m^n\right)\right).\end{aligned} \tag{5.6}$$

Since $Z^{(n)} = z_m^n \geq z_k^n$ for all k, and since $\psi'' > 0$ is non-increasing, it follows that

$$\begin{aligned}\psi'\left(z_{m+1}^n\right) - \psi'\left(z_m^n\right) \leq -\psi''\left(Z^{(n)}\right)\left(z_m^n - z_{m+1}^n\right),\\ \psi'\left(z_{m-1}^n\right) - \psi'\left(z_m^n\right) \leq -\psi''\left(Z^{(n)}\right)\left(z_m^n - z_{m-1}^n\right).\end{aligned} \tag{5.7}$$

https://doi.org/10.1051/m2an/2013126 Published online by Cambridge University Press

In particular, these terms are non-positive, and so (5.6) implies

$$(1-\Lambda\tau)\widetilde{\mathrm{W}}_{m,m}Z^{(n)} \le \sigma_m Z^{(n-1)}, \quad \text{with} \quad \sigma_m = \widetilde{\mathrm{W}}_{m,m-1} + \widetilde{\mathrm{W}}_{m,m} + \widetilde{\mathrm{W}}_{m,m+1}.$$

From the explicit form of $\widetilde{\mathrm{W}}$ in (4.8), we obtain $\sigma_m \le 3\alpha(\boldsymbol{\xi})\widetilde{\mathrm{W}}_{m,m}$, and by means of the induction hypotheses, we conclude the rough bound

$$Z^{(n)} \le 6\alpha(\boldsymbol{\xi})Z^{n-1} \le 6\alpha(\boldsymbol{\xi})(1-\Lambda\tau)^{-(n-1)} \max_k z_k^0 \le Z_T^*. \tag{5.8}$$

We return to (5.6), insert (5.7), use that ψ'' is non-increasing, and find after some manipulations:

$$\begin{aligned}(1-\Lambda\tau)\sigma_m Z^{(n)} \le \sigma_m Z^{(n-1)} &+ (1-\gamma_m)\left(\frac{1-\Lambda\tau}{6}\delta_{m+1}^2 - \tau\psi''\left(Z_T^*\right)\right)\left(z_m^n - z_{m-1}^n\right)\\ &+ (1+\gamma_m)\left(\frac{1-\Lambda\tau}{6}\delta_{m-1}^2 - \tau\psi''\left(Z_T^*\right)\right)\left(z_m^n - z_{m+1}^n\right).\end{aligned}$$

The inverse CFL condition (5.3) implies non-positivity of the last two terms, so (5.8) refines to

$$Z^{(n)} \le (1-\Lambda\tau)^{-1}Z^{(n-1)} \le (1-\Lambda\tau)^{-n}\max_k z_k^0.$$

If the maximum is attained at one of the boundary points, $m=1$ or $m=K$, then a similar calculation can be carried out using (4.10) or (4.11), respectively, instead of (4.9). □

Similarly to the minimum principle, one obtains the following maximum principle. Notice that the inverse CFL condition is the same as before, *i.e.*, it involves $\min_x u_\Delta^0$ and not $\max_x u_\Delta^0$.

Proposition 5.4 (maximum principle)**.** *Let all the hypotheses of Proposition 5.3 hold, and assume – in addition to* (5.3) *– that*

$$(1+\lambda\tau)\overline{\delta}(\boldsymbol{\xi})^2 \le 6\tau\psi''(Z_T^*), \quad \textit{where} \quad \lambda := \max_{x\in I} V_{xx} \ge 0.$$

Then, for every n with $n\tau \le T$,

$$\max_x u_\Delta^n \le \mathrm{e}^{2n\tau\lambda}\max_x u_\Delta^0.$$

We omit the proof, which is very similar to the one given above.

5.3. Regularity

Below, we derive a bound on the time-integrated *total variation* of the considered discrete solution u_Δ which is independent on the discretization Δ. This bound provides the key a priori estimate for the convergence proof in the next section. The use of total variation (instead of *e.g.* Sobolev norms) is natural in our context, since it can be directly evaluated on the piecewise constant profiles u_Δ^n.

Recall that the total variation of a function $f \in L^1(I)$ is given by

$$\{f\}_{\mathrm{TV}} := \sup\left\{\int_a^b f(x)\varphi'(x)\,\mathrm{d}x \,\middle|\, \varphi\in\mathrm{Lip}(I),\ \varphi(a)=\varphi(b)=0,\ \sup_{x\in I}|\varphi(x)|\le 1\right\}. \tag{5.9}$$

If f is a piecewise constant function, taking values f_k on intervals $(x_{k-1}, x_k]$, with our usual convention $a = x_0 < x_1 < \cdots < x_K = b$, then the integral in (5.9) amounts to

$$\int_a^b f(x)\varphi'(x)\,\mathrm{d}x = \sum_{k=1}^K \left[f(x)\varphi(x)\right]_{x=x_{k-1}+0}^{x_k-0} = \sum_{k=1}^{K-1}(f_k - f_{k+1})\varphi(x_k),$$

https://doi.org/10.1051/m2an/2013126 Published online by Cambridge University Press

where we have used that $\varphi(a) = \varphi(b) = 0$. Consequently, for such f, the supremum in (5.9) equals

$$\{f\}_{\mathrm{TV}} = \sum_{k=1}^{K-1} |f_{k+1} - f_k| \tag{5.10}$$

and is attained for every $\varphi \in \mathrm{Lip}(I)$ with $\varphi(x_k) = \mathrm{sgn}(f_k - f_{k+1})$ at $k = 1, \dots, K-1$.

Proposition 5.5. *For every $N \in \mathbb{N}$,*

$$\tau \sum_{n=1}^{N} \{\mathrm{P}\,(u_\Delta^n)\}_{\mathrm{TV}}^2 \le \mathcal{C}\left(\mathbf{E}_{\boldsymbol{\xi}}\left(u_\Delta^0\right), N\tau\right), \tag{5.11}$$

where

$$\mathcal{C}(E,T) = 2M\left[2(E - \underline{\mathbf{E}}) + MT \sup_{x\in I} |V_x(x)|^2\right]. \tag{5.12}$$

Proof. Fix some time index n. Let $\mathrm{X}^n = \mathbf{X}_{\boldsymbol{\xi}}[\vec{x}_\Delta^n]$ be the inverse distribution function of u_Δ^n. Further, let $\vec{\mathrm{y}} \in \mathbb{R}^{K-1}$ with components $y_k \in [-1,1]$ be given, and introduce $\mathrm{Y} : [0,M] \to \mathbb{R}$ by

$$\mathrm{Y} = \sum_{k=1}^{K-1} y_k \theta_k.$$

Note that Y does not belong to $\mathfrak{X}_{\boldsymbol{\xi}}$, but $\mathrm{Y}(0) = \mathrm{Y}(M) = 0$ instead. With $y_0 = y_K = 0$,

$$\begin{aligned}
\sum_{k=1}^{K-1} \left[\partial_{\vec{\mathrm{x}}} \mathbb{E}_{\boldsymbol{\xi}}\left(\vec{x}_\Delta^n\right)\right]_k y_k &= \sum_{k=1}^{K-1} \left(-\psi\left(\frac{x_{k+1}^n - x_k^n}{\delta_{k+1}}\right) + \psi\left(\frac{x_k^n - x_{k-1}^n}{\delta_k}\right) + \int_0^M V_x\left(\mathrm{X}^n\right) \theta_k \,\mathrm{d}\xi \right) y_k \\
&= \sum_{k=1}^{K} \delta_k \psi\left(\frac{x_k^n - x_{k-1}^n}{\delta_k}\right) \frac{y_k - y_{k-1}}{\delta_k} + \int_0^M V_x\left(\mathrm{X}^n\right) \sum_{k=1}^{K-1} y_k \theta_k \,\mathrm{d}\xi \\
&= \int_0^M \psi'\left(\mathrm{X}_\xi^n\right) \mathrm{Y}_\xi \,\mathrm{d}\xi + \int_0^M V_x\left(\mathrm{X}^n\right) \mathrm{Y} \,\mathrm{d}\xi \\
&= \int_a^b \psi'\left(\mathrm{X}_\xi^n \circ \left(\mathrm{X}^n\right)^{-1}\right) \frac{\mathrm{Y}_\xi \circ \left(\mathrm{X}^n\right)^{-1}}{\mathrm{X}_\xi^n \circ \left(\mathrm{X}^n\right)^{-1}} \,\mathrm{d}x + \int_0^M V_x\left(\mathrm{X}^n\right) \mathrm{Y} \,\mathrm{d}\xi.
\end{aligned} \tag{5.13}$$

On one hand, we have by definition of ψ and by (2.1) that

$$\psi\left(\mathrm{X}_\xi^n \circ \left(\mathrm{X}^n\right)^{-1}\right) = -\mathrm{P}\left(\frac{1}{\mathrm{X}_\xi^n \circ \left(\mathrm{X}^n\right)^{-1}}\right) = -\mathrm{P}\left(u_\Delta^n\right) \quad \text{a.e. on } I. \tag{5.14}$$

On the other hand, $\varphi := \mathrm{Y} \circ \left(\mathrm{X}^n\right)^{-1}$ is a (piecewise affine) Lipschitz function with

$$\varphi_x = \frac{\mathrm{Y}_\xi \circ \left(\mathrm{X}^n\right)^{-1}}{\mathrm{X}_\xi^n \circ \left(\mathrm{X}^n\right)^{-1}}. \tag{5.15}$$

Moreover, by the Cauchy–Schwarz inequality,

$$-\sum_{k=1}^{K-1} \left[\partial_{\vec{\mathrm{x}}} \mathbb{E}_{\boldsymbol{\xi}}\left(\vec{x}_\Delta^n\right)\right]_k y_k \le \left(\vec{\mathrm{y}}^T \mathrm{W} \vec{\mathrm{y}}\right)^{1/2} \left(\left[\partial_{\vec{\mathrm{x}}} \mathbb{E}_{\boldsymbol{\xi}}\left(\vec{x}_\Delta^n\right)\right]^T \mathrm{W}^{-1} \left[\partial_{\vec{\mathrm{x}}} \mathbb{E}_{\boldsymbol{\xi}}\left(\vec{x}_\Delta^n\right)\right]\right)^{1/2}, \tag{5.16}$$

https://doi.org/10.1051/m2an/2013126 Published online by Cambridge University Press

and $\vec{y}^T \mathrm{W}\vec{y} \le M$ by (3.8), and since $|y_k| \le 1$. Using (5.14), (5.15), (5.16) and $|\mathrm{Y}| \le 1$ in (5.13), we eventually arrive at

$$\int_a^b \mathrm{P}\left(u_\Delta^n\right) \varphi_x \, \mathrm{d}x \le \left(M \left[\partial_{\vec{\mathrm{x}}} \mathbb{E}_{\boldsymbol{\xi}} \left(\vec{\mathrm{x}}_\Delta^n\right)\right]^T \mathrm{W}^{-1} \left[\partial_{\vec{\mathrm{x}}} \mathbb{E}_{\boldsymbol{\xi}} \left(\vec{\mathrm{x}}_\Delta^n\right)\right] \right)^{1/2} + M \sup_{x \in I} |V_x(x)|. \tag{5.17}$$

The function $\mathrm{P}(u_\Delta^n)$ is piecewise constant on the intervals $(x_{k-1}^n, x_k^n]$. And φ is Lipschitz with

$$\sup_{x \in I} |\varphi(x)| = \sup_{\xi \in [0,M]} |\mathrm{Y}(\xi)| = \max_k |y_k| \le 1, \quad \varphi(a) = \mathrm{Y}(0) = 0 = \mathrm{Y}(M) = \varphi(b).$$

Moreover, the values $\varphi(x_k^n) = y_k$ can be chosen arbitrarily in $[-1, 1]$ for $k = 1, \dots, K-1$. Thus – recalling our discussion preceeding Proposition 5.5 – the right-hand side of inequality (5.17) is a bound on $\{P(u_\Delta^n)\}_{\mathrm{TV}}$. To obtain (5.11), take the square on both sides, sum the resulting inequalities from $n = 1$ to $n = N$, and apply the energy estimate (4.15). □

6. Convergence

In this section, we prove Theorem 1.1. Let a time horizon $T > 0$ and an initial condition $u^0 \in L^1(I)$ with $\mathbf{E}(u^0) < \infty$ be given. We consider a family $\Delta_j = (\tau_j, \boldsymbol{\xi}_j)$ of time-space discretizations, with $j \in \mathbb{N}$. Accordingly, we denote by K_j the number of nodes of $\boldsymbol{\xi}_j$, and N_j is the smallest integer with $\tau_j N_j \ge T$. Throughout this section, we assume all the hypotheses of Theorem 1.1:

- $\tau_j \downarrow 0$ and $\overline{\delta}(\boldsymbol{\xi}_j) \downarrow 0$ as $j \to \infty$;
- the initial conditions $u_{\Delta_j}^0 \in \mathrm{D}_{\boldsymbol{\xi}_j}^M(I)$ converge to u^0 weakly in $L^1(I)$;
- uniformly in $j \in \mathbb{N}$,

$$\alpha(\boldsymbol{\xi}_j) \le \overline{\alpha} < \infty, \quad \mathbf{E}\left(u_{\Delta_j}^0\right) \le \overline{\mathbf{E}} < \infty, \quad \overline{\delta}(\boldsymbol{\xi}_j)^2 \le 6\psi''\left(\frac{6\overline{\alpha}\mathrm{e}^{2\Lambda T}}{\min_x u_{\Delta_j}^0}\right)\tau_j. \tag{6.1}$$

Denote by $(u_{\Delta_j}^n)_{n=0}^N$ the corresponding discrete solutions obtained as in Proposition 4.1, and introduce the time-interpolated functions $\bar{u}_{\Delta_j} : [0,T] \to \mathrm{D}_{\boldsymbol{\xi}_j}^M(I)$ by

$$\bar{u}_{\Delta_j}(t;x) = u_{\Delta_j}^n(x) \quad \text{for all } t \in ((n-1)\tau_j, n\tau_j] \cap [0,T]. \tag{6.2}$$

The following preliminary result plays an important role in the convergence proof.

Lemma 6.1. *With the maximal mesh width $\overline{\delta}(\vec{\mathrm{x}})$ defined in (3.4), we have*

$$\max_{n \le N_j} \overline{\delta}\left(\vec{\mathrm{x}}_{\Delta_j}^n\right) \to 0 \quad \text{as } j \to \infty. \tag{6.3}$$

Proof. This is a consequence of the minimum principle and hypothesis (6.1). It follows from the assumption $\lim_{r \downarrow 0} \mathrm{P}'(r) < \infty$ in (1.2) that

$$\psi''\left(\frac{\mathrm{e}^{2\Lambda T}}{\min_x u_{\Delta_j}^0}\right) \le C\left(\frac{\min_x u_{\Delta_j}^0}{\mathrm{e}^{2\Lambda T}}\right)^2$$

for some appropriate constant C. Recalling that $(x_k - x_{k-1})u_k = \delta_k$, Proposition 5.3 thus implies – uniformly in j and $n \le N_j$ – that

$$\overline{\delta}\left(\vec{\mathrm{x}}_{\Delta_j}^n\right)^2 \le \left(\frac{\overline{\delta}(\boldsymbol{\xi}_j)}{\min_x u_{\Delta_j}^n}\right)^2 \le \overline{\delta}(\boldsymbol{\xi}_j)^2 \left(\frac{\mathrm{e}^{2\Lambda T}}{\min_x u_{\Delta_j}^0}\right)^2 \le \frac{C\overline{\delta}(\boldsymbol{\xi}_j)^2}{\psi''\left(\frac{\mathrm{e}^{2\Lambda T}}{\min_x u_{\Delta_j}^0}\right)} \le 6C\tau_j,$$

using hypothesis (6.1). Since $\tau_j \downarrow 0$ as $j \to \infty$, this proves the claim. □

https://doi.org/10.1051/m2an/2013126 Published online by Cambridge University Press

6.1. Compactness

From now on, denote by $\mathcal{P}(I)$ the metric space of non-negative measures on I with total mass M, endowed with the Wasserstein distance $\mathbf{W}_2$. The space $\mathrm{D}^M(I)$ of densities naturally embeds into $\mathcal{P}(I)$.

The following weak convergence result is a well-known consequence of the energy estimate (4.14).

Proposition 6.2. *Every subsequence of $(\bar{u}_{\Delta_j})_{j\in\mathbb{N}}$ contains a sub-subsequence that converges uniformly w.r.t. $t \in [0,T]$ in $\mathbf{W}_2$ to a limit curve $u_* \in C^{1/2}([0,T];\mathcal{P}(I))$.*

For a proof, see [1], Proposition 3.3.1. In essence, one shows that (4.14) yields a j-uniform control on the modulus of Hölder-1/2 continuity of the $\bar{u}_{\Delta_j}$, which allows to invoke the Arzelà–Ascoli theorem.

A stronger compactness result is needed for the convergence proof.

Proposition 6.3. *Every subsequence of $(\bar{u}_{\Delta_j})_{j\in\mathbb{N}}$ contains a sub-subsequence such that the respective $\bar{u}_{\Delta_j}$ converge to some u_*, and the $\mathrm{P}(\bar{u}_{\Delta_j})$ converge to $\mathrm{P}(u_*)$, both strongly in $L^1([0,T]\times I)$.*

The proof of this proposition is an application of the Aubin-Lions compactness principle. Specifically, we use:

Theorem 6.4 (Adapted from Theorem 2 in [20])**.** *Assume that:*

(1) There is a normal coercive integrand $\mathfrak{F} : L^1(I) \to [0,\infty]$, i.e., *$\mathfrak{F}$ is measurable, lower semi-continuous and has compact sublevels in $L^1(I)$, for which the following is true:*

$$\sup_{j\in\mathbb{N}} \int_0^T \mathfrak{F}\left(\bar{u}_{\Delta_j}(t)\right)\,\mathrm{d}t < \infty. \tag{6.4}$$

(2) The $\bar{u}_{\Delta_j}$ are integral equicontinuous with respect to $\mathbf{W}_2$,

$$\limsup_{h\downarrow 0} \sup_{j\in\mathbb{N}} \int_0^{T-h} \mathbf{W}_2\left(\bar{u}_{\Delta_j}(t+h), \bar{u}_{\Delta_j}(t)\right)\,\mathrm{d}t = 0. \tag{6.5}$$

Then the sequence $(\bar{u}_{\Delta_j})_{j\in\mathbb{N}}$ is relatively compact in $L^1([0,T]\times I)$.

Let $\mathfrak{F} : L^1(I) \to \mathbb{R}\cup\{+\infty\}$ be given by $\mathfrak{F}(u) = \{\mathrm{P}(u)\}^2_{\mathrm{TV}}$ – possibly $+\infty$ – if u is non-negative with $\int_I u(x)\,\mathrm{d}x = M$, and by $+\infty$ otherwise.

Lemma 6.5. *$\mathfrak{F}$ is lower semi-continuous and has relatively compact sublevels.*

Proof of Lemma 6.5. Let $A_c := \mathfrak{F}^{-1}((-\infty;c]) \subset L^1(I)$ be a sublevel of $\mathfrak{F}$. By [11], Theorem 1.19, the set $B_c := \{\mathrm{P}(u)\,|\,u \in A_c\}$ is relatively compact in $L^1(I)$; here we use that our domain I is an interval, so that $\{\mathrm{P}(u)\}^2_{\mathrm{TV}} \le c$ and $\int_I u(x)\,\mathrm{d}x = M$ induce a uniform bound on the BV-norm of $\mathrm{P}(u)$. Thus, if (u_ℓ) is a sequence in A_c, converging to u_0 in $L^1(I)$, then also $(\mathrm{P}(u_\ell))$ converges to $\mathrm{P}(u_0)$ in $L^1(I)$. By lower semi-continuity of the total variation $\{\cdot\}_{\mathrm{TV}}$ ([11], Thm. 1.9), the lower semi-continuity of $\mathfrak{F}$ follows.

To conclude compactness of A_c, it suffices to prove that the mapping $u \mapsto \mathrm{P}(u)$ is $L^1(I)$-continuously invertible. For that, let a sequence $(f_\ell)_{\ell\in\mathbb{N}}$ in B_c be given, which converges to some f_0 in $L^1(I)$. Since the map $r \mapsto \mathrm{P}(r)$ is strictly increasing, positive, and continuous with superlinear growth it possesses a strictly increasing, positive and continuous inverse with sublinear growth. Hence, there are a uniquely determined sequence of functions $u_\ell \in A_c$ such that $\mathrm{P}(u_\ell) = f_\ell$ for all $\ell \in \mathbb{N}$, and a unique $u_0 \in A_c$ with $\mathrm{P}(u_0) = f_0$. We wish to show that u_ℓ converges to u_0 in $L^1(I)$. By standard arguments, we can assume without loss of generality that the f_ℓ converge to f_0 pointwise a.e. By continuous invertibility of $r \mapsto \mathrm{P}(r)$, the u_j converge to u_0 pointwise a.e. Moreover, by construction,

$$\sup_{\ell\in\mathbb{N}} \int_a^b \mathrm{P}(u_\ell(x))\,\mathrm{d}x = \sup_{\ell\in\mathbb{N}} \int_a^b f_\ell(x)\,\mathrm{d}x < \infty,$$

so we can invoke Vitali's theorem – recall the superlinear growth of P – to conclude strong convergence of u_ℓ to u_0. □

https://doi.org/10.1051/m2an/2013126 Published online by Cambridge University Press

Proof of Proposition 6.3. It suffices to show that every subsequence of $(\bar{u}_{\Delta_j})_{j\in\mathbb{N}}$ contains a sub-subsequence which is relatively compact. In view of Proposition 6.2, we may thus assume – without loss of generality – that $(\bar{u}_{\Delta_j})_{j\in\mathbb{N}}$ converges uniformly w.r.t. $t\in[0,T]$ in $\mathcal{P}(I)$ to a curve $u_*\in C^{1/2}([0,T];\mathcal{P}(I))$. The verification of (6.5) then becomes an easy exercise, which is left to the reader. And estimate (6.4) is a direct consequence of the regularity estimate (5.11).

Thus Theorem 6.4 applies and provides relative compactness of $(\bar{u}_{\Delta_j})_{j\in\mathbb{N}}$ in $L^1([0,T]\times I)$. Since L^1-convergence implies weak convergence, it actually follows that $\bar{u}_{\Delta_j}$ converges to u_* in $L^1([0,T]\times I)$. Without loss of generality, we may even assume that $\bar{u}_{\Delta_j}$ converges to u_* a.e. on $[0,T]\times I$. By continuity of P, also $\mathrm{P}(\bar{u}_{\Delta_j})$ converges to $\mathrm{P}(u_*)$ a.e. on $[0,T]\times I$. Further, (5.10) implies

$$\int_0^T\int_I \mathrm{P}\left(\bar{u}_{\Delta_j}(t;x)\right)^2\,\mathrm{d}x\,\mathrm{d}t \le 2(b-a)\tau\sum_{n=1}^{N_j}\left[\mathrm{P}\left(\frac{M}{b-a}\right)^2+\left\{\mathrm{P}\left(\bar{u}^n_{\Delta_j}\right)\right\}^2_{\mathrm{TV}}\right],$$

which is j-uniformly bounded because of the regularity estimate (5.11). By the growth property of P, we can invoke Vitali's theorem to conclude that $\mathrm{P}(\bar{u}_{\Delta_j})$ tends to $\mathrm{P}(u_*)$ in $L^1([0,T]\times I)$. □

6.2. Weak formulation

Combining the compactness results from Propositions 6.2 and 6.3, we know that every subsequence of $(\Delta_j)_{j\in\mathbb{N}}$ contains a sub-subsequence for which $\bar{u}_{\Delta_j}$ converges to some limit

$$u_*\in C^{1/2}([0,T];\mathcal{P}(I))\cap L^1([0,T]\times I),$$

uniformly w.r.t. $t\in[0,T]$ in $\mathcal{P}(I)$, and strongly in $L^1([0,T]\times I)$; finally, also $\mathrm{P}(\bar{u}_{\Delta_j})$ converges to $\mathrm{P}(u_*)$ in $L^1([0,T]\times I)$. To simplify notations, we denote that sub-subsequence simply by $(\bar{u}_\Delta)$, bearing in mind that $\Delta=\Delta_j$. In this subsection, we prove that every such limit u_* is a weak solution to the initial value problem (1.1).

Proposition 6.6. *u_* satisfies the weak formulation*

$$\int_0^T\int_I u_*\partial_t\varphi\,\mathrm{d}x\,\mathrm{d}t=\int_0^T\int_I u_*V_x\partial_x\varphi\,\mathrm{d}x\,\mathrm{d}t-\int_0^T\int_I \mathrm{P}(u_*)\partial_{xx}\varphi\,\mathrm{d}x\,\mathrm{d}t \tag{6.6}$$

for all test functions φ from

$$\mathcal{D}:=\left\{\varphi\in C^\infty([0,T]\times I)\,\middle|\,\operatorname{supp}\varphi\subset(0,T)\times I,\ \varphi_x(t;a)=\varphi_x(t;b)=0\ \textit{f.a. } t\in[0,T]\right\}. \tag{6.7}$$

Moreover, u_* attains the initial datum u^0 weakly-⋆ as $t\downarrow 0$.

Remark 6.7. Since weak solutions to (1.1) are unique (this follows, *e.g.*, by metric contraction of the gradient flow), we conclude *a posteriori* that the entire sequence $(u_{\Delta_j})_{j\in\mathbb{N}}$ converges to the solution u_*.

Remark 6.8. By our definition (6.7) of test functions, the weak formulation (6.6) automatically induces homogeneous Neumann boundary conditions on u.

The weak formulation (6.6) is obtained in the limit $j\to\infty$ from a certain fully discrete variant of the weak formulation. The latter is derived by studying suitable variations of the minimizers $u^n_{\Delta_j}$. In order to discuss this perturbation, let $\rho\in C^\infty(I)$ with $\rho_x(a)=\rho_x(b)=0$ be given, and let $\kappa>0$ be such that

$$|\rho_x(x)|\le\kappa,\ |\rho_{xx}(x)|\le\kappa,\ |\rho_{xxx}(x)|\le\kappa\quad\text{for all } x\in I. \tag{6.8}$$

Further, fix $j\in\mathbb{N}$ and also $n\in\mathbb{N}$. We shall omit the index j in the following.

https://doi.org/10.1051/m2an/2013126 Published online by Cambridge University Press

Introduce the functionals $\mathbf{A} : \mathfrak{X} \to \mathbb{R}$ and $\mathbf{A}_{\boldsymbol{\xi}} : \mathfrak{x}_{\boldsymbol{\xi}} \to \mathbb{R}$ by

$$\mathbf{A}(\mathrm{X}) = \int_0^M \rho(\mathrm{X})\,\mathrm{d}\xi \quad \text{and} \quad \mathbf{A}_{\boldsymbol{\xi}}(\vec{\mathrm{x}}) = \mathbf{A}(\mathbf{X}_{\boldsymbol{\xi}}[\vec{\mathrm{x}}]).$$

The derivative of $\mathbf{A}_{\boldsymbol{\xi}}$ is given by

$$\big[\partial_{\vec{\mathrm{x}}}\mathbf{A}_{\boldsymbol{\xi}}(\vec{\mathrm{x}})\big]_k = \int_0^M \rho_x\big(\mathbf{X}_{\boldsymbol{\xi}}[\vec{\mathrm{x}}]\big)\theta_k\,\mathrm{d}\xi \quad \text{for } k = 1, \dots, K-1.$$

The variations we study are those induced by the *gradient vector field* $\vec{v} := \mathrm{W}^{-1}\partial_{\vec{\mathrm{x}}}\mathbf{A}_{\boldsymbol{\xi}}$, *i.e.*,

$$\mathrm{W}\vec{v}(\vec{\mathrm{x}}) = \partial_{\vec{\mathrm{x}}}\mathbf{A}_{\boldsymbol{\xi}}(\vec{\mathrm{x}}). \tag{6.9}$$

This vector field has a nice asymptotic expansion.

Lemma 6.9. *For every $\vec{\mathrm{x}} \in \mathfrak{x}_{\boldsymbol{\xi}}$, define $\vec{\rho_x}(\vec{\mathrm{x}}) \in \mathbb{R}^{K-1}$ by*

$$\vec{\rho_x}(\vec{\mathrm{x}}) = \big(\rho_x(x_1), \rho_x(x_2), \dots, \rho_x(x_{K-1})\big). \tag{6.10}$$

Then the residual vector $\nu = \vec{v}(\vec{\mathrm{x}}) - \vec{\rho_x}(\vec{\mathrm{x}})$ satisfies

$$\nu^T \mathrm{W}\nu \le C\kappa^2\alpha(\boldsymbol{\xi})M\overline{\delta}(\vec{\mathrm{x}})^2, \tag{6.11}$$

with κ defined in (6.8) and some universal constant C.

Proof. We start by estimating

$$\mu := \mathrm{W}\nu = \mathrm{W}\big(\vec{v}(\vec{\mathrm{x}}) - \vec{\rho_x}[\vec{\mathrm{x}}]\big) = \partial_{\vec{\mathrm{x}}}\mathbf{A}_{\boldsymbol{\xi}}(\vec{\mathrm{x}}) - \mathrm{W}\vec{\rho_x}[\vec{\mathrm{x}}].$$

By definition of $\mathbf{X}_{\boldsymbol{\xi}}[\vec{\mathrm{x}}]$ and a Taylor expansion of ρ_x, we have

$$\rho_x\left(\mathbf{X}_{\boldsymbol{\xi}}[\vec{\mathrm{x}}](\xi)\right) = \rho_x(x_m) + \rho_{xx}(\hat{x})(x_m - x_{m-1})\theta_{m-1}(\xi) \tag{6.12}$$

$$= \rho_x(x_{m-1}) + \rho_{xx}(\check{x})(x_m - x_{m-1})\theta_m(\xi) \tag{6.13}$$

for all $\xi \in [\xi_{m-1}, \xi_m]$, where $\hat{x}, \check{x} \in I$ denote suitable intermediate values, depending on ξ. Multiply (6.12) by θ_m, multiply (6.13) by θ_{m-1}, and sum these up to obtain, recalling (6.8), that

$$\rho_x\left(\mathbf{X}_{\boldsymbol{\xi}}[\vec{\mathrm{x}}](\xi)\right) = \rho_x(x_m)\theta_m(\xi) + \rho_x(x_{m-1})\theta_{m-1}(\xi) + R_m(\xi), \quad \text{with} \quad |R_m(\xi)| \le \kappa\overline{\delta}(\vec{\mathrm{x}}). \tag{6.14}$$

Since any component of W is given by $\mathrm{W}_{k,j} = \int_0^M \theta_k\theta_j\,\mathrm{d}\xi$, we calculate

$$\mu_k = \int_{\xi_{k-1}}^{\xi_{k+1}} \left(\rho_x\left(\mathbf{X}_{\boldsymbol{\xi}}[\vec{\mathrm{x}}]\right) - \sum_{j=k-1}^{k+1} \rho_x(x_j)\theta_j\right)\theta_k\,\mathrm{d}\xi.$$

Substitute (6.14) into the integral to obtain

$$|\mu_k| \le C_1\kappa(\delta_k + \delta_{k+1})\overline{\delta}(\vec{\mathrm{x}})$$

for every $k = 1, \dots, K-1$, with some universal constant C_1. Recalling the lower estimate on W in (3.9), it follows for $\nu = \mathrm{W}^{-1}\mu$ that

$$\nu^T\mathrm{W}\nu = \mu^T\mathrm{W}^{-1}\mu \le \frac{6}{\underline{\delta}(\boldsymbol{\xi})}\sum_{k=1}^{K-1}\mu_k^2 \le \left(\frac{6C_1^2\kappa^2}{\underline{\delta}(\boldsymbol{\xi})}\sum_{k=1}^{K-1}(\delta_{k+1} + \delta_k)^2\right)\overline{\delta}(\vec{\mathrm{x}})^2 \le 24C_1^2\kappa^2\alpha(\boldsymbol{\xi})M\overline{\delta}(\vec{\mathrm{x}})^2,$$

proving our claim (6.11). □

https://doi.org/10.1051/m2an/2013126 Published online by Cambridge University Press

Lemma 6.10. *With κ satisfying (6.8), the following estimate holds:*

$$\frac{1}{\tau}\left(\mathbf{A}_{\boldsymbol{\xi}}\left(\vec{\mathrm{x}}_\Delta^n\right)-\mathbf{A}_{\boldsymbol{\xi}}\left(\vec{\mathrm{x}}_\Delta^{n-1}\right)\right)\le -\partial_{\vec{\mathrm{x}}}\mathbb{E}_{\boldsymbol{\xi}}\left(\vec{\mathrm{x}}_\Delta^n\right)^T\vec{\rho_x}\left(\vec{\mathrm{x}}_\Delta^n\right) \tag{6.15}$$

$$+\frac{\kappa}{2\tau}\left(\vec{\mathrm{x}}_\Delta^n-\vec{\mathrm{x}}_\Delta^{n-1}\right)^T\mathrm{W}\left(\vec{\mathrm{x}}_\Delta^n-\vec{\mathrm{x}}_\Delta^{n-1}\right) \tag{6.16}$$

$$+C\left(M\alpha(\boldsymbol{\xi})\right)^{1/2}\kappa\left(\partial_{\vec{\mathrm{x}}}\mathbb{E}_{\boldsymbol{\xi}}\left(\vec{\mathrm{x}}_\Delta^n\right)^T\mathrm{W}^{-1}\partial_{\vec{\mathrm{x}}}\mathbb{E}_{\boldsymbol{\xi}}\left(\vec{\mathrm{x}}_\Delta^n\right)\right)^{1/2}\overline{\delta}\left(\vec{\mathrm{x}}_\Delta^n\right). \tag{6.17}$$

Proof. A Taylor expansion of ρ yields for arbitrary $\mathrm{X},\mathrm{X}'\in\mathfrak{X}$:

$$\mathbf{A}(\mathrm{X}')\ge\mathbf{A}(\mathrm{X})+\int_0^M\rho_x(\mathrm{X})\cdot(\mathrm{X}'-\mathrm{X})\,\mathrm{d}\xi-\frac{\kappa}{2}\int_0^M(\mathrm{X}'-\mathrm{X})^2\,\mathrm{d}\xi.$$

With $\mathrm{X}=\mathbf{X}_{\boldsymbol{\xi}}[\vec{\mathrm{x}}_\Delta^n]$ and $\mathrm{X}'=\mathbf{X}_{\boldsymbol{\xi}}[\vec{\mathrm{x}}_\Delta^{n-1}]$, we obtain

$$\frac{1}{\tau}\left(\mathbf{A}_{\boldsymbol{\xi}}\left(\vec{\mathrm{x}}_\Delta^n\right)-\mathbf{A}_{\boldsymbol{\xi}}\left(\vec{\mathrm{x}}_\Delta^{n-1}\right)\right)\le-\sum_{m=1}^{K-1}\left(\frac{1}{\tau}\left[\vec{\mathrm{x}}_\Delta^n-\vec{\mathrm{x}}_\Delta^{n-1}\right]_m\int_0^M\rho_x\left(\mathbf{X}_{\boldsymbol{\xi}}\left[\vec{\mathrm{x}}_\Delta^n\right]\right)\theta_m\,\mathrm{d}\xi\right)$$
$$+\frac{\kappa}{2\tau}\left(\vec{\mathrm{x}}_\Delta^n-\vec{\mathrm{x}}_\Delta^{n-1}\right)^T\mathrm{W}\left(\vec{\mathrm{x}}_\Delta^n-\vec{\mathrm{x}}_\Delta^{n-1}\right).$$

Using the definition (6.9) of $\vec{v}$, the symmetry $\mathrm{W}^T=\mathrm{W}$, and the Euler–Lagrange equations (4.3), the sum can be rewritten as

$$\sum_{m=1}^{K-1}\frac{1}{\tau}\left[\vec{\mathrm{x}}_\Delta^n-\vec{\mathrm{x}}_\Delta^{n-1}\right]_m\left[\partial_{\vec{\mathrm{x}}}\mathbf{A}_{\boldsymbol{\xi}}\left(\vec{\mathrm{x}}_\Delta^n\right)\right]_m=\frac{1}{\tau}\left(\vec{\mathrm{x}}_\Delta^n-\vec{\mathrm{x}}_\Delta^{n-1}\right)^T\mathrm{W}\vec{v}\left(\vec{\mathrm{x}}_\Delta^n\right)=\partial_{\vec{\mathrm{x}}}\mathbb{E}_{\boldsymbol{\xi}}\left(\vec{\mathrm{x}}_\Delta^n\right)^T\vec{v}\left(\vec{\mathrm{x}}_\Delta^n\right).$$

With the notations introduced in Lemma 6.9, we obtain further

$$\partial_{\vec{\mathrm{x}}}\mathbb{E}_{\boldsymbol{\xi}}\left(\vec{\mathrm{x}}_\Delta^n\right)^T\vec{v}\left(\vec{\mathrm{x}}_\Delta^n\right)=\partial_{\vec{\mathrm{x}}}\mathbb{E}_{\boldsymbol{\xi}}\left(\vec{\mathrm{x}}_\Delta^n\right)^T\vec{\rho_x}\left(\vec{\mathrm{x}}_\Delta^n\right)+\partial_{\vec{\mathrm{x}}}\mathbb{E}_{\boldsymbol{\xi}}\left(\vec{\mathrm{x}}_\Delta^n\right)^T\nu_\Delta^n$$
$$\le\partial_{\vec{\mathrm{x}}}\mathbb{E}_{\boldsymbol{\xi}}\left(\vec{\mathrm{x}}_\Delta^n\right)^T\vec{\rho_x}\left(\vec{\mathrm{x}}_\Delta^n\right)+\left(\partial_{\vec{\mathrm{x}}}\mathbb{E}_{\boldsymbol{\xi}}\left(\vec{\mathrm{x}}_\Delta^n\right)^T\mathrm{W}^{-1}\partial_{\vec{\mathrm{x}}}\mathbb{E}_{\boldsymbol{\xi}}\left(\vec{\mathrm{x}}_\Delta^n\right)\right)^{1/2}\left(\left(\nu_\Delta^n\right)^T\mathrm{W}\nu_\Delta^n\right)^{1/2},$$

where the Cauchy–Schwarz inequality has been applied in the last step. The claim (6.15) now follows directly from the estimate (6.11). □

Lemma 6.11. *For every $\vec{\mathrm{x}}\in\mathfrak{x}_{\boldsymbol{\xi}}$,*

$$\mathbf{A}_{\boldsymbol{\xi}}(\vec{\mathrm{x}})=\int_I\rho(x)\mathbf{u}_{\boldsymbol{\xi}}[\vec{\mathrm{x}}](x)\,\mathrm{d}x, \tag{6.18}$$

$$\partial_{\vec{\mathrm{x}}}\mathbb{E}_{\boldsymbol{\xi}}(\vec{\mathrm{x}})^T\vec{\rho_x}(\vec{\mathrm{x}})=-\int_I\mathrm{P}\left(\mathbf{u}_{\boldsymbol{\xi}}[\vec{\mathrm{x}}](x)\right)\rho_{xx}(\hat{x})\,\mathrm{d}x+\int_I V_x(x)\rho_x(\check{x})\mathbf{u}_{\boldsymbol{\xi}}[\vec{\mathrm{x}}](x)\,\mathrm{d}x, \tag{6.19}$$

where $\hat{x},\check{x}\in I$ are x-dependent quantities satisfying

$$|\hat{x}-x|,\,|\check{x}-x|<\overline{\delta}(\vec{\mathrm{x}}).$$

Proof. Relation (6.18) is a direct consequence of

$$\int_{\xi_{k-1}}^{\xi_k}\rho\big(x_k\theta_k+x_{k-1}\theta_{k-1}\big)\,\mathrm{d}\xi=\int_{x_{k-1}}^{x_k}\rho(x)\frac{\delta_k}{x_k-x_{k-1}}\,\mathrm{d}x,$$

https://doi.org/10.1051/m2an/2013126 Published online by Cambridge University Press

which follows by a change of variables $x = x_k\theta_k(\xi) + x_{k-1}\theta_{k-1}(\xi)$. To prove (6.19), first observe that (3.12) implies

$$\partial_{\vec{\mathrm{x}}}\mathbb{E}_{\boldsymbol{\xi}}(\vec{\mathrm{x}})^T\vec{\rho_x}(\vec{\mathrm{x}}) = -\sum_{k=1}^{K-1}\left[\psi'\left(\frac{x_{k+1}-x_k}{\delta_{k+1}}\right) - \psi'\left(\frac{x_k-x_{k-1}}{\delta_k}\right)\right]\rho_x(x_k) + \int_0^M V_x\left(\mathbf{X}_{\boldsymbol{\xi}}[\vec{\mathrm{x}}]\right)\sum_{k=1}^{K-1}\rho_x(x_k)\theta_k\,\mathrm{d}\xi.$$

We consider both terms on the right hand side separately. Using that $\psi'(1/r) = -P(r)$ and that $\rho_x(x_0) = \rho_x(x_K)$, we see that the first equals to

$$\sum_{k=1}^{K}\psi'\left(\frac{x_k-x_{k-1}}{\delta_k}\right)\left(\rho_x(x_k)-\rho_x(x_{k-1})\right) = -\sum_{k=1}^{K}\int_{x_{k-1}}^{x_k}\mathrm{P}\left(\frac{\delta_k}{x_k-x_{k-1}}\right)\rho_{xx}(\hat{x}_k)\,\mathrm{d}x,$$

with a suitable $\hat{x}_k \in (x_{k-1}, x_k)$ by the intermediate value theorem. For the second term, we perform a change of variables:

$$\int_{\xi_{k-1}}^{\xi_k} V_x\big(x_k\theta_k + x_{k-1}\theta_{k-1}\big)\big(\rho_x(x_k)\theta_k + \rho_x(x_{k-1})\theta_{k-1}\big)\,\mathrm{d}\xi = \int_{x_{k-1}}^{x_k} V_x(x)\rho_x(\check{x})\frac{\delta_k}{x_k-x_{k-1}}\,\mathrm{d}x,$$

with some x-dependent intermediate value $\check{x} \in (x_{k-1}, x_k)$. Summation over $k = 1, \dots, K$ provides (6.19). □

Lemma 6.12. *Let $\vartheta \in C_c^\infty(0,T)$ be a non-negative test function of compact support in $(0,T)$. Then*

$$\int_0^T\int_I \vartheta'(t)\rho(x)u_*(t;x)\,\mathrm{d}x\,\mathrm{d}t \le \int_0^T\int_I \vartheta(t)\left[-P(u_*)\rho_{xx} + V_x\rho_x u\right]\,\mathrm{d}x\,\mathrm{d}t. \tag{6.20}$$

Proof. Multiply inequality (6.15)–(6.17) by $\tau\vartheta(n\tau) \ge 0$, and sum over $n = 1, \dots, N_\tau$. On the left-hand side, it follows by means of (6.18) that

$$\tau\sum_{n=1}^{N_\tau}\vartheta(n\tau)\frac{\mathbf{A}_{\boldsymbol{\xi}}\left(\vec{\mathrm{x}}_\Delta^n\right) - \mathbf{A}_{\boldsymbol{\xi}}\left(\vec{\mathrm{x}}_\Delta^{n-1}\right)}{\tau} = -\tau\sum_{n=0}^{N_\tau-1}\frac{\vartheta((n+1)\tau)-\vartheta(n\tau)}{\tau}\mathbf{A}_{\boldsymbol{\xi}}\left(\vec{\mathrm{x}}_\Delta^n\right) = -\int_0^T\int_I \widehat{\vartheta'_\tau}(t)\rho(x)\bar{u}_\Delta(t;x)\,\mathrm{d}x\,\mathrm{d}t,$$

where the sequence of piecewise constant functions $\widehat{\vartheta'_\tau}$ converge to ϑ' uniformly on $[0,T]$. The strong convergence of $\bar{u}_\Delta$ to u_* in L^1 is sufficient to pass to the limit $j \to \infty$.

On the right-hand side, the first term can be rewritten using (6.19):

$$\tau\sum_{n=1}^{N_\tau}\vartheta(n\tau)\big(\partial_{\vec{\mathrm{x}}}\mathbb{E}_{\boldsymbol{\xi}}(\vec{\mathrm{x}}_\Delta^n)^T\vec{\rho_x}(\vec{\mathrm{x}}_\Delta^n)\big) = \int_0^T\bar{\vartheta}_\tau(t)\int_I\left[-P(\bar{u}_\Delta)\rho_{xx}(\hat{x}) + V_x(x)\rho_x(\check{x})\bar{u}_\Delta\right]\,\mathrm{d}x\,\mathrm{d}t,$$

where, for $(n-1)\tau < t \le n\tau$, we know that $|x-\hat{x}|, |x-\check{x}| < \overline{\delta}(\vec{\mathrm{x}}_\Delta^n)$. The convergence (6.3) implies that $\rho_{xx}(\hat{x})$ and $\rho_x(\check{x})$ converge to their respective limits $\rho_{xx}(x)$ and $\rho_x(x)$ uniformly in $x \in I$. Likewise, the piecewise constant interpolants $\bar{\vartheta}_\tau$ converge to ϑ uniformly on $[0,T]$. The strong convergence of $\bar{u}_\Delta$ and of $P(\bar{u}_\Delta)$ to their respective limits u_* and $P(u_*)$ in $L^1([0,T]\times I)$ thus suffices to pass to the limit with the integral.

Finally we observe that the remainder terms, resulting from (6.16) and (6.17), vanish in the limit $j \to \infty$. This is an easy consequence of (6.1), (6.3) and of the estimates (4.14) and (4.15). □

https://doi.org/10.1051/m2an/2013126 Published online by Cambridge University Press

Proof of Proposition 6.6. First, observe that inequality (6.20) is actually an equality, since ρ can be replaced by $-\rho$ everywhere. Next, recall that functions φ of the type $\varphi(t,x) = \vartheta(t)\rho(x)$, where $\vartheta \in C_c^\infty(0,T)$ is non-negative, and $\rho \in C^\infty(I)$ satisfies $\rho_x(a) = \rho_x(b) = 0$, are dense in the set $\mathcal{D}$ of test functions. Thus, the weak formulation (6.6) holds for all $\varphi \in \mathcal{D}$.

It remains to prove that the solution attains the initial datum u^0 weakly-$\star$ as $t \downarrow 0$. However, this is a trivial consequence of the Hölder regularity of $u_* \in C^{1/2}([0,T];\mathcal{P}(I))$, of the uniform convergence of $(\bar{u}_{\Delta_j})_{j\in\mathbb{N}}$ to u_* w.r.t. $t \in [0,T]$ in $\mathcal{P}(I)$, and of the approximation of u^0 by $u^0_{\Delta_j}$. □

7. Numerical results and proof of consistency

7.1. Implementation

7.1.1. Choice of the initial condition

The numerical scheme is phrased in Lagrangian coordinates: the discretization $\boldsymbol{\xi} = (\xi_0, \xi_1, \dots, \xi_K)$ of the reference domain $[0,M]$ is fixed, whereas the corresponding grid points $\vec{\mathrm{x}}^n = (x_1^n, \dots, x_{K-1}^n) \in \mathfrak{x}_{\boldsymbol{\xi}}$ on the interval I evolve in (discrete) time. In the numerical experiments that follows, our choice for the discretization of the initial condition is to use an equidistant grid $\vec{\mathrm{x}}^0$ with K vertices on I,

$$x_k^0 = a + k(b-a)/K,$$

and an accordingly adapted mesh $\boldsymbol{\xi}$ on $[0,M]$, with

$$\xi_k = U^0(x_k^0), \quad \text{where} \quad U^0(x) = \int_a^x u^0(y)\,\mathrm{d}y \quad \text{for all } x \in I$$

is the initial datum's distribution function. This discretization has the property that

$$\int_{x_{k-1}^0}^{x_k^0} u^0(x)\,\mathrm{d}x = \int_{x_{k-1}^0}^{x_k^0} u_\Delta^0(x)\,\mathrm{d}x \quad \text{for all } k = 1, \dots, K.$$

7.1.2. Time stepping

Each (time) step in the numerical scheme consists of solving the system (4.3) of Euler–Lagrange equations. In practice, this is done with a damped Newton method, which guarantees that the constraint $\vec{\mathrm{x}}_\Delta^n \in \mathfrak{x}_{\boldsymbol{\xi}}$ – *i.e.*, that $a < x_1^n < \cdots < x_{K-1}^n < b$ – is propagated from the $n-1$st to the nth iterate. To be more precise, recall that

$$\partial_{\vec{\mathrm{x}}}\mathbb{E}_{\boldsymbol{\xi},\tau}(\vec{\mathrm{x}}) = \frac{1}{\tau}\mathrm{W}(\vec{\mathrm{x}} - \vec{\mathrm{x}}_\Delta^{n-1}) + \partial_{\vec{\mathrm{x}}}\mathbb{E}_{\boldsymbol{\xi}}(\vec{\mathrm{x}})$$

is the functional whose unique root in $\mathfrak{x}_{\boldsymbol{\xi}}$ defines the nth time iterate $\vec{\mathrm{x}}_\Delta^n$, and that

$$\partial_{\vec{\mathrm{x}}}^2\mathbb{E}_{\boldsymbol{\xi},\tau}(\vec{\mathrm{x}}) = \frac{1}{\tau}\mathrm{W} + \partial_{\vec{\mathrm{x}}}^2\mathbb{E}_{\boldsymbol{\xi}}(\vec{\mathrm{x}})$$

is its Jacobian. Given $\vec{\mathrm{x}}_\Delta^{n-1}$, we calculate $\vec{\mathrm{x}}_\Delta^n$ by means of the following algorithm:

$\vec{\mathrm{x}} := \vec{\mathrm{x}}_\Delta^{n-1}$;
repeat
  $\mathrm{d}\vec{\mathrm{x}} := -\left(\partial_{\vec{\mathrm{x}}}^2\mathbb{E}_{\boldsymbol{\xi},\tau}(\vec{\mathrm{x}})\right)^{-1}\partial_{\vec{\mathrm{x}}}\mathbb{E}_{\boldsymbol{\xi},\tau}(\vec{\mathrm{x}})$;
  while $\vec{\mathrm{x}} + \mathrm{d}\vec{\mathrm{x}} \notin \mathfrak{x}_{\boldsymbol{\xi}}$ $\mathrm{d}\vec{\mathrm{x}} := 0.5\,\mathrm{d}\vec{\mathrm{x}}$; end;
  $\vec{\mathrm{x}} := \vec{\mathrm{x}} + \mathrm{d}\vec{\mathrm{x}}$;
until $\|\mathrm{d}\vec{\mathrm{x}}\|_{l^1} <$ tol and $\|\partial_{\vec{\mathrm{x}}}\mathbb{E}_{\boldsymbol{\xi},\tau}(\vec{\mathrm{x}})\|_{l^1} <$ tol;
$\vec{\mathrm{x}}_\Delta^n := \vec{\mathrm{x}}$.

https://doi.org/10.1051/m2an/2013126 Published online by Cambridge University Press

In our experiments, we use $\mathsf{tol} = 10^{-8}$. For the evaluation of $\partial_{\vec{x}}\mathbb{E}_{\boldsymbol{\xi},\tau}(\vec{x})$ above, an explicit expression for the integrals

$$\partial_{x_m}\left(\int_0^M V\circ\mathbf{X}_{\boldsymbol{\xi}}[\vec{x}]\,\mathrm{d}\xi\right) = \int_0^M V_x\circ\mathbf{X}_{\boldsymbol{\xi}}[\vec{x}]\theta_m(\xi)\,\mathrm{d}\xi$$

is needed, see (3.12). Denoting by $\mathfrak{V}$ an anti-derivative of V, one finds

$$\begin{aligned}\int_{\xi_{m-1}}^{\xi_m} V_x\circ\mathbf{X}_{\boldsymbol{\xi}}[\vec{x}]\theta_m(\xi)\,\mathrm{d}\xi &= \frac{\delta_m}{x_m-x_{m-1}}\int_{x_{m-1}}^{x_m} V_x(x)\frac{x-x_{m-1}}{x_m-x_{m-1}}\,\mathrm{d}x\\ &= \frac{\delta_m}{x_m-x_{m-1}}\left(V(x_m)-\frac{\mathfrak{V}(x_m)-\mathfrak{V}(x_{m-1})}{x_m-x_{m-1}}\right),\end{aligned}$$

and analogously for the integral from ξ_m to ξ_{m+1}. In combination, we obtain

$$\begin{aligned}\partial_{x_k}\int_0^M V\circ\mathbf{X}_{\boldsymbol{\xi}}\left[\vec{x}\right](\xi)\,\mathrm{d}\xi = {}& \frac{-\delta_k}{(x_k-x_{k-1})^2}\left(\mathfrak{V}(x_k)-\mathfrak{V}(x_{k-1})\right)+\frac{\delta_{k-1}}{(x_{k+1}-x_k)^2}\left(\mathfrak{V}(x_{k+1})-\mathfrak{V}(x_k)\right)\\ &+V(x_k)\left(\frac{\delta_k}{x_k-x_{k-1}}-\frac{\delta_{k+1}}{x_{k+1}-x_k}\right).\end{aligned}$$

A similar expression is available for the respective contribution to the Hessian $\partial^2_{\vec{x}}\mathbb{E}_{\boldsymbol{\xi},\tau}$:

$$\begin{aligned}&\partial_{x_m}\partial_{x_k}\left(\int_0^M V\circ\mathbf{X}_{\boldsymbol{\xi}}\left[\vec{x}\right]\,\mathrm{d}\xi\right)\\ &\quad=\begin{cases}\frac{2\delta_k}{(x_k-x_{k-1})^3}\left(\mathfrak{V}(x_k)-\mathfrak{V}(x_{k-1})\right)+\frac{2\delta_{k+1}}{(x_{k+1}-x_k)^3}\left(\mathfrak{V}(x_{k+1})-\mathfrak{V}(x_k)\right)\\ -2V(x_k)\left(\frac{\delta_k}{(x_k-x_{k-1})^2}+\frac{\delta_{k+1}}{(x_{k+1}-x_k)^2}\right)-V_x\left(\frac{\delta_k}{(x_k-x_{k-1})^2}-\frac{\delta_{+1}k}{(x_{k+1}-x_k)^2}\right), & m=k\\ -\frac{2\delta_k}{(x_k-x_{k-1})^3}\left(\mathfrak{V}(x_k)-\mathfrak{V}(x_{k-1})\right)+\frac{\delta_k}{(x_k-x_{k-1})^2}\left(V(x_k)+V(x_{k-1})\right), & m=k-1\\ 0, & \text{otherwise.}\end{cases}\end{aligned}$$

7.2. Numerical experiments and convergence

The following numerical experiments are performed for the porous medium equation with quadratic nonlinearity,

$$\partial_t u = (u^2)_{xx} + (V_x u)_x,$$

on the interval $I=[-1,1]$. For the potential V, we choose

$$V(x) = -\frac{1}{\pi}\cos(\pi x),$$

and as initial datum, we take the following function of unit mass $M=1$:

$$u^0(x) = C\big(-\cos(2\pi x)+1.5\big)\big((x+0.5)^4+1\big) \quad \text{with} \quad C=\frac{240-280\pi^2+423\pi^4}{80\pi^4}. \tag{7.1}$$

7.2.1. Reference solution

Our numerical reference resolution is calculated with $K=5000$ spatial grid points and a time step size $\tau=10^{-2}$. Figure 2/left shows snapshots of the reference solution's spatial density after the first couple of time steps. One observes the typical behaviour for nonlinear drift diffusion equations: on a very short time scale, diffusion reduces the extrema of the initial mass distribution; subsequently, the drift dominates and transports the mass towards the equilibrium (dotted line) on a longer time scale. Figure 2/right displays the corresponding particle trajectories in the Lagrangian picture, *i.e.*, how the points x_k^n move with (discrete) time n for fixed k.

https://doi.org/10.1051/m2an/2013126 Published online by Cambridge University Press

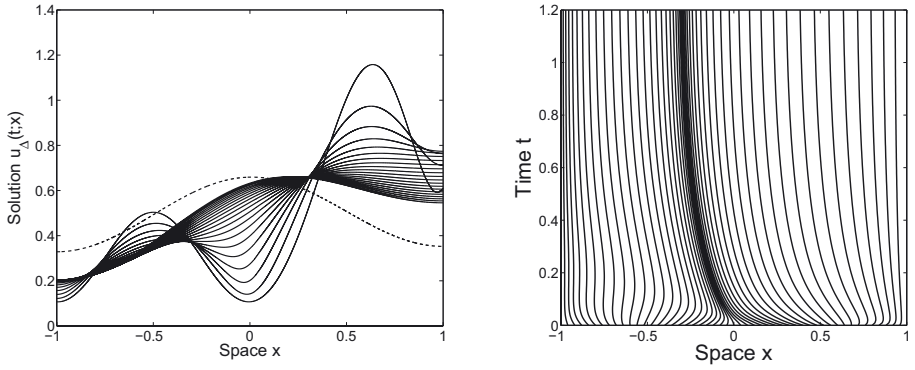


FIGURE 2. Left: evolution of the (reference) solution u_Δ with initial condition (7.1) at times $t = 0, \tau, \ldots, 20\tau = 0.2$, with time step $\tau = 10^{-2}$ and $K = 5000$ grid points. The dotted line shows the stationary solution. Right: associated particle trajectories.

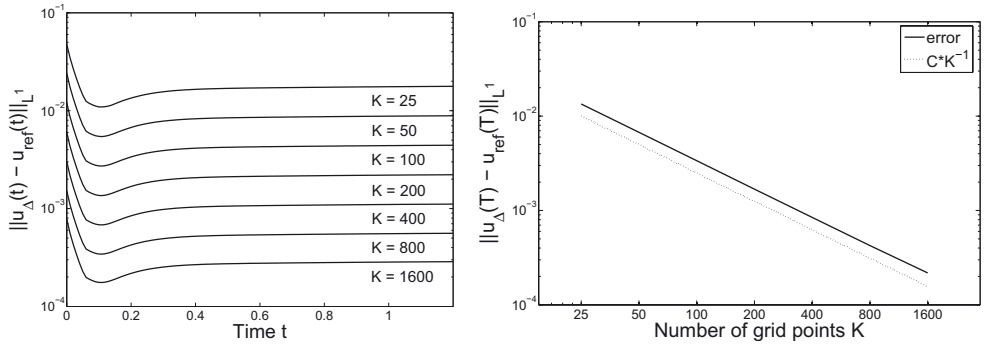


FIGURE 3. Numerical error analysis with fixed time step $\tau = 10^{-2}$, using $K = 25, 50, 100, 200, 400, 800, 1600$ grid points. Left: evolution of the L^1-error $\|\bar{u}_\Delta(t) - u_{\mathrm{ref}}(t)\|_{L^1(I)}$. Right: order of convergence at terminal time $T = 0.2$.

7.2.2. Fixed τ

In a first series of experiments, we fix the time step $\tau = 10^{-2}$ and vary the number of spatial grid points K. In Figure 3/left, the corresponding L^1-distances to the reference solution u_{ref} obtained in 7.2.1 above are shown as a function of time. Note that the counter-intuitive dramatic decay of the error for small times is explained by the strong contractivity of the nonlinear diffusion in L^1 in a neighborhood of the initial condition. Figure 3/right shows the L^1-errors at $T = 0.2$. The observed convergence rate is of order K^{-1}.

7.2.3. Fixed parabolic mesh ratio

Next, we study the decay of the L^1-error under mutual refinement of space and time. As it is standard in numerical experiments on parabolic equations, we fix the parabolic mesh ratio $K^2\tau$. The value of this ratio is chosen such that the inverse CFL condition is satisfied in every experiment. In Figure 4/left, the error is plotted at the fixed terminal time $T = 0.2$ for various choices of τ. The observed order of convergence is $\sqrt{\tau}$, which is in agreement with the result of experiment 7.2.2.

7.2.4. Weakly convergent initial datum

In order to illustrate that it sufficies to approximate the original initial condition u^0 by its discretizations u^0_Δ just *weakly* in $L^1(I)$, we use perturbed discrete initial data $u^0_{\Delta,\varepsilon}$ that are biased by high-frequency oscillations

https://doi.org/10.1051/m2an/2013126 Published online by Cambridge University Press

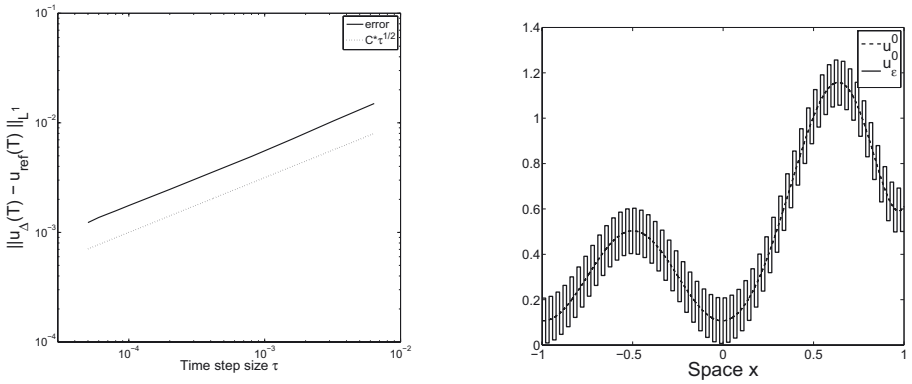


FIGURE 4. Left: numerical error analysis with fixed parabolic mesh ratio $K^2\tau \approx 0.257$, analyzing the L^1-error $\|\bar{u}_\Delta(T) - u_{\mathrm{ref}}(T)\|_{L^1(I)}$ at terminal time $T = 0.2$ using $\tau = 5\times10^{-5}, 10^{-4}, 5\times10^{-4}, 10^{-3}, 5\times10^{-3}, 10^{-2}, 5\times10^{-2}, 10^{-1}$. Right: initial condition $u^0_{\Delta,\varepsilon}$ with high frequency perturbation.

of fixed amplitude 0.1, as indicated in Figure 4/right. As expected, the perturbation becomes almost invisible already after the first time step, and the discrete solution $u_{\Delta,\varepsilon}$ is indistinguishable from the one computed with unperturbed initial conditions u^0_Δ.

7.2.5. A discontinuous initial datum

For the last two series of experiments, we change the initial condition u^0. This first series is carried out with the discontinuous inital datum

$$u^0(x) = \begin{cases} 0.1, & \text{if } |x| > 0.75 \text{ or } |x| < 0.25, \\ 0.9, & \text{otherwise.} \end{cases} \tag{7.2}$$

Similar to experiment 7.2.2, we fix $\tau = 10^{-2}$ and vary the number of grid points K. Figure 5/right displays the corresponding L^1-error over the time interval $t \in [0, 0.8]$. In contrast to experiment 7.2.2, the approximation error is zero initially, since the step function u^0 can be discretized exactly. However, the error jumps to a positive value (that is of the same order as the initial error in experiment 7.2.2) in the first time step. Afterwards, the qualitative behaviour is very similar to that in experiment 7.2.2. The observed order of convergence (at $T = 0.2$) is again K^{-1}.

7.2.6. A non-positive initial datum

For this last series of experiment, we consider the initial condition

$$u^0(x) = (-\cos(2\pi x) + 1.5)\left((x+0.5)^4 + 1\right) \times \begin{cases} -(x-0.5)(x+0.5) & |x| \le 0.5 \\ 0 & |x| > 0.5 \end{cases}, \tag{7.3}$$

which vanishes outside of the subinterval $[-0.5, 0.5] \subset I$. The numerical scheme is not directly applicable to u^0, but to any of its strictly positive approximations $u^0 + \varepsilon$, see Figure 6/left. The qualitative numerical results at $T = 0.6$ for various choices of $\varepsilon > 0$ are given in Figure 6/right.

7.3. Order of consistency

The experimental observations that the discrete solutions seem to approximate the reference solution with an error of order $K\tau$ can be supported theoretically by the following consistency consideration.

https://doi.org/10.1051/m2an/2013126 Published online by Cambridge University Press

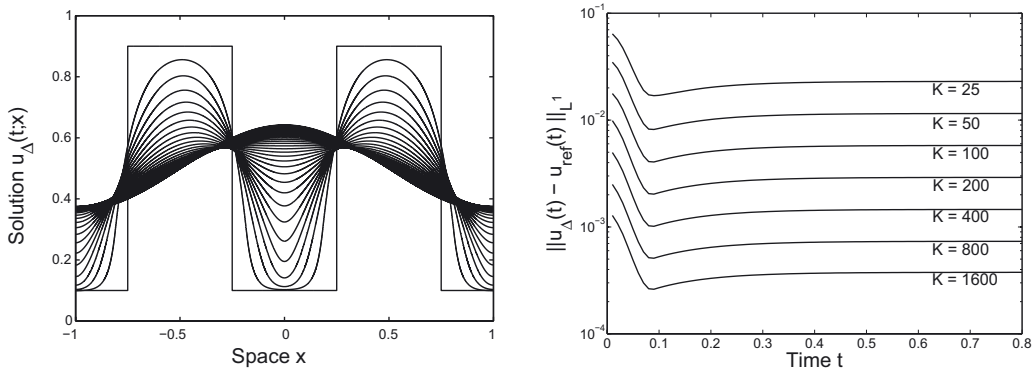


FIGURE 5. Left: evolution of the solution u_Δ with initial condition (7.2) at times $t = 0, \tau, \ldots, 40\tau = 0.2$, with time step $\tau = 5 \times 10^{-3}$ and $K = 5000$ grid points. Right: numerical error analysis for discrete solutions with the discontinuous initial datum from (7.2), using a fixed time step $\tau = 10^{-2}$ and varying $K = 25, 50, 100, 200, 400, 800, 1600$.

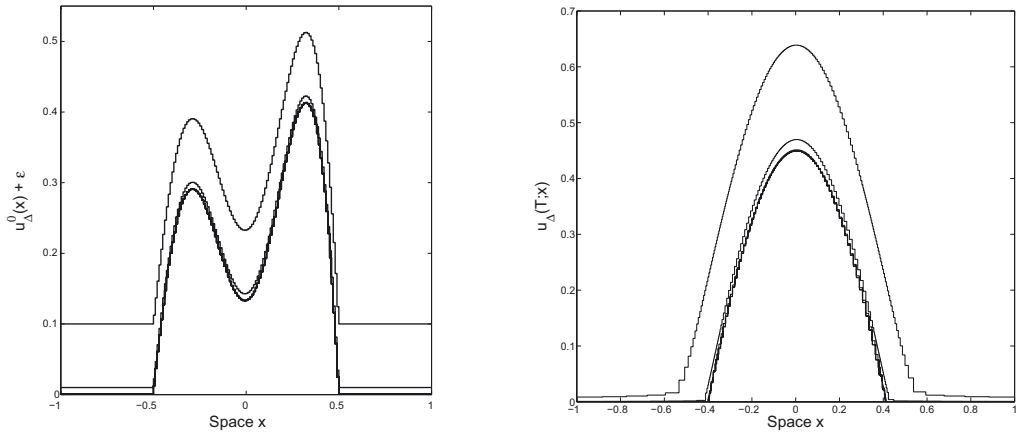


FIGURE 6. The merely non-negative initial condition u^0 from (7.3) is approximated by strictly positive data $u^0 + \varepsilon$. Left: discrete initial profiles for $\varepsilon = 10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}, 10^{-5}$. Right: qualitative behaviour of corresponding discrete solutions at $T = 0.6$, using $\tau = 10^{-3}$, $K = 200$.

Assume that $X : [0,T] \times [0,M] \to I$ is a smooth solution of (1.12). Consider a discretization $\Delta = (\tau; \boldsymbol{\xi})$ that is equidistant w.r.t. ξ, *i.e.*, $\xi_k = Mk/K$ with some $K \in \mathbb{N}$ for all $k = 0, 1, 2, \ldots, K$, and consequently $\delta := \delta_1 = \cdots = \delta_K$.

From X, we define a discrete "pseudo-solution" $\vec{\mathrm{x}}_\Delta$ by restriction, *i.e.*, $x_k^n := \mathrm{X}(n\tau, k\delta)$. We are going to show that $\vec{\mathrm{x}}_\Delta$ satisfies the discrete evolution equation (4.3) up to an error of order $\delta(\mathcal{O}(\tau) + \mathcal{O}(\delta^2))$. To this end, we perform a Taylor expansion of X around a fixed point $p := (n\tau, k\delta)$ w.r.t. ξ:

$$x_{k\pm1}^n = \mathrm{X}(p) \pm \delta \mathrm{X}_\xi(p) + \frac{\delta^2}{2}\mathrm{X}_{\xi\xi}(p) \pm \frac{\delta^3}{6}\mathrm{X}_{\xi\xi\xi}(p) + \mathcal{O}(\delta^4).$$

https://doi.org/10.1051/m2an/2013126 Published online by Cambridge University Press

For the diffusion term, we find

$$\psi'\left(\frac{x^n_{k\pm1}-x^n_k}{\delta}\right)=\psi'\left(\mathrm{X}_\xi(p)\right)\pm\frac{\delta}{2}\psi''\left(\mathrm{X}_\xi(p)\right)\mathrm{X}_{\xi\xi}(p)$$
$$+\frac{\delta^2}{2}\left(\frac{1}{3}\psi''\left(\mathrm{X}(p)\right)\mathrm{X}_{\xi\xi\xi}(p)+\frac{1}{4}\psi'''\left(\mathrm{X}(p)\right)\mathrm{X}_{\xi\xi}(p)^2\right)+\mathcal{O}\left(\delta^3\right),$$

so that

$$\psi'\left(\frac{x^n_{k+1}-x^n_k}{\delta}\right)-\psi'\left(\frac{x^n_k-x^n_{k-1}}{\delta}\right)=\delta\psi''\left(\mathrm{X}_\xi(p)\right)\mathrm{X}_{\xi\xi}(p)+\mathcal{O}\left(\delta^3\right).$$

For the drift term, we obtain

$$\begin{aligned}\int_{k\delta}^{(k\pm1)\delta}V_x\left(\mathbf{X}_{\boldsymbol{\xi}}[\vec{\mathrm{x}}_\Delta](\xi)\right)\theta_k(\xi)\,\mathrm{d}\xi&=\pm\delta\int_0^1 V_x\left((1-s)x^n_k+sx^n_{k\pm1}\right)(1-s)\,\mathrm{d}s\\&=\pm\delta\int_0^1\left[V_x\left(\mathrm{X}(p)\right)\pm s\delta V_{xx}\left(\mathrm{X}(p)\right)\mathrm{X}_\xi(p)+\mathcal{O}\left(\delta^2\right)\right](1-s)\,\mathrm{d}s\\&=\pm\frac{\delta}{2}V_x\left(\mathrm{X}(p)\right)+\frac{\delta^2}{6}V_{xx}\left(\mathrm{X}(p)\right)\mathrm{X}_\xi(p)+\mathcal{O}\left(\delta^3\right),\end{aligned}$$

so that

$$\int_{(k-1)\delta}^{(k+1)\delta}V_x\left(\mathbf{X}_{\boldsymbol{\xi}}[\vec{\mathrm{x}}_\Delta](\xi)\right)\theta_k(\xi)\,\mathrm{d}\xi=\delta V_x\left(\mathrm{X}(p)\right)+\mathcal{O}\left(\delta^3\right).$$

In combination,

$$\left[\partial_{\vec{\mathrm{x}}}\mathbb{E}_{\boldsymbol{\xi}}\left(\vec{\mathrm{x}}\right)\right]_k=\delta\left(\psi'\left(\mathrm{X}_\xi(p)\right)_\xi+V_x\left(\mathrm{X}(p)\right)+\mathcal{O}\left(\delta^2\right)\right).\tag{7.4}$$

Moreover, we have

$$\left[\mathrm{W}\vec{\mathrm{x}}^n_\Delta\right]_k=\frac{\delta}{6}x^n_{k-1}+\frac{2\delta}{3}x^n_k+\frac{\delta}{6}x^n_{k+1}=\delta\mathrm{X}(p)+\mathcal{O}\left(\delta^3\right),$$

and likewise

$$\left[\mathrm{W}\vec{\mathrm{x}}^{n-1}_\Delta\right]_k=\delta\mathrm{X}(p')+\mathcal{O}\left(\delta^3\right),$$

where $p'=((n-1)\tau,k\delta)$. Finally, using

$$\mathrm{X}(p)-\mathrm{X}(p')=\tau\mathrm{X}_t(p)+\mathcal{O}\left(\tau^2\right),$$

we obtain the relation

$$\frac{1}{\tau}\left[\mathrm{W}\left(\vec{\mathrm{x}}^n_\Delta-\vec{\mathrm{x}}^{n-1}_\Delta\right)\right]_k=\delta\left(\mathrm{X}_t(p)+\mathcal{O}\left(\delta^2\right)+\mathcal{O}(\tau)\right).\tag{7.5}$$

Using the continuous evolution equation (1.12) in (7.5) and (7.4) leads to

$$\frac{1}{\tau}\left[\mathrm{W}\left(\vec{\mathrm{x}}^n_\Delta-\vec{\mathrm{x}}^{n-1}_\Delta\right)\right]_k=-\left[\partial_{\vec{\mathrm{x}}}\mathbb{E}_{\boldsymbol{\xi}}\left(\vec{\mathrm{x}}\right)\right]_k+\delta\left(\mathcal{O}(\tau)+\mathcal{O}\left(\delta^2\right)\right).$$

for all admissible k and n.

Acknowledgements. The authors would like to thank the two anonymous referees for their careful reading of the manuscript and their manifold suggestions for improvement.

https://doi.org/10.1051/m2an/2013126 Published online by Cambridge University Press

References

[1] L. Ambrosio, N. Gigli and G. Savaré, Gradient flows in metric spaces and in the space of probability measures. *Lectures in Mathematics*. ETH Zürich, Birkhäuser Verlag, Basel (2005).

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

[3] J.-D. Benamou and Y. Brenier, A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem. *Numer. Math.* **84** (2000) 375–393.

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

[5] C.J. Budd, G.J. Collins, W.Z. Huang and R.D. Russell, Self-similar numerical solutions of the porous-medium equation using moving mesh methods. *R. Soc. Lond. Philos. Trans. Ser. A Math. Phys. Eng. Sci.* **357** (1999) 1047–1077.

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

[7] J.A. Carrillo and J.S. Moll, Numerical simulation of diffusive and aggregation phenomena in nonlinear continuity equations by evolving diffeomorphisms. *SIAM J. Sci. Comput.* **31** (2009/2010) 4305–4329.

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

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

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

[11] E. Giusti, Minimal surfaces and functions of bounded variation, vol. 80, *Monographs in Mathematics*. Birkhäuser Verlag, Basel (1984).

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

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

[14] D. Kinderlehrer and N.J. Walkington, Approximation of parabolic equations using the Wasserstein metric. *ESAIM: M2AN* **33** (1999) 837–852.

[15] M. Leven, *Gradientenfluß-basierte diskretisierung parabolischer gleichungen*, diplomarbeit, Universität Bonn (2002).

[16] R.C. MacCamy and E. Socolovsky, A numerical procedure for the porous media equation. Hyperbolic partial differential equations, II. *Comput. Math. Appl.* **11** (1985) 315–319.

[17] R.J. McCann, A convexity principle for interacting gases. *Adv. Math.* **128** (1997) 153–179.

[18] F. Otto, The geometry of dissipative evolution equations: the porous medium equation. *Comm. Partial Differential Equations* **26** (2001) 101–174.

[19] T. Roessler, *Discretizing the porous medium equation based on its gradient flow structure – a consistency paradox*, Technical report 150, Sonderforschungsbereich 611, May 2004. Available online at http://sfb611.iam.uni-bonn.de/uploads/150-komplett.pdf.

[20] R. Rossi and G. Savaré, Tightness, integral equicontinuity and compactness for evolution problems in Banach spaces. *Ann. Sc. Norm. Super. Pisa Cl. Sci.* **2** (2003) 395–431.

[21] G. Russo, Deterministic diffusion of particles. *Commun. Pure Appl. Math.* **43** (1990) 697–733.

[22] S. Serfaty, Gamma-convergence of gradient flows on Hilbert and metric spaces and applications, *Discrete Contin. Dyn. Syst.* **31** (2011) 1427-1451.

[23] C. Villani, Topics in optimal transportation, in vol. 58 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI (2003).

[24] M. Westdickenberg and J. Wilkening, Variational particle schemes for the porous medium equation and for the system of isentropic Euler equations. *ESAIM: M2AN* **44** (2010) 133–166.

https://doi.org/10.1051/m2an/2013126 Published online by Cambridge University Press