

# On the total variation Wasserstein gradient flow and the TV-JKO scheme

Guillaume CARLIER\*      Clarice POON<sup>†</sup>

July 21, 2018

## Abstract

We study the JKO scheme for the total variation, characterize the optimizers, prove some of their qualitative properties (in particular a form of maximum principle and in some cases, a minimum principle as well). Finally, we establish a convergence result as the time step goes to zero to a solution of a fourth-order nonlinear evolution equation, under the additional assumption that the density remains bounded away from zero, this lower bound is shown in dimension one and in the radially symmetric case.

**Keywords:** total variation, Wasserstein gradient flows, JKO scheme, fourth-order evolution equations.

**MS Classification:** 35G31, 49N15.

## 1 Introduction

Variational schemes based on total variation are extremely popular in image processing for denoising purposes, in particular the seminal work of Rudin, Osher and Fatemi [25] has been extremely influential and is still the object of an intense stream of research, see [10] and the references therein. Continuous-time counterparts are well-known to be related to the  $L^2$  gradient flow of the total variation, see Bellettini, Caselles and Novaga [3] and the mean-curvature

---

\*Ceremade, UMR CNRS 7534, Université Paris Dauphine, Pl. de Lattre de Tassigny, 75775, Paris Cedex 16, France, and MOKAPLAN, INRIA-Paris, E-mail: carlier@ceremade.dauphine.fr

<sup>†</sup>Centre for Mathematical Sciences, University of Cambridge, Wilberforce Rd, Cambridge CB3 0WA, United Kingdom, Email: C.M.H.S.Poon@maths.cam.ac.uk

flow, see Evans and Spruck [14]. The gradient flow of the total variation for other Hilbertian structures may be natural as well and in particular the  $H^{-1}$  case, leads to a singular fourth-order evolution equation studied by Giga and Giga [15], Giga, Kuroda and Matsuoka [16]. In the present work, we consider another metric, namely the Wasserstein one.

Given an open subset  $\Omega$  of  $\mathbb{R}^d$  and  $\rho \in L^1(\Omega)$ , recall that the total variation of  $\rho$  is given by

$$J(\rho) := \sup \left\{ \int_{\Omega} \operatorname{div}(z) \rho : z \in C_c^1(\Omega), \|z\|_{L^\infty} \leq 1 \right\} \quad (1.1)$$

and  $\operatorname{BV}(\Omega)$  is by definition the subspace of  $L^1(\Omega)$  consisting of those  $\rho$ 's in  $L^1(\Omega)$  such that  $J(\rho)$  is finite. The following fourth-order nonlinear evolution equation

$$\partial_t \rho + \operatorname{div} \left( \rho \nabla \operatorname{div} \left( \frac{\nabla \rho}{|\nabla \rho|} \right) \right) = 0, \text{ in } (0, T) \times \Omega, \rho|_{t=0} = \rho_0, \quad (1.2)$$

supplemented by the zero-flux boundary condition

$$\rho \nabla \operatorname{div} \left( \frac{\nabla \rho}{|\nabla \rho|} \right) \cdot \nu = 0 \text{ on } \partial\Omega \quad (1.3)$$

has been proposed in [7] for the purpose of denoising image densities. Numerical schemes for approximating the solutions of this equation have been investigated in [7, 13, 4]. One should consider weak solutions and in particular interpret the nonlinear term  $\operatorname{div}(\frac{\nabla \rho}{|\nabla \rho|})$  as the negative of an element of the subdifferential of  $J$  at  $\rho$ .

At least formally, when  $\rho_0$  is a probability density on  $\Omega$ , (1.2)-(1.3) can be viewed as the Wasserstein gradient flow of  $J$  (we refer to the textbooks of Ambrosio, Gigli, Savaré [1] and Santambrogio [26], for a detailed exposition). Following the seminal work of Jordan, Kinderlehrer and Otto [17] for the Fokker-Planck equation, it is reasonable to expect that solutions of (1.2) can be obtained, at the limit  $\tau \rightarrow 0^+$ , of the JKO Euler implicit scheme:

$$\rho_0^\tau = \rho_0, \rho_{k+1}^\tau \in \operatorname{argmin} \left\{ \frac{1}{2\tau} W_2^2(\rho_k^\tau, \rho) + J(\rho), \rho \in \operatorname{BV}(\Omega) \cap \mathcal{P}_2(\overline{\Omega}) \right\} \quad (1.4)$$

where  $\mathcal{P}_2(\overline{\Omega})$  is the space of Borel probability measures  $\overline{\Omega}$  with finite second moment and  $W_2$  is the quadratic Wasserstein distance:

$$W_2^2(\rho_0, \rho_1) := \inf_{\gamma \in \Pi(\rho_0, \rho_1)} \left\{ \int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^2 d\gamma(x, y) \right\}, \quad (1.5)$$

$\Pi(\rho_0, \rho_1)$  denoting the set of transport plans between  $\rho_0$  and  $\rho_1$  i.e. the set of probability measures on  $\mathbb{R}^d \times \mathbb{R}^d$  having  $\rho_0$  and  $\rho_1$  as marginals. Our aim is to study in detail the discrete TV-JKO scheme (1.4) as well as its connection with (suitable weak solutions) of the PDE (1.2). Although the assertion that (1.2) is the TV Wasserstein gradient flow is central to the numerical schemes described in [7, 13, 4], there has been so far, to the best of our knowledge, no theoretical justification of this fact.

Fourth-order equations which are Wasserstein gradient flows of functionals involving the gradient of  $\rho$ , such as the Dirichlet energy or the Fisher information, have been studied by McCann, Matthes and Savaré [22] who found a new method, *the flow interchange technique*, to prove higher-order estimates, we refer to [18] for a recent reference on this topic. The total variation is however too singular for such arguments to be directly applicable, as far as we know. We shall prove the convergence of JKO steps as  $\tau \rightarrow 0^+$  under the extra assumption that densities remain bounded away from zero. Whether this extra assumption is reasonable or not is related to a minimum principle issue, interesting in its own right, namely the monotonicity of the infimum along JKO steps. We shall see that, in a convex domain, JKO steps obey a maximum principle (the maximum of the density is non-increasing along JKO steps). The corresponding minimum principle seems more difficult to prove and we have been able to establish it only in some particular cases, namely in dimension one and in the radially symmetric case, even though we conjecture it is satisfied in more general situations.

The paper is organized as follows. In section 2, we start with the discussion of a few examples. Section 3 establishes optimality conditions for JKO steps thanks to an entropic regularization scheme. Section 4 is devoted to some properties of solutions of JKO steps and in particular a maximum principle based on a result of [11], we also establish a minimum principle in dimension one and in the radially symmetric case. Finally, in section 5, we prove a conditional convergence result, we establish convergence of the JKO scheme, as  $\tau \rightarrow 0^+$ , under the extra assumption that the density remains away from zero, this covers the unidimensional case as well as the radially symmetric case when the initial condition is strictly positive.

## 2 Some examples

We first recall the Kantorovich dual formulation of  $W_2^2$ :

$$\frac{1}{2}W_2^2(\mu_0, \mu_1) = \sup \left\{ \int_{\mathbb{R}^d} \psi d\mu_0 + \int_{\mathbb{R}^d} \varphi d\mu_1 : \psi(x) + \varphi(y) \leq \frac{|x - y|^2}{2} \right\} \quad (2.1)$$

an optimal pair  $(\psi, \varphi)$  for this problem is called a pair of Kantorovich potentials. The existence of Kantorovich potentials is well-known and such potentials can be taken to be conjugates of each other, i.e. such that

$$\varphi(x) = \inf_{y \in \mathbb{R}^d} \left\{ \frac{1}{2} |x - y|^2 - \psi(y) \right\}, \quad \psi(y) = \inf_{x \in \mathbb{R}^d} \left\{ \frac{1}{2} |x - y|^2 - \varphi(x) \right\},$$

which implies that  $\varphi$  and  $\psi$  are semi-concave (more precisely  $\frac{1}{2}|\cdot|^2 - \varphi$  is convex). If  $\mu_1$  is absolutely continuous with respect to the  $d$ -dimensional Lebesgue measure,  $\varphi$  is differentiable  $\mu_1$  a.e. and the map  $T = \text{id} - \nabla \varphi$  is the gradient of a convex function pushing forward  $\mu_1$  to  $\mu_0$  which is in fact the optimal transport between  $\mu_0$  and  $\mu_1$  thanks to Brenier's theorem [5]. In such a case, we will simply refer to  $\varphi$  as a Kantorovich potential between  $\mu_1$  and  $\mu_0$ . We refer the reader to [28] and [26] for details.

In this section, we will consider some explicit examples which rely on the following sufficient optimality condition (details for a rigorous derivation of the Euler-Lagrange equation for JKO steps will be given in section 3) in the case of the whole space i.e.  $\Omega = \mathbb{R}^d$ . Let us also recall that by Sobolev inequality  $\text{BV}(\mathbb{R}^d)$  is continuously embedded in  $L^{\frac{d}{d-1}}(\mathbb{R}^d)$ .

**Lemma 2.1.** *Let  $\rho_0 \in \mathcal{P}_2(\mathbb{R}^d)$ ,  $\tau > 0$  and  $\Omega = \mathbb{R}^d$  (so  $J$  is the total variation on the whole space), if  $\rho_1 \in \text{BV}(\mathbb{R}^d) \cap \mathcal{P}_2(\mathbb{R}^d)$  is such that*

$$\frac{\varphi}{\tau} + \text{div}(z) \geq 0, \quad \text{with equality } \rho_1\text{-a.e.} \quad (2.2)$$

where  $\varphi$  is a Kantorovich potential between  $\rho_1$  and  $\rho_0$  and  $z \in C^1(\mathbb{R}^d)$ , with  $\|z\|_{L^\infty} \leq 1$ ,  $\text{div}(z) \in L^d(\mathbb{R}^d)$  (so that  $\text{div}(z)\rho_1 \in L^1(\mathbb{R}^d)$ ), and

$$J(\rho_1) = \int_{\mathbb{R}^d} \text{div}(z)\rho_1. \quad (2.3)$$

Then, setting

$$\Phi_{\tau, \rho_0}(\rho) := \frac{1}{2\tau} W_2^2(\rho_0, \rho) + J(\rho), \quad \forall \rho \in \text{BV}(\mathbb{R}^d) \cap \mathcal{P}_2(\mathbb{R}^d) \quad (2.4)$$

one has

$$\Phi_{\tau, \rho_0}(\rho_1) \leq \Phi_{\tau, \rho_0}(\rho), \quad \forall \rho \in \text{BV}(\mathbb{R}^d) \cap \mathcal{P}_2(\mathbb{R}^d).$$

*Proof.* For all  $\rho \in \text{BV}(\mathbb{R}^d) \cap \mathcal{P}_2(\mathbb{R}^d)$ ,  $J(\rho) \geq \int_{\mathbb{R}^d} \text{div}(z)\rho = J(\rho_1) + \int_{\mathbb{R}^d} \text{div}(z)(\rho - \rho_1)$ , and it follows from the Kantorovich duality formula that

$$\frac{1}{2\tau} W_2^2(\rho_0, \rho) \geq \frac{1}{2\tau} W_2^2(\rho_0, \rho_1) + \int_{\mathbb{R}^d} \frac{\varphi}{\tau} (\rho - \rho_1).$$

The claim then directly follows from (2.2).  $\square$

## 2.1 The case of a characteristic function

A simple illustration of Lemma 2.1 in dimension 1 concerns the case of a uniform  $\rho_0$ , (here and in the sequel we shall denote by  $\chi_A$  the characteristic function of the set  $A$ ):

$$\rho_0 = \rho_{\alpha_0}, \quad \alpha_0 > 0, \quad \rho_\alpha := \frac{1}{2\alpha} \chi_{[-\alpha, \alpha]}.$$

It is natural to make the ansatz that the minimizer of  $\Phi_{\tau, \rho_0}$  defined by (2.4) remains of the form  $\rho_1 = \rho_{\alpha_1}$  for some  $\alpha_1 > \alpha_0$ . The optimal transport between  $\rho_{\alpha_1}$  and  $\rho_0$  being the linear map  $T = \frac{\alpha_0}{\alpha_1} \text{id}$ , a direct computation gives

$$\Phi_{\tau, \rho_0}(\rho_{\alpha_1}) = \frac{1}{\alpha_1} + \frac{1}{6\tau}(\alpha_1 - \alpha_0)^2$$

which is minimal when  $\alpha_1$  is the only root in  $(\alpha_0, +\infty)$  of

$$\alpha_1^2(\alpha_1 - \alpha_0) = 3\tau. \quad (2.5)$$

To check that this is the correct guess, we shall check that the conditions of Lemma 2.1 are met. It is easy to check that the potential defined by

$$\varphi(x) = \frac{1}{2\alpha_1}(\alpha_1 - \alpha_0)x^2 - \frac{3\tau}{2\alpha_1}$$

is a Kantorovich potential between  $\rho_1 = \rho_{\alpha_1}$  and  $\rho_0$ . Define<sup>1</sup> then  $z_1$  by

$$\tau z_1(x) := -\frac{(\alpha_1 - \alpha_0)}{6\alpha_1}x^3 + \frac{3\tau x}{2\alpha_1}, \quad x \in [-\alpha_1, \alpha_1]$$

extended by 1 on  $[\alpha_1, +\infty)$  and  $-1$  on  $(-\infty, -\alpha_1]$ . By construction  $-1 \leq z_1 \leq 1$  (use the fact that it is odd and nondecreasing on  $[0, \alpha_1]$  thanks to (2.5)), also  $z_1'(\pm\alpha_1) = 0$  so that  $z_1 \in C^1(\mathbb{R})$  and  $z_1(\alpha_1) = 1$ ,  $z_1(-\alpha_1) = -1$  and one easily checks that  $J(\rho_1) = -\int_{\mathbb{R}} z_1 D\rho_1 = \int_{\mathbb{R}} z_1' \rho_1$  (here and in the sequel  $D\rho_1$  denotes the Radon measure which is the distributional derivative of the BV function  $\rho_1$ ). Moreover  $\tau z_1' + \varphi \geq 0$  with an equality on  $[-\alpha_1, \alpha_1]$ . The optimality of  $\rho_1 = \rho_{\alpha_1}$  then directly follows from Lemma 2.1.

Of course, the argument can be iterated so as to obtain the full TV-JKO sequence:

$$\rho_{k+1}^\tau = \operatorname{argmin} \Phi_{\tau, \rho_k^\tau} = \left( \frac{\alpha_{k+1}^\tau}{\alpha_k^\tau} \text{id} \right)_\# \rho_k^\tau = \left( \frac{\alpha_{k+1}^\tau}{\alpha_0} \text{id} \right)_\# \rho_0$$

---

<sup>1</sup>The guess for this construction is by integrating the Euler-Lagrange equation on the support of  $\rho_{\alpha_1}$ .

where  $\alpha_k^\tau$  is defined inductively by

$$(\alpha_{k+1}^\tau - \alpha_k^\tau)(\alpha_{k+1}^\tau)^2 = 3\tau, \quad \alpha_0^\tau = \alpha_0$$

which is nothing but the implicit Euler discretization of the ODE

$$\alpha' \alpha^2 = 3, \quad \alpha(0) = \alpha_0,$$

whose solution is  $\alpha(t) = (\alpha_0^3 + 9t)^{\frac{1}{3}}$ . Extending  $\rho_k^\tau$  in a piecewise constant way:  $\rho^\tau(t) = \rho_{k+1}^\tau$  for  $t \in (k\tau, (k+1)\tau]$ , it is not difficult to check that  $\rho^\tau$  converges (in  $L^\infty((0, T), (\mathcal{P}_2(\mathbb{R}), W_2))$  and in  $L^p((0, T) \times \mathbb{R})$  for any  $p \in (1, \infty)$  and any  $T > 0$ ) to  $\rho$  given by  $\rho(t, \cdot) = (\frac{\alpha(t)}{\alpha_0} \text{id})_\# \rho_0$ . Since  $v(t, x) = \frac{\alpha'(t)}{\alpha(t)} x$  is the velocity field associated to  $X(t, x) = \frac{\alpha(t)}{\alpha_0} x$ ,  $\rho$  solves the continuity equation

$$\partial_t \rho + (\rho v)_x = 0.$$

In addition,  $\rho v = -\rho z_{xx}$  where

$$z(t, x) = \frac{-\alpha'(t)}{6\alpha(t)} x^3 + \frac{3x}{2\alpha(t)}, \quad x \in [-\alpha(t), \alpha(t)],$$

extended by 1 (respectively  $-1$ ) on  $[\alpha(t), +\infty)$  (respectively  $(-\infty, -\alpha(t)]$ ). The function  $z$  is  $C^1$ ,  $\|z\|_{L^\infty} \leq 1$  and  $z \cdot D\rho = -|D\rho|$  (in the sense of measures). In other words the limit  $\rho$  of  $\rho^\tau$  satisfies

$$\partial_t \rho - (\rho z_{xx})_x = 0$$

with  $|z| \leq 1$  and  $z \cdot D\rho = -|D\rho|$  which is the natural weak form of (1.2) since  $z_{xx} = \nabla \text{div}(z)$  in dimension one.

## 2.2 Instantaneous creation of discontinuities

We now consider the case where  $\rho_0(x) = (1 - |x|)_+$  and will show that the JKO scheme instantaneously creates a discontinuity at the level of  $\rho_1$ , the minimizer of  $\Phi_{\tau, \rho_0}$  when  $\tau$  is small enough. We indeed look for  $\rho_1$  in the form:

$$\rho_1(x) = \begin{cases} 1 - \beta/2 & \text{if } |x| < \beta, \\ (1 - |x|)_+ & \text{if } |x| \geq \beta, \end{cases}$$

for some well-chosen  $\beta \in (0, 1)$ . The optimal transport map  $T$  between such a  $\rho_1$  and  $\rho_0$  is odd and given explicitly by

$$T(x) = \begin{cases} 1 - \sqrt{1 - x(2 - \beta)} & \text{if } x \in [0, \beta), \\ x & \text{if } x \geq \beta. \end{cases}$$

The Kantorovich potential which vanishes at  $\beta$  (extended in an even way to  $\mathbb{R}_-$ ) is then given by

$$\varphi(x) = \begin{cases} \frac{x^2}{2} - x - \frac{(1-x(2-\beta))^{3/2}}{3(1-\beta/2)} + C & \text{if } x \in [0, \beta), \\ 0 & \text{if } x > \beta, \end{cases}$$

where

$$C = -\frac{\beta^2}{2} + \beta + \frac{2(1-\beta)^3}{3(2-\beta)}.$$

Let us now integrate  $\tau z' = -\varphi$  on  $[0, \beta]$  with initial condition  $z(0) = 0$ , i.e. for  $x \in [0, \beta]$

$$\begin{aligned} \tau z(x) = & -\frac{x^3}{6} + \frac{x^2}{2} - \frac{4}{15(2-\beta)^2} [1 - (1-2\beta)x]^{\frac{5}{2}} \\ & + \left( \frac{\beta^2}{2} - \beta - \frac{2(1-\beta)^3}{3(2-\beta)} \right) x + \frac{4}{15(2-\beta)^2} \end{aligned}$$

Note that  $z$  is nondecreasing on  $[0, \beta]$  (because  $\varphi(0) < 0$ ,  $\varphi(\beta) = 0$  and  $\varphi$  is convex on  $[0, \beta]$  so that  $\varphi \leq 0$  on  $[0, \beta]$ ), our aim now is to find  $\beta \in (0, 1)$  in such a way that  $z(\beta) = 1$  i.e. replacing in the previous formula

$$\tau = \frac{\beta^3}{3} - \frac{\beta^2}{2} + \frac{4(1-(1-\beta)^5)}{15(2-\beta)^2} - \frac{2(1-\beta)^3\beta}{3(2-\beta)}$$

the right hand-side is a continuous function of  $\beta \in [0, 1]$  taking value 0 for  $\beta = 0$  and  $\frac{1}{10}$  for  $\beta = 1$ , hence as soon as  $10\tau < 1$  one may find a  $\beta \in (0, 1)$  such that indeed  $z(\beta) = 1$ . Extend then  $z$  by 1 on  $[\beta, +\infty)$  and to  $\mathbb{R}_-$  in an odd way. We then have built a function  $z$  which is  $C^1$  ( $\varphi(\beta) = 0$ ), such that  $|z| \leq 1$ ,  $z \cdot D\rho_1 = -|D\rho_1|$  and such that  $z' + \frac{\varphi}{\tau} = 0$ . Thanks to Lemma 2.1, we conclude that  $\rho_1$  is optimal. This example shows that discontinuities may appear at the very first iteration of the TV-JKO scheme.

### 3 Euler-Lagrange equation for JKO steps

The aim of this section is to establish optimality conditions for (3.1). Despite the fact that it is a convex minimization problem, it involves two nonsmooth terms  $J$  and  $W_2^2(\rho_0, \cdot)$ , so some care should be taken of to justify rigorously the arguments. In the next subsection, we introduce an entropic regularization, the advantage of this strategy is that the minimizer will be positive everywhere, giving some differentiability of the transport term.

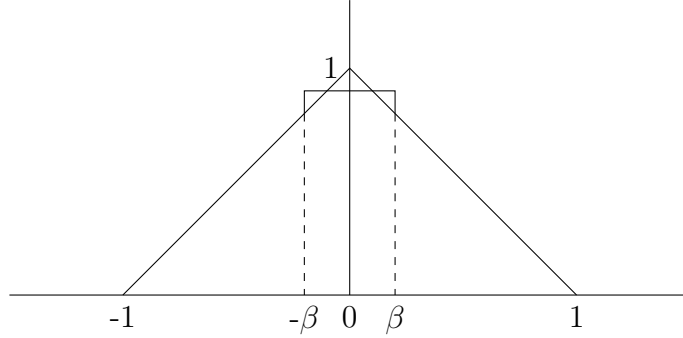


Figure 1: The probability density functions  $\rho_0$  and  $\rho_1$  from section 2.2

### 3.1 Entropic approximation

In this whole section, we assume that  $\Omega$  is an open bounded connected (not necessarily convex) subset of  $\mathbb{R}^d$  with Lipschitz boundary and denote by  $\mathcal{P}_{\text{ac}}(\Omega)$  the set of Borel probability measures on  $\Omega$  that are absolutely continuous with respect to the Lebesgue measure (and will use the same notation for  $\mu \in \mathcal{P}_{\text{ac}}(\Omega)$  both for the measure  $\mu$  and its density). Given  $\rho_0 \in \mathcal{P}_{\text{ac}}(\Omega)$  and  $\tau > 0$ , we consider one step of the TV-JKO scheme:

$$\inf_{\rho \in \mathcal{P}_{\text{ac}}(\Omega)} \left\{ \frac{1}{2\tau} W_2^2(\rho_0, \rho) + J(\rho) \right\}. \quad (3.1)$$

It is easy by the direct method of the calculus of variations to see that (3.1) has at least one solution, moreover  $J$  being convex and  $\rho \mapsto W_2^2(\rho, \rho_0)$  being strictly convex whenever  $\rho_0 \in \mathcal{P}_{\text{ac}}(\Omega)$  (see [26]), the minimizer is in fact unique, and in the sequel we denote it by  $\rho_1$ . Given  $h > 0$  we consider the following approximation of (3.1):

$$\inf_{\rho \in \mathcal{P}_{\text{ac}}(\Omega)} \left\{ \mathcal{F}_h(\rho) := \frac{1}{2\tau} W_2^2(\rho_0, \rho) + J(\rho) + h\mathcal{E}(\rho) \right\} \quad (3.2)$$

where

$$\mathcal{E}(\rho) := \int_{\Omega} \rho(x) \log(\rho(x)) dx.$$

It is easy to see that (3.2) admits a unique solution  $\rho_h$ . Moreover, since  $\Omega$  is bounded,  $\mathcal{E}$  is lower bounded, hence  $J(\rho_h)$  is bounded. Recalling that the embedding  $BV(\Omega) \subset L^p(\Omega)$  is compact for every  $p \in [1, \frac{d}{d-1})$ , one may therefore (up to extraction) assume that  $\rho_h$  converges as  $h \rightarrow 0$  a.e. and strongly in  $L^p(\Omega)$  for every  $p \in [1, \frac{d}{d-1})$  to some  $\rho_1$ , which, by a standard



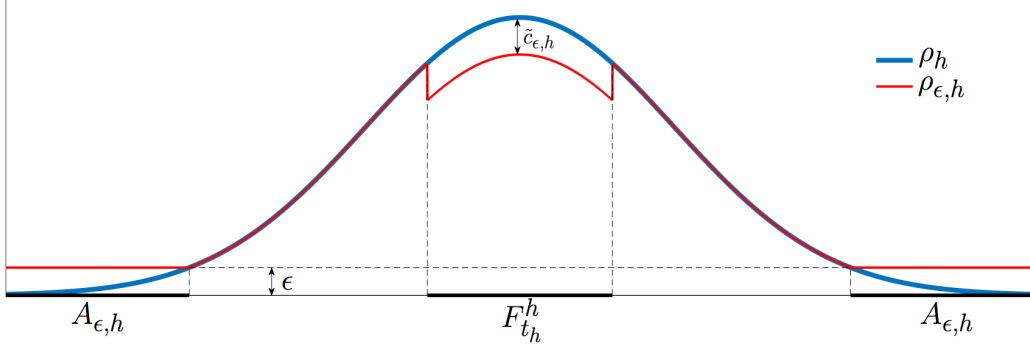


Figure 2: The perturbation  $\rho_{\varepsilon,h}$  from (3.5).

$\Gamma$ -convergence argument, is easily seen to be the solution of (3.1). The advantage of this regularization is that not only each  $\rho_h$  is bounded from below but also that  $h \log(\rho_h)$  is bounded from below uniformly in  $h$  (but not in  $\tau$  which is fixed throughout this section):

**Proposition 3.1.** *Up to passing to a subsequence, the family  $\beta_h := h \log(\rho_h)$  is uniformly bounded from below. Moreover,  $\beta_h$  is bounded in  $L^p(\Omega)$  for any  $p > 1$  and  $\max(0, \beta_h)$  converges strongly to 0 in  $L^p(\Omega)$  for any  $p > 1$ .*

*Proof.* Let  $t_h > 0$  be such that the set  $F_{t_h}^h := \{\rho_h > t_h\}$  has positive measure and finite perimeter (recall that  $\rho_h \in \text{BV}$ ). Let us assume that there is an  $\varepsilon \in (0, 1)$  such that

$$\varepsilon \leq \frac{t_h |F_{t_h}^h|}{2|\Omega|}, \quad (3.3)$$

and

$$|A_{\varepsilon,h}| > 0 \text{ with } A_{\varepsilon,h} := \{\rho_h \leq \varepsilon\}. \quad (3.4)$$

We aim to show that  $\varepsilon$  cannot be arbitrarily small. Define then  $\mu_{\varepsilon,h} := \max(\rho_h, \varepsilon)$  that is  $\varepsilon$  on  $A_{\varepsilon,h}$  and  $\rho_h$  elsewhere. Defining  $c_{\varepsilon,h} := \int_{\Omega} (\mu_{\varepsilon,h} - \rho_h)$  and observing that  $c_{\varepsilon,h} \leq \varepsilon |\Omega|$ , we see that (3.3) implies that  $c_{\varepsilon,h} \leq \frac{1}{2} t_h |F_{t_h}^h|$  and  $t_h \geq 2\varepsilon$  so that  $A_{\varepsilon,h}$  and  $F_{t_h}^h$  are disjoint. Finally, set

$$\rho_{\varepsilon,h} := \mu_{\varepsilon,h} - c_{\varepsilon,h} \frac{\chi_{F_{t_h}^h}}{|F_{t_h}^h|}. \quad (3.5)$$

See Figure 2, where we set  $\tilde{c}_{\varepsilon,h} := c_{\varepsilon,h} / |F_{t_h}^h|$ .

By construction  $\rho_{\varepsilon,h} \in \mathcal{P}(\Omega)$  hence  $0 \leq \mathcal{F}_h(\rho_{\varepsilon,h}) - \mathcal{F}_h(\rho_h)$ , in this difference we have four terms, namely

- the Wasserstein term, which, using the Kantorovich duality formula (2.1) and the fact that  $\Omega$  is bounded can be estimated in terms of  $\|\rho_{\varepsilon,h} - \rho_h\|_{L^1} = 2c_{\varepsilon,h}$ :

$$\frac{1}{2\tau}W_2^2(\rho_{\varepsilon,h}, \rho_0) - \frac{1}{2\tau}W_2^2(\rho_h, \rho_0) \leq \frac{C}{\tau}c_{\varepsilon,h}. \quad (3.6)$$

for a constant  $C$  that depends on  $\Omega$  but neither on  $\varepsilon$  nor  $h$ ,

- the TV term:  $J(\rho_{\varepsilon,h}) - J(\rho_h)$ : outside  $F_{t_h}^h$  we have replaced  $\rho_h$  by a 1-Lipschitz function of  $\rho_h$  which decreases the TV semi-norm, on  $F_{t_h}^h$  on the contrary we have created a jump of magnitude  $c_{\varepsilon,h}/|F_{t_h}^h|$  so

$$J(\rho_{\varepsilon,h}) - J(\rho_h) \leq c_{\varepsilon,h} \frac{\text{Per}(F_{t_h}^h)}{|F_{t_h}^h|} \quad (3.7)$$

where  $\text{Per}(F_{t_h}^h) = J(\chi_{F_{t_h}^h})$  denotes the perimeter of  $F_{t_h}^h$  (in  $\Omega$ ),

- the entropy variation on  $A_{\varepsilon,h}$ , on this set both  $\rho_{\varepsilon,h}$  and  $\rho_h$  are less than  $\varepsilon$  so that  $(1 + \log(t)) \leq (1 + \log(\varepsilon))$  whenever  $t \in [\rho_h, \rho_{\varepsilon,h}]$  which by the mean value theorem yields

$$\int_{A_{\varepsilon,h}} (\rho_{\varepsilon,h} \log(\rho_{\varepsilon,h}) - \rho_h \log(\rho_h)) \leq (1 + \log(\varepsilon))c_{\varepsilon,h} \quad (3.8)$$

- the last term is the entropy variation on  $F_{t_h}^h$ . It is convenient to split  $F_{t_h}^h$  into  $F_{t_h}^h \cap \{\rho_{\varepsilon,h} \geq \frac{1}{e}\}$  and  $F_{t_h}^h \cap \{\rho_{\varepsilon,h} < \frac{1}{e}\}$ . The entropy variation on the first part is easy to control. Indeed,  $t \mapsto t \log(t)$  is nondecreasing on  $[\frac{1}{e}, +\infty)$ . Since, on  $F_{t_h}^h \cap \{\rho_{\varepsilon,h} \geq \frac{1}{e}\}$ ,  $\rho_h \geq \rho_{\varepsilon,h} \geq \frac{1}{e}$ , we have  $(\rho_{\varepsilon,h} \log(\rho_{\varepsilon,h}) - \rho_h \log(\rho_h)) \leq 0$ . As for the second part, we observe that  $F_{t_h}^h \cap \{\rho_{\varepsilon,h} < \frac{1}{e}\} \subset \{\rho_h \leq \frac{1}{e} + \frac{t_h}{2}\}$ , so on this set, both  $\rho_{\varepsilon,h}$  and  $\rho_h$  remain in the interval  $[\frac{t_h}{2}, \frac{1}{e} + \frac{t_h}{2}]$ . We thus have

$$\int_{F_{t_h}^h} (\rho_{\varepsilon,h} \log(\rho_{\varepsilon,h}) - \rho_h \log(\rho_h)) \leq C_h(t_h)c_{\varepsilon,h}, \quad (3.9)$$

where

$$C_h(t_h) := \max \left\{ |1 + \log(t)| : \frac{t_h}{2} \leq t \leq \frac{1}{e} + \frac{t_h}{2} \right\}. \quad (3.10)$$

Putting together (3.6)-(3.7)-(3.8)-(3.9), we arrive at

$$0 \leq \left( \frac{C}{\tau} + \frac{\text{Per}(F_{t_h}^h)}{|F_{t_h}^h|} + hC_h(t_h) + h \log(\varepsilon) + h \right) c_{\varepsilon,h}$$

which for small enough  $\varepsilon$  is possible only when  $c_{\varepsilon,h} = 0$  i.e.  $|A_{\varepsilon,h}| = 0$ . More precisely, either we have the lower bound:

$$h \log(\rho_h) \geq -\frac{C}{\tau} - hC_h(t_h) - \frac{\text{Per}(F_{t_h}^h)}{|F_{t_h}^h|} - h \quad (3.11)$$

or (3.3) is impossible i.e.  $\rho_h \geq \frac{t_h |F_{t_h}^h|}{2|\Omega|}$ . To prove that  $\beta_h = h \log(\rho_h)$  is bounded from below uniformly in  $h$ , it is therefore enough to show that we can find a family  $t_h$ , bounded and bounded away from 0, such that  $|F_{t_h}^h|$  remains bounded away from 0, and  $\text{Per}(F_{t_h}^h)$  is uniformly bounded from above as  $h \rightarrow 0$ . First note that, since  $J(\rho_h)$  is bounded, there exists  $\rho$  such that  $\rho_h \rightarrow \rho$  in  $L^1$  and a.e. up to a subsequence, note also that  $\rho \in \text{BV}$  and  $\rho$  is a probability density. Setting  $F_t := \{\rho > t\}$ ,  $F_t^h := \{\rho_h > t\}$ , if  $s > t$ , since  $\rho_h$  converges a.e. to  $\rho$ , we have a.e.  $\liminf_h \chi_{F_t^h} \geq \chi_{F_s}$ . It then follows from Fatou's Lemma that when  $s > t$ ,  $\liminf_h |F_t^h| \geq |F_s|$ , hence choosing  $0 < \beta_1 < \beta_2 < \beta$  so that  $|F_{\beta}| > 0$ , we deduce that there exists  $h_0 > 0$  and  $c_1 > 0$  such that for all  $t \in [\beta_1, \beta_2]$  and all  $h \in (0, h_0]$ , we have  $c_1 \leq |F_t^h| \leq |\Omega|$ . For an upper bound on perimeters, we observe that since  $J(\rho_h) \leq C$ , thanks to the co-area formula, we have

$$\int_{\beta_1}^{\beta_2} \text{Per}(F_t^h) dt \leq J(\rho_h) \leq C.$$

So, there exists  $t_h \in [\beta_1, \beta_2]$  such that  $\text{Per}(F_{t_h}^h) \leq C/(\beta_2 - \beta_1)$ .

Finally, since  $\rho_h$  converges in  $L^1$ , we may assume that, up to a subsequence,  $\rho_h \leq \phi$  for some  $\phi \in L^1$  (see Theorem IV.9 in [6]). Then, by Dominated convergence and since  $\log(\max(\phi, 1)) \in L^p(\Omega)$  for every  $p > 1$ , we have that  $\log(\max(\rho_h, 1))$  converges a.e. and in  $L^p$ , in particular this implies that  $\max(0, \beta_h)$  converges to 0 strongly in  $L^p(\Omega)$ , and we have just seen that  $\min(0, \beta_h)$  is bounded in  $L^\infty(\Omega)$ . □

Let us also recall some well-known facts (see [9]) about the total variation functional  $J$  viewed as a convex l.s.c. and one-homogeneous functional on  $L^{\frac{d}{d-1}}(\Omega)$ . Define

$$\Gamma_d := \left\{ \xi \in L^d(\Omega) : \exists z \in L^\infty(\Omega, \mathbb{R}^d), \|z\|_{L^\infty} \leq 1, \text{div}(z) = \xi, z \cdot \nu = 0 \text{ on } \partial\Omega \right\} \quad (3.12)$$

where  $\text{div}(z) = \xi$ ,  $z \cdot \nu = 0$  on  $\partial\Omega$  are to be understood in the weak sense

$$\int_{\Omega} \xi u = - \int_{\Omega} z \cdot \nabla u, \quad \forall u \in C^1(\overline{\Omega}).$$

Note that  $\Gamma_d$  is closed and convex in  $L^d(\Omega)$  and  $J$  is its support function:

$$J(\mu) = \sup_{\xi \in \Gamma_d} \int_{\Omega} \xi \mu, \quad \forall \mu \in L^{\frac{d}{d-1}}(\Omega). \quad (3.13)$$

As for the Wasserstein term, recalling Kantorovich dual formulation (2.1), the derivative of the Wasserstein term  $\rho \mapsto W_2^2(\rho_0, \rho)$  term will be expressed in terms of a Kantorovich potential between  $\rho$  and  $\rho_0$ .

We then have the following characterization for  $\rho_h$ :

**Proposition 3.2.** *There exists  $z_h \in L^\infty(\Omega, \mathbb{R}^d)$  such that  $\operatorname{div}(z_h) \in L^p(\Omega)$  for every  $p \in [1, +\infty)$ ,  $\|z_h\|_{L^\infty} \leq 1$ ,  $z_h \cdot \nu = 0$  on  $\partial\Omega$ ,  $J(\rho_h) = \int_{\Omega} \operatorname{div}(z_h) \rho_h$  and*

$$\frac{\varphi_h}{\tau} + \operatorname{div}(z_h) + h \log(\rho_h) = 0, \quad \text{a.e. in } \Omega \quad (3.14)$$

where  $\varphi_h$  is the Kantorovich potential between  $\rho_h$  and  $\rho_0$ .

*Proof.* Let  $\mu \in L^\infty(\Omega) \cap \operatorname{BV}(\Omega)$  such that  $\int_{\Omega} \mu = 0$ . Thanks to Proposition 3.1, we know that  $\rho_h$  is bounded away from 0 hence for small enough  $t > 0$ ,  $\rho_h + t\mu$  is positive hence a probability density. Also, as a consequence of Theorem 1.52 in [26], we have that

$$\lim_{t \rightarrow 0^+} \frac{1}{2t} [W_2^2(\rho_0, \rho_h + t\mu) - W_2^2(\rho_0, \rho_h)] = \int_{\Omega} \varphi_h \mu \quad (3.15)$$

where  $\varphi_h$  is the (unique up to an additive constant) Kantorovich potential between  $\rho_h$  and  $\rho_0$ , in particular  $\varphi_h$  is Lipschitz and semi concave ( $D^2\varphi_h \leq \operatorname{id}$  in the sense of measures and  $\operatorname{id} - \nabla\varphi_h$  is the optimal transport between  $\rho_h$  and  $\rho_1$ ). By the optimality of  $\rho_h$  and the fact that  $J$  is a semi-norm, we get

$$J(\mu) \geq J(\rho_h + \mu) - J(\rho_h) \geq \lim_{t \rightarrow 0^+} t^{-1} (J(\rho_h + t\mu) - J(\rho_h)) \geq \int_{\Omega} \xi_h \mu, \quad (3.16)$$

where

$$\xi_h := -\frac{\varphi_h}{\tau} - h \log(\rho_h).$$

Since  $\varphi_h$  is defined up to an additive constant, we may chose it in such a way that  $\xi_h$  has zero mean, doing so, (3.16) holds for any  $\mu \in L^\infty(\Omega) \cap \operatorname{BV}(\Omega)$  (not necessarily with zero mean). Being Lipschitz,  $\varphi_h$  is bounded, also observe that  $h(\log(\rho_h))_+ = h \log(\max(1, \rho_h))$  is in  $L^p(\Omega)$  for every  $p \in [1, +\infty)$  since  $\rho_h \in L^{\frac{d}{d-1}}(\Omega)$  and  $h \log(\rho_h)_- = -h \log(\min(1, \rho_h))$  is  $L^\infty(\Omega)$  thanks to Proposition 3.1, hence we have  $\xi_h \in L^p(\Omega)$  for every  $p \in [1, +\infty)$ .

By approximation and observing that  $\xi_h \in L^d(\Omega)$ , (3.16) extends to all  $\mu \in L^{\frac{d}{d-1}}(\Omega)$ . In particular, we have

$$\sup_{\xi \in \Gamma_d} \int_{\Omega} \xi \mu \geq \int_{\Omega} \xi_h \mu$$

but since  $\Gamma_d$  is convex and closed in  $L^d(\Omega)$ , it follows from Hahn-Banach's separation theorem that  $\xi_h \in \Gamma_d$ . Finally, getting back to (3.16) (without the zero mean restriction on  $\mu$ ) and taking  $\mu = -\rho_h$  gives  $J(\rho_h) \leq \int_{\Omega} \xi_h \rho_h$ , and we then deduce that this should be an equality.  $\square$

### 3.2 Euler-Lagrange equation

We are now in position to rigorously establish the Euler-Lagrange equation for (3.1):

**Theorem 3.3.** *If  $\rho_1$  solves (3.1), there exists  $\varphi$  a Kantorovich potential between  $\rho_1$  and  $\rho_0$  (in particular  $-\nabla\varphi$  is the optimal transport between  $\rho_1$  and  $\rho_0$ ),  $\beta \in L^\infty(\Omega)$ ,  $\beta \geq 0$  and  $z \in L^\infty(\Omega, \mathbb{R}^d)$  such that*

$$\frac{\varphi}{\tau} + \operatorname{div}(z) = \beta, \quad z \cdot \nu = 0 \text{ on } \partial\Omega, \quad (3.17)$$

and

$$\beta \rho_1 = 0, \quad \|z\|_{L^\infty} \leq 1, \quad J(\rho_1) = \int_{\Omega} \operatorname{div}(z) \rho_1. \quad (3.18)$$

*Remark 3.4.* It is not difficult (since (3.1) is a convex problem) to check that (3.17)-(3.18) are also sufficient optimality conditions. The main point here is that the right hand side  $\beta$  in (3.17) which is a multiplier associated with the nonnegativity constraint is better than a measure, it is actually an  $L^\infty$  function.

*Proof.* As in section 3.1, we denote by  $\rho_h$  the solution of the entropic approximation (3.2). Up to passing to a subsequence (not explicitly written), we may assume that  $\rho_h$  converges a.e. and strongly in  $L^p(\Omega)$  (for any  $p \in [1, \frac{d}{d-1})$ ) to  $\rho_1$  (the solution of (3.1), again by a standard  $\Gamma$ -convergence argument). We then rewrite the Euler-Lagrange equation from Proposition 3.2 as

$$\frac{\varphi_h}{\tau} + \operatorname{div}(z_h) + \beta_h^+ = \beta_h^-, \quad (3.19)$$

where  $\beta_h^+ := h \log(\max(\rho_h, 1))$ ,  $\beta_h^- := -h \log(\min(\rho_h, 1))$ , and

$$\|z_h\|_{L^\infty} \leq 1, \quad z_h \cdot \nu = 0 \text{ on } \partial\Omega \text{ and } J(\rho_h) = \int_{\Omega} \operatorname{div}(z_h) \rho_h. \quad (3.20)$$

It follows from Proposition 3.1 that  $\beta_h^+$  converges to 0 strongly in any  $L^p$ ,  $p \in [1, +\infty)$  and that  $\beta_h^-$  is bounded in  $L^\infty$ . Up to subsequences, we may therefore assume that  $z_h$  and  $\beta_h^-$  weakly-\* converge in  $L^\infty$  respectively to some  $z$  and  $\beta$  with  $\|z\|_{L^\infty} \leq 1$ ,  $z \cdot \nu = 0$  on  $\partial\Omega$  and  $\beta \geq 0$ . As for the Kantorovich potentials  $\varphi_h$ , since the transport map  $(\text{id} - \nabla\varphi_h)$  a.e. takes values in  $\Omega$  we have  $\|\nabla\varphi_h\|_{L^\infty} \leq \text{diam}(\Omega)$ , hence  $\varphi_h$  is an equi-Lipschitz family because  $\Omega$  is bounded. Moreover  $\int_\Omega \varphi_h = \tau \int_\Omega (\beta_h^- - \beta_h^+)$  which remains bounded, hence we may assume that  $\varphi_h$  converges uniformly to some potential  $\varphi$  and it is well-known (see [26]) that  $\varphi$  is a Kantorovich potential between  $\rho_1$  and  $\rho_0$ . Letting  $h$  tend to 0 gives (3.17).

Since  $\rho_h$  converges strongly in  $L^1$  to  $\rho_1$  and  $\beta_h^-$  converges weakly-\* to  $\beta$  in  $L^\infty$  we have

$$\int_\Omega \rho_1 \beta = \lim_h \int_\Omega \rho_h \beta_h^- = \lim_h h \int_\Omega \rho_h |\log(\min(1, \rho_h))| = 0,$$

hence  $\beta \rho_1 = 0$ . Thanks to (3.13), we obviously have  $J(\rho_1) \geq \int_\Omega \text{div}(z) \rho_1$  (since  $\text{div}(z) \in L^\infty$ ,  $\text{div}(z) \in \Gamma_d$ ), for the converse inequality, it is enough to observe that

$$J(\rho_1) \leq \liminf_h J(\rho_h) = \liminf_h \int_\Omega \text{div}(z_h) \rho_h$$

and that  $\text{div}(z_h) = -\frac{\varphi_h}{\tau} - \beta_h^+ + \beta_h^-$  converges to  $\text{div}(z)$  weakly in  $L^q$  for every  $q \in [1, +\infty)$ . Since  $\rho_h$  converges strongly to  $\rho_1$  in  $L^q$  when  $q \in [1, \frac{d}{d-1})$  we deduce that  $J(\rho_1) = \int_\Omega \text{div}(z) \rho_1$  which completes the proof of (3.18).  $\square$

A first consequence of the high integrability of  $\text{div}(z)$  is that one can give a meaning to  $z \cdot \nabla u$  for any  $u \in \text{BV}(\Omega)$ . Indeed, if  $q \in [\frac{d}{d-1}, +\infty]$  and  $q'$  denotes its conjugate exponent, following Anzellotti [2], if  $u \in \text{BV}(\Omega) \cap L^q(\Omega)$  and  $\sigma \in L^\infty(\Omega, \mathbb{R}^d)$  is such that  $\text{div}(\sigma) \in L^{q'}(\Omega)$ , one can define the distribution  $\sigma \cdot Du$  by

$$\langle \sigma \cdot Du, v \rangle = - \int_\Omega \text{div}(\sigma) uv - \int_\Omega u \sigma \cdot \nabla v, \quad \forall v \in C_c^1(\Omega).$$

Then  $\sigma \cdot Du$  is a Radon measure which satisfies  $|\sigma \cdot Du| \leq \|\sigma\|_{L^\infty} |Du|$  (in the sense of measures) hence is absolutely continuous with respect to  $|Du|$ . Moreover one can also define a weak notion of normal trace of  $\sigma$ ,  $\sigma \cdot \nu \in L^\infty(\partial\Omega)$  such that the following integration by parts formula holds

$$\int_\Omega \sigma \cdot Du = - \int_\Omega \text{div}(\sigma) u + \int_{\partial\Omega} u(\sigma \cdot \nu).$$

We refer to [2] for proofs. These considerations of course apply to  $\sigma = z$  and  $u = \rho_1 \in \text{BV}(\Omega)$  and in particular enable one to see  $z \cdot D\rho_1$  as a measure and to interpret the optimality condition  $J(\rho_1) = \int_{\Omega} \text{div}(z)\rho_1$  as  $|D\rho_1| = -z \cdot D\rho_1$  in the sense of measures. Finally, the fact that  $\text{div}(z) \in L^\infty$  in Theorem 3.3 and the theory of variational mean curvature (see Tamanini [27], Massari [20, 21], Theorem 3.6 of Gonzalez and Massari [19]) allows for conclusions about the regularity of the level sets,  $F_t = \{\rho_1 > t\}$  of  $\rho_1$ , the solution of (3.1), we do not elaborate this regularity (which, anyway, only holds for fixed time step  $\tau > 0$ ) further here.

## 4 Maximum and minimum principles for JKO steps

Throughout this section, we further assume that  $\Omega$  is a convex open bounded subset of  $\mathbb{R}^d$ , our aim is to establish bounds on the TV-JKO iterates given by (3.1). Since, the TV-JKO scheme aims at minimizing total variation at the fastest rate in the Wasserstein metric, it is natural to wonder whether when the initial condition is bounded from above and from below then the JKO-iterates remain so (with the same bounds). We shall answer affirmatively for the upper bound (maximum principle), as for the propagation of the lower bound (minimum principle), we have been able to prove it only in special cases (dimension one and radially symmetric setting).

### 4.1 Convexity along generalized geodesics

Our aim is to deduce some bounds on  $\rho_1$  from bounds on  $\rho_0$ . To do so, we shall combine some convexity arguments and a remarkable BV estimate due to De Philippis et al. [11]. First we recall the notion of generalized geodesic from Ambrosio, Gigli and Savaré [1]. Given  $\bar{\mu}$ ,  $\mu_0$  and  $\mu_1$  in  $\mathcal{P}_{\text{ac}}(\Omega)$ , and denoting by  $T_0$  (respectively  $T_1$ ) the optimal transport (Brenier) map between  $\bar{\mu}$  and  $\mu_0$  (respectively  $\mu_1$ ), the *generalized geodesic with base  $\bar{\mu}$*  joining  $\mu_0$  to  $\mu_1$  is by definition the curve of measures:

$$\mu_t := ((1-t)T_0 + tT_1)_{\#}\bar{\mu}, \quad t \in [0, 1]. \quad (4.1)$$

A key property of these curves introduced in [1] is the strong convexity of the squared distance estimate:

$$W_2^2(\bar{\mu}, \mu_t) \leq (1-t)W_2^2(\bar{\mu}, \mu_0) + tW_2^2(\bar{\mu}, \mu_1) - t(1-t)W_2^2(\mu_0, \mu_1). \quad (4.2)$$

It is well-known that if  $G : \mathbb{R}_+ \rightarrow \mathbb{R} \cup \{+\infty\}$  is a proper convex lower semi-continuous (l.s.c.) internal energy density, bounded from below such that  $G(0) = 0$  and which satisfies McCann's condition (see [23])

$$\lambda \in \mathbb{R}_+ \rightarrow \lambda^d G(\lambda^{-d}) \text{ is convex nonincreasing} \quad (4.3)$$

then defining the generalized geodesic curve  $(\mu_t)_{t \in [0,1]}$  by (4.1), one has

$$\int_{\Omega} G(\mu_t(x)) dx \leq (1-t) \int_{\Omega} G(\mu_0(x)) dx + t \int_{\Omega} G(\mu_1(x)) dx. \quad (4.4)$$

In particular  $L^p$  and uniform bounds are stable along generalized geodesics:

$$\|\mu_t\|_{L^p}^p \leq (1-t)\|\mu_0\|_{L^p}^p + t\|\mu_1\|_{L^p}^p, \quad \|\mu_t\|_{L^\infty} \leq \max(\|\mu_0\|_{L^\infty}, \|\mu_1\|_{L^\infty}), \quad (4.5)$$

and

$$\int_{\Omega} \mu_t(x) \log(\mu_t(x)) dx \leq (1-t) \int_{\Omega} \mu_0(x) \log(\mu_0(x)) dx + t \int_{\Omega} \mu_1(x) \log(\mu_1(x)) dx \quad (4.6)$$

An immediate consequence of (4.2) (see chapter 4 of [1] for general contraction estimates) is the following

**Lemma 4.1.** *Let  $K$  be a nonempty subset of  $\mathcal{P}_{ac}(\Omega)$ , let  $\mu_0 \in K$ ,  $\mu_1 \in \mathcal{P}_{ac}(\Omega)$ , if  $\hat{\mu}_1 \in \operatorname{argmin}_{\mu \in K} W_2^2(\mu_1, \mu)$  is a Wasserstein projection of  $\mu_1$  onto  $K$ , and if the generalized geodesic with base  $\mu_1$  joining  $\mu_0$  to  $\hat{\mu}_1$  remains in  $K$  then*

$$W_2^2(\mu_0, \hat{\mu}_1) \leq W_2^2(\mu_0, \mu_1) - W_2^2(\mu_1, \hat{\mu}_1). \quad (4.7)$$

*Proof.* Since  $\mu_t \in K$  we have  $W_2^2(\mu_1, \hat{\mu}_1) \leq W_2^2(\mu_1, \mu_t)$ , applying (4.2) to the generalized geodesics with base  $\mu_1$  joining  $\mu_0$  to  $\hat{\mu}_1$  we thus get

$$(1-t)W_2^2(\mu_1, \hat{\mu}_1) \leq (1-t)W_2^2(\mu_1, \mu_0) - t(1-t)W_2^2(\mu_0, \hat{\mu}_1),$$

dividing by  $(1-t)$  and then taking  $t = 1$  therefore gives the desired result.  $\square$

The other result we shall use to derive bounds is a BV estimate of De Philippis et al. [11], which states that, given,  $\mu \in \mathcal{P}_{ac}(\Omega) \cap \operatorname{BV}(\Omega)$ , and  $G : \mathbb{R}_+ \rightarrow \mathbb{R} \cup \{+\infty\}$ , proper convex l.s.c., the solution of

$$\inf_{\rho \in \mathcal{P}_{ac}(\Omega)} \left\{ \frac{1}{2} W_2^2(\mu, \rho) + \int_{\Omega} G(\rho(x)) dx \right\} \quad (4.8)$$

is BV with the bound

$$J(\rho) \leq J(\mu). \quad (4.9)$$



Taking in particular,

$$G(\rho) := \begin{cases} 0 & \text{if } \rho \leq M, \\ +\infty & \text{otherwise,} \end{cases}$$

this implies that the Wasserstein projection of  $\mu$  onto the set defined by the constraint  $\rho \leq M$  has a smaller total variation than  $\mu$ .

## 4.2 Maximum principle

**Theorem 4.2.** *Let  $\rho_0 \in \mathcal{P}_{\text{ac}}(\Omega) \cap L^\infty(\Omega)$  and let  $\rho_1$  be the solution of (3.1), then  $\rho_1 \in L^\infty(\Omega)$  with*

$$\|\rho_1\|_{L^\infty(\Omega)} \leq \|\rho_0\|_{L^\infty(\Omega)}. \quad (4.10)$$

*Proof.* Thanks to (4.5) the set  $K := \{\rho \in \mathcal{P}_{\text{ac}}(\Omega) : \rho \leq \|\rho_0\|_{L^\infty(\Omega)} \text{ a.e.}\}$  has the property that the generalized geodesics (with any base) joining two of its points remains in  $K$ . Let then  $\hat{\rho}_1$  be the  $W_2$  projection of  $\rho_1$  onto  $K$  i.e. the solution of  $\inf_{\rho \in K} W_2^2(\rho_1, \rho)$ . Thanks to Lemma 4.1 we have  $W_2^2(\rho_0, \hat{\rho}_1) \leq W_2^2(\rho_0, \rho_1) - W_2^2(\rho_1, \hat{\rho}_1)$  and thanks to Theorem 1.1 of [11],  $J(\hat{\rho}_1) \leq J(\rho_1)$ . The optimality of  $\rho_1$  for (3.1) therefore implies  $W_2(\rho_1, \hat{\rho}_1) = 0$  i.e.  $\rho_1 \leq \|\rho_0\|_{L^\infty(\Omega)}$ . □

*Remark 4.3.* In section 3, we have used an approximation of (3.1) with an additional small entropy term, the same bound as in Theorem 4.2 will remain valid in this case. Indeed, consider a proper convex l.s.c. and bounded from below internal energy density  $G$  and consider given  $h \geq 0$ , the variant of (3.1)

$$\inf_{\rho \in \mathcal{P}_{\text{ac}}(\Omega)} \left\{ \frac{1}{2\tau} W_2^2(\rho_0, \rho) + J(\rho) + h \int_{\Omega} G(\rho(x)) dx \right\}. \quad (4.11)$$

Then we claim that the solution  $\rho_h$  still satisfies  $\rho_h \leq \|\rho_0\|_{L^\infty(\Omega)}$ . Indeed we have seen in the previous proof that the Wasserstein projection  $\hat{\rho}_h$  of  $\rho_h$  onto the constraint  $\rho \leq \|\rho_0\|_{L^\infty(\Omega)}$  both diminishes  $J$  and the Wasserstein distance to  $\rho_0$ . It turns out that it also diminishes the internal energy. Indeed, thanks to Proposition 5.2 of [11], there is a measurable set  $A$  such that  $\hat{\rho}_h = \chi_A \rho_h + \chi_{\Omega \setminus A} \|\rho_0\|_{L^\infty}$ , it thus follows that  $|\Omega \setminus A| \|\rho_0\|_{L^\infty} = \int_{\Omega \setminus A} \rho_h$ . So, from the convexity of  $G$  and Jensen's inequality,

$$\int G(\hat{\rho}_h) = \int_A G(\rho_h) + |\Omega \setminus A| G\left(|\Omega \setminus A|^{-1} \int_{\Omega \setminus A} \rho_h\right) \leq \int G(\rho_h),$$

thus yielding the same conclusion as above.

### 4.3 Minimum principle in special cases

In dimension one, it turns out that we can obtain bounds from below by the same convexity arguments as for the maximum principle of Theorem 4.2:

**Proposition 4.4.** *Assume that  $d = 1$ , that  $\Omega$  is a bounded interval and that  $\rho_0 \geq \alpha > 0$  a.e. on  $\Omega$  then the solution  $\rho_1$  of (3.1) also satisfies  $\rho_1 \geq \alpha > 0$  a.e. on  $\Omega$ .*

*Proof.* The proof is similar to that of Theorem 4.2 but using the Wasserstein projection on the set  $K := \{\rho \in \mathcal{P}_{ac}(\Omega) : \rho \geq \alpha\}$ , the only thing to check to be able to use Lemma 4.1 is that for any basepoint  $\bar{\mu}$  and any  $\mu_0$  and  $\mu_1$  in  $K$ , the generalized geodesic with base point  $\bar{\mu}$  joining  $\mu_0$  to  $\mu_1$  remains in  $K$ . The optimal transport maps  $T_0$  and  $T_1$  from  $\bar{\mu}$  to  $\mu_0$  and  $\mu_1$  respectively are nondecreasing and continuous and setting  $T_t := (1 - t)T_0 + tT_1$ , one has

$$\bar{\mu} = \mu_t(T_t)T'_t = \mu_0(T_0)T'_0 = \mu_1(T_1)T'_1 = (1 - t)\mu_0(T_0)T'_0 + t\mu_1(T_1)T'_1 \geq \alpha T'_t$$

which is easily seen to imply that  $\mu_t \geq \alpha$  a.e..  $\square$

As a consequence of the previous minimum principle, integrating the Euler-Lagrange equation one can deduce higher regularity for the dual variable  $z$ :

**Corollary 4.5.** *Assume that  $d = 1$  and  $\Omega$  is a bounded interval. If  $\rho_1$  solves (3.1) and  $z$  is as in Theorem 3.3 then  $z \in W_0^{1,\infty}(\Omega)$ . If in addition  $\rho_0 \geq \alpha > 0$  a.e. on  $\Omega$ , then  $z \in W^{3,\infty}(\Omega)$ .*

*Proof.* The first claim is obvious because both  $\varphi$  and  $\beta$  ( $\varphi$ ,  $\beta$  and  $z$  are as in Theorem 3.3) are bounded hence so is  $z'$ . As for the second one when  $\rho_0 \geq \alpha > 0$ , thanks to Proposition 4.4, we also have  $\rho_1 \geq \alpha$  hence  $\beta = 0$  in (3.17) and in this case  $\text{div}(z) = z' = -\frac{\varphi}{\tau}$  is Lipschitz i.e.  $z \in W^{2,\infty}$ . One can actually go one step further because  $x - \varphi'(x) = T(x)$  where  $T$  is the optimal (monotone) transport between  $\rho_1$  and  $\rho_0$ . This map is explicit in terms of the cumulative distribution function of  $\rho_1$ ,  $F_1$ , and  $F_0^{-1}$  the inverse of  $F_0$ , the cumulative distribution function of  $\rho_0$ , namely  $T = F_0^{-1} \circ F_1$ . But  $F_1$  is Lipschitz since its derivative is  $\rho_1$  which is BV hence bounded and  $F_0^{-1}$  is Lipschitz as well since  $\rho_0 \geq \alpha > 0$ . This gives that  $\varphi \in W^{2,\infty}$  hence  $z \in W^{3,\infty}$ .  $\square$

The proof of Proposition 4.4 unfortunately does not generalize to higher dimensions, because densities which are bounded from below by  $\alpha$  are not stable by generalized geodesics. In the radially symmetric case, we can use the Euler-Lagrange equation to derive a minimum principle. We believe that JKO steps preserve lower bounds in more general situations but have not been able to prove it.

**Proposition 4.6.** *Assume that  $\Omega = B(0, R)$  is the ball centered at 0 of radius  $R > 0$  in  $\mathbb{R}^d$ , and that  $\rho_0$  is radially symmetric with  $\rho_0 \geq \alpha > 0$  a.e. on  $\Omega$  then the solution  $\rho_1$  of (3.1) also satisfies  $\rho_1 \geq \alpha > 0$  a.e. on  $\Omega$ .*

*Proof.* Let us write  $\rho_0(x) = \tilde{\rho}_0(r)$  with  $r = |x| \in [0, R]$ , since (3.1) is invariant by rotation and strictly convex, it is easy to see that its unique solution  $\rho_1$  is also radially symmetric, let us write it as  $\rho_1(x) = \tilde{\rho}_1(r)$ . Denoting by  $c_d$  the  $(d-1)$ -Hausdorff measure of the unit sphere  $S^{d-1}$ , and setting  $\tilde{\mu}_0 := c_d r^{d-1} \tilde{\rho}_0$ ,  $\tilde{\mu}_1 := c_d r^{d-1} \tilde{\rho}_1$ , observe that  $\tilde{\rho}_1$  is the minimizer of the one-dimensional convex functional

$$\mathcal{F}^{\text{rad}}(\tilde{\rho}) := \frac{1}{2\tau} W_2^2(\tilde{\mu}_0, c_d r^{d-1} \tilde{\rho}) + c_d \int_0^R r^{d-1} |D\tilde{\rho}|$$

among nonnegative densities  $\tilde{\rho}$  on  $(0, R)$  such that  $c_d \int_0^R r^{d-1} \tilde{\rho} = 1$  and  $r^{d-1} D\tilde{\rho}$  is a bounded Radon measure on  $(0, R)$ . Arguing as in the proof of Theorem 3.3, the minimizer  $\tilde{\rho}_1$  is characterized by the Euler-Lagrange equation

$$(\tilde{z} r^{d-1})' + \frac{\tilde{\varphi}}{\tau} r^{d-1} = \tilde{\beta} \geq 0, \quad \tilde{\beta} \in L^\infty(0, R), \quad \tilde{\beta} \tilde{\rho}_1 = 0, \quad (4.12)$$

where  $\tilde{\varphi}$  is a Kantorovich potential between  $\tilde{\mu}_1$  and  $\tilde{\mu}_0$  and  $\tilde{z} \in L^\infty(0, R)$  is such that

$$|\tilde{z}| \leq 1 \text{ a.e. and } \int_0^R r^{d-1} |D\tilde{\rho}_1| = \int_0^R (\tilde{z} r^{d-1})' \tilde{\rho}_1. \quad (4.13)$$

Note that (4.12) implies that  $r^{d-1} \tilde{z}$  is Lipschitz so that  $\tilde{z}$  is locally Lipschitz and

$$\int_0^R r^{d-1} |D\tilde{\rho}_1| = - \int_0^R r^{d-1} \tilde{z} D\tilde{\rho}_1 \quad (4.14)$$

Since  $\tilde{\rho}_1 \in \text{BV}_{\text{loc}}(0, R)$ , we can perform a Hahn-Jordan decomposition of  $D\tilde{\rho}_1$ :

$$D\tilde{\rho}_1 = \nu^+ - \nu^-, \quad \nu^+ \geq 0, \quad \nu^- \geq 0, \quad \nu^+ \perp \nu^-, \quad (4.15)$$

and set

$$A := \text{spt}(|D\tilde{\rho}_1|) = A^+ \cup A^- \text{ with } A^+ := \text{spt}(\nu^+), \quad A^- := \text{spt}(\nu^-). \quad (4.16)$$

Next, we observe that, using (4.14), we have  $|D\tilde{\rho}_1| = \nu^+ + \nu^- = -\tilde{z}(\nu^+ - \nu^-)$ , we thus deduce that  $\tilde{z} = -1 = \min \tilde{z}$   $\nu^+$ -a.e and since  $\tilde{z}$  is continuous we actually have  $\tilde{z} = -1$  on  $A^+ = \text{spt}(\nu^+)$ . In a similar way,  $\tilde{z} = 1 = \max \tilde{z}$  on  $A^- := \text{spt}(\nu^-)$ .

Now let us show that  $\tilde{\rho}_1 \geq \alpha$ . Assume, by contradiction, that the set where  $\tilde{\rho}_1 < \alpha$  has positive measure in  $(0, R)$ , and let  $r_0 \in (0, R)$  be a continuity point of  $\tilde{\rho}_1$  such that  $\tilde{\rho}_1(r_0) < \alpha$ , define then

$$\begin{aligned} a_- &:= \inf\{r \in (0, r_0) : \tilde{\rho}_1 \leq \alpha \text{ on } [r, r_0]\}, \\ a_+ &:= \sup\{r \in (r_0, R) : \tilde{\rho}_1 \leq \alpha \text{ on } [r_0, r]\}. \end{aligned}$$

We then have  $0 \leq a_- < a_+ \leq R$ . Let us assume that  $a_- > 0$ , we claim then that  $a_- \in A^-$  since otherwise,  $\tilde{\rho}_1$  would be nondecreasing in a neighbourhood of  $a_-$  which would imply  $\tilde{\rho}_1(a_- - \varepsilon) \leq \alpha$  for small  $\varepsilon > 0$ , contradicting the definition of  $a_-$ , we thus have  $\tilde{z}(a_-) = 1$ . Since  $\tilde{\rho}_1$  is BV in a neighbourhood of  $a_-$ , it has a right and a left limit at  $a_-$ , again by minimality of  $a_-$ , the left limit of  $\tilde{\rho}_1$  at  $a_-$  cannot be strictly smaller than  $\alpha$ , so there is an  $\varepsilon > 0$  such that  $\tilde{\rho}_1 > 0$  on  $I_- := [a_- - \varepsilon, a_-)$ . Hence on  $I_-$ , (4.12) becomes

$$\tilde{z}' + \frac{d-1}{r}\tilde{z} + \frac{\tilde{\varphi}}{\tau} = 0, \quad (4.17)$$

moreover, on  $I_-$ ,  $\tilde{\varphi}$  is actually of class  $C^1$  with  $\tilde{\varphi}'(r) = r - \tilde{T}(r)$  where  $\tilde{T}$  is the (continuous) optimal transport between  $\tilde{\mu}_1$  and  $\tilde{\mu}_0$  obtained by the relation  $F_{\tilde{\mu}_0} \circ \tilde{T} = F_{\tilde{\mu}_1}$  (where  $F_{\tilde{\mu}_i}$  is the cumulative distribution function of  $\tilde{\mu}_i$  for  $i = 0, 1$ ). One can therefore differentiate (4.17) on  $I_-$  so as to obtain

$$\tilde{z}'' + \frac{d-1}{r}\tilde{z}' - \frac{(d-1)}{r^2}\tilde{z}(r) + \frac{r - \tilde{T}(r)}{\tau} = 0, \forall r \in I_-. \quad (4.18)$$

Since  $\tilde{z}$  is maximal at  $a_-$ , we first have

$$\lim_{\delta \rightarrow 0^+} \delta^{-1}[\tilde{z}(a_-) - \tilde{z}(a_- - \delta)] = -\frac{(d-1)\tilde{z}(a_-)}{a_-} - \frac{\tilde{\varphi}(a_-)}{\tau} \geq 0$$

but recalling (4.12) we also have

$$\begin{aligned} 0 &\geq \limsup_{\delta \rightarrow 0^+} \delta^{-1}[\tilde{z}(a_- + \delta) - \tilde{z}(a_-)] \\ &\geq \lim_{\delta \rightarrow 0^+} \delta^{-1} \int_{a_-}^{a_- + \delta} [-(d-1)s^{-1}\tilde{z}(s) - \tau^{-1}\tilde{\varphi}(s)] ds \\ &= -\frac{(d-1)\tilde{z}(a_-)}{a_-} - \frac{\tilde{\varphi}(a_-)}{\tau} \end{aligned}$$

which shows that  $\tilde{z}$  is differentiable at  $a_-$  with  $\tilde{z}'(a_-) = 0$ , this enables us to deduce that  $\tilde{z}''(a_-) := \lim_{\delta \rightarrow 0^+} \tilde{z}''(a_- - \delta) \leq 0$ , with (4.18) this gives

$$\tilde{T}(a_-) - a_- = \tau \left( \tilde{z}''(a_-) - \frac{(d-1)}{a_-^2} \right) \leq 0.$$

If  $a_- = 0$ , since  $\tilde{T}(0) = 0$ , the same conclusion is reached with an equality. In a similar way, we obtain  $\tilde{T}(a_+) \geq a_+$  (again with an equality in case  $a_+ = R$ ). Using the fact that  $\tilde{\rho}_1 \leq \alpha$  on  $(a_-, a_+)$  (with strict inequality in a neighbourhood of  $r_0$ ) together with  $F_{\tilde{\mu}_0} \circ \tilde{T} = F_{\tilde{\mu}_1}$  and  $\tilde{\rho}_0 \geq \alpha$ , we get

$$\begin{aligned} \alpha c_d \frac{(a_+^d - a_-^d)}{d} &> F_{\tilde{\mu}_1}(a_+) - F_{\tilde{\mu}_1}(a_-) = F_{\tilde{\mu}_0}(\tilde{T}(a_+)) - F_{\tilde{\mu}_0}(\tilde{T}(a_-)) \\ &\geq F_{\tilde{\mu}_0}(a_+) - F_{\tilde{\mu}_0}(a_-) \geq \alpha c_d \frac{(a_+^d - a_-^d)}{d} \end{aligned}$$

which yields the desired contradiction.  $\square$

Let us remark that the proof of Proposition 4.6 gives an alternative proof of the minimum principle in dimension one.

## 5 Convergence of the TV-JKO scheme under a lower bound estimate

We are now interested in the convergence of the TV-JKO scheme to a solution of the fourth-order nonlinear equation (1.2) as the time step  $\tau$  goes to 0. Throughout this section, we assume that  $\Omega$  is a bounded open convex subset of  $\mathbb{R}^d$  and that the initial condition  $\rho_0$  satisfies

$$\rho_0 \in \mathcal{P}_{\text{ac}}(\Omega) \cap BV(\Omega) \cap L^\infty(\Omega), \quad \rho_0 \geq \alpha > 0 \text{ a.e. on } \Omega. \quad (5.1)$$

We fix a time horizon  $T$ , and for small  $\tau > 0$ , define the sequence  $\rho_k^\tau$  by

$$\rho_0^\tau = \rho_0, \quad \rho_{k+1}^\tau \in \operatorname{argmin} \left\{ \frac{1}{2\tau} W_2^2(\rho_k^\tau, \rho) + J(\rho), \quad \rho \in BV \cap \mathcal{P}_{\text{ac}}(\Omega) \right\} \quad (5.2)$$

for  $k = 0, \dots, N_\tau$  with  $N_\tau := \lceil \frac{T}{\tau} \rceil$ . Thanks to Theorem 4.2, (5.1) ensures that the JKO-iterates  $\rho_k^\tau$  defined by (5.2) also remain bounded  $\rho_k^\tau \leq \|\rho_0\|_{L^\infty(\Omega)}$ . We shall also assume that  $\rho_k^\tau$  remains bounded from below by  $\alpha$ :

$$\rho_k^\tau \geq \alpha > 0 \text{ a.e. in } \Omega, \text{ for every } k \text{ and } \tau \quad (5.3)$$

which holds, as we have seen in subsection 4.3 when  $d = 1$  or when  $\Omega$  is a ball and  $\rho_0$  is radially symmetric.

We extend this discrete sequence by piecewise constant interpolation i.e.

$$\rho^\tau(t, x) = \rho_{k+1}^\tau(x), \quad t \in (k\tau, (k+1)\tau], \quad k = 0, \dots, N_\tau, \quad x \in \Omega. \quad (5.4)$$

We shall see that  $\rho^\tau$  converges to a solution  $\rho$  of

$$\partial_t \rho + \operatorname{div} \left( \rho \nabla \operatorname{div} \left( \frac{\nabla \rho}{|\nabla \rho|} \right) \right) = 0, \quad (t, x) \in (0, T) \times \Omega, \quad \rho|_{t=0} = \rho_0, \quad (5.5)$$

with the no-flux boundary condition

$$\rho \nabla \operatorname{div} \left( \frac{\nabla \rho}{|\nabla \rho|} \right) \cdot \nu = 0, \quad \text{on } (0, T) \times \partial\Omega. \quad (5.6)$$

Let us introduce the spaces

$$\begin{aligned} H_{\operatorname{div}}^1(\Omega) &:= \{z \in L^2(\Omega, \mathbb{R}^d) : \operatorname{div}(z) \in L^2(\Omega)\}, \\ H_{\operatorname{div}}^2(\Omega) &:= \{z \in H_{\operatorname{div}}^1(\Omega) : \operatorname{div}(z) \in H^1(\Omega)\} \end{aligned}$$

Since  $\rho$  is no more than BV in  $x$ , one has to be slightly cautious in the meaning of  $\operatorname{div}(\frac{\nabla \rho}{|\nabla \rho|})$  which be conveniently done by interpreting this term as the negative of an element in the subdifferential of  $J$  (in the  $L^2$  sense). For every  $\rho \in \operatorname{BV}(\Omega) \cap L^2(\Omega)$  let us define

$$\partial J(\rho) := \{\operatorname{div}(z) : z \in H_{\operatorname{div}}^1(\Omega), \|z\|_{L^\infty} \leq 1, z \cdot \nu = 0 \text{ on } \partial\Omega, J(\rho) = \int_{\Omega} \operatorname{div}(z) \rho\}.$$

This leads to the following definition:

**Definition 5.1.** *A weak solution of (5.5)-(5.6) is a  $\rho \in L^\infty((0, T), \operatorname{BV}(\Omega) \cap L^\infty(\Omega)) \cap C^0([0, T], (\mathcal{P}(\overline{\Omega}), W_2))$  such that there exists  $z \in L^\infty((0, T) \times \Omega) \cap L^2((0, T), H_{\operatorname{div}}^2(\Omega))$  with*

$$\operatorname{div}(z(t, \cdot)) \in \partial J(\rho(t, \cdot)) \text{ for a.e. } t \in (0, T), \quad (5.7)$$

and  $\rho$  is a weak solution of

$$\partial_t \rho - \operatorname{div} \left( \rho \nabla \operatorname{div}(z) \right) = 0, \quad \rho|_{t=0} = \rho_0, \quad \rho \nabla \operatorname{div}(z) \cdot \nu = 0 \text{ on } (0, T) \times \partial\Omega. \quad (5.8)$$

i.e. for every  $u \in C_c^\infty([0, T] \times \overline{\Omega})$

$$\int_0^T \int_{\Omega} (\partial_t u \rho - \rho \nabla \operatorname{div}(z) \cdot \nabla u) dx dt = - \int_{\Omega} u(0, x) \rho_0(x) dx.$$

We then have

**Theorem 5.2.** *If  $\rho_0$  satisfies (5.6) and the JKO iterates  $\rho_k^\tau$  obey the lower bound (5.3), there exists a vanishing sequence of time steps  $\tau_n \rightarrow 0$  such that the sequence  $\rho^{\tau_n}$  constructed by (5.2)-(5.4) converges strongly in  $L^p((0, T) \times (0, 1))$  for any  $p \in [1, +\infty)$  and in  $L^\infty((0, T), (\mathcal{P}(\overline{\Omega}), W_2))$  to a weak solution of (5.5)-(5.6).*

*Proof.* First,  $\rho_0$  being  $L^\infty$ , we have a uniform  $L^\infty$  bound on  $\rho^\tau$  thanks to Theorem 4.2, and from our extra lower bound assumption (5.3) we have

$$M := \|\rho_0\|_{L^\infty} \geq \rho^\tau(t, x) \geq \alpha, \quad t \in [0, T], \quad \text{a.e. } x \in \Omega. \quad (5.9)$$

Moreover, by construction of the TV-JKO scheme (5.2), one has

$$\frac{1}{2\tau} \sum_{k=0}^{N_\tau} W_2^2(\rho_k^\tau, \rho_{k+1}^\tau) \leq J(\rho_0), \quad \sup_{t \in [0, T]} J(\rho^\tau(t, \cdot)) \leq J(\rho_0) \quad (5.10)$$

By using an Aubin-Lions type compactness Theorem of Savaré and Rossi (Theorem 2 in [24]), the fact that the embedding of  $BV(\Omega)$  into  $L^p(\Omega)$  is compact for every  $p \in [1, \frac{d}{d-1})$  as well as a refinement of Arzèla-Ascoli Theorem (Proposition 3.3.1 in [1]), one obtains (see section 4 of [12] or section 5 of [8] for details) that, up to taking suitable sequence of vanishing times steps  $\tau_n \rightarrow 0$ , we may assume that

$$\rho^\tau \rightarrow \rho \text{ a.e. in } (0, T) \times \Omega \text{ and in } L^p((0, T) \times \Omega), \quad \forall p \in [1, \frac{d}{d-1}) \quad (5.11)$$

and

$$\sup_{t \in [0, T]} W_2(\rho^\tau(t, \cdot), \rho(t, \cdot)) \rightarrow 0 \text{ as } \tau \rightarrow 0, \quad (5.12)$$

for some limit curve  $\rho \in C^{0, \frac{1}{2}}([0, T], (\mathcal{P}(\overline{\Omega}), W_2)) \cap L^q((0, T) \times \Omega)$ . From (5.9) and Lebesgue's dominated convergence Theorem, we deduce that the convergence in (5.11) actually holds for any  $p \in [1, +\infty)$ . It also follows from (5.9) and (5.10), that  $\rho \in L^\infty((0, T), BV(\Omega) \cap L^\infty(\Omega))$  and that  $\rho \geq \alpha$ .

We deduce from the fact that  $\rho_k^\tau \geq \alpha > 0$  and Theorem 3.3 that for each  $k = 0, \dots, N_\tau$ , there exists  $z_k^\tau \in L^\infty(\Omega, \mathbb{R}^d)$  such that  $\text{div}(z_k^\tau) \in W^{1, \infty}(\Omega)$  and

$$\|z_k^\tau\|_{L^\infty} \leq 1, \quad z_k^\tau \cdot \nu = 0 \text{ on } \partial\Omega, \quad J(\rho_k^\tau) = \int_{\Omega} \text{div}(z_k^\tau) \rho_k^\tau, \quad (5.13)$$

and the optimal (backward) optimal transport  $T_{k+1}^\tau$  from  $\rho_{k+1}^\tau$  to  $\rho_k^\tau$  is related to  $z_{k+1}^\tau$  by

$$\text{id} - T_{k+1}^\tau = -\tau \nabla \text{div}(z_{k+1}^\tau). \quad (5.14)$$

We extend  $z_k^\tau$  in a piecewise constant way i.e. set

$$z^\tau(t, x) = z_{k+1}^\tau(x), \quad t \in (k\tau, (k+1)\tau], \quad k = 0, \dots, N_\tau, \quad x \in \Omega. \quad (5.15)$$

We then observe that

$$\begin{aligned}
W_2^2(\rho_k^\tau, \rho_{k+1}^\tau) &= \int_{\Omega} |x - T_{k+1}^\tau(x)|^2 \rho_{k+1}^\tau(x) dx \\
&= \tau^2 \int_{\Omega} |\nabla \operatorname{div}(z_{k+1}^\tau)|^2 \rho_{k+1}^\tau(x) dx \\
&\geq \alpha \tau^2 \int_{\Omega} |\nabla \operatorname{div}(z_{k+1}^\tau)|^2 dx
\end{aligned}$$

Thanks to (5.10) we thus deduce that  $\nabla \operatorname{div} z^\tau$  is bounded in  $L^2((0, T) \times \Omega)$ , since  $\operatorname{div}(z^\tau)$  has zero-mean, with Poincaré-Wirtinger inequality, we obtain

$$\int_0^T \|\operatorname{div}(z^\tau)\|_{H^1(\Omega)}^2 dt \leq C. \quad (5.16)$$

We may therefore assume (up to further suitable extractions) that there is some  $z \in L^\infty((0, T) \times \Omega) \cap L^2((0, T), H_{\operatorname{div}}^2(\Omega))$  such that  $z^\tau$  converges to  $z$  weakly  $*$  in  $L^\infty((0, T) \times \Omega)$  and  $(\operatorname{div}(z^\tau), \nabla \operatorname{div}(z^\tau))$  converges weakly in  $L^2((0, T) \times \Omega)$  to  $(\operatorname{div}(z), \nabla \operatorname{div}(z))$ . Of course  $\|z\|_{L^\infty} \leq 1$  and  $z(t, \cdot) \cdot \nu = 0$  on  $\partial\Omega$  for a.e.  $t$ . Note also that  $\rho^\tau \nabla \operatorname{div}(z^\tau)$  converges weakly in  $L^1((0, T) \times \Omega)$  to  $\rho \nabla \operatorname{div}(z)$ .

The limiting equation can now be derived using standard computations (see the proof of Theorem 5.1 of the seminal work [17], or chapter 8 of [26]): Let  $u \in C_c^2([0, T) \times \bar{\Omega})$  and observe that

$$\int_0^T \int_{\Omega} \partial_t u \rho^\tau dx dt = \sum_{k=1}^{N_\tau} \left( \int_{\Omega} u(k\tau, x) (\rho_k^\tau(x) - \rho_{k+1}^\tau(x)) dx \right) - \int_{\Omega} u(0, x) \rho_1^\tau(x) dx.$$

Recalling that  $\rho_k^\tau = T_{k+1}^\tau \# \rho_{k+1}^\tau$ , and applying Taylor's theorem, we have

$$\begin{aligned}
&\sum_{k=1}^{N_\tau} \left( \int_{\Omega} u(k\tau, x) (\rho_k^\tau(x) - \rho_{k+1}^\tau(x)) dx \right) \\
&= \sum_{k=1}^{N_\tau} \left( \int_{\Omega} ((T_{k+1}^\tau(x) - x) \cdot \nabla u(k\tau, x) + \tilde{R}_\tau(x)) \rho_{k+1}^\tau(x) dx \right) \\
&= \sum_{k=1}^{N_\tau} \left( \int_{\Omega} (\tau(\nabla \operatorname{div}(z_{k+1}^\tau)) \cdot \nabla u(k\tau, x) + \tilde{R}_\tau(x)) \rho_{k+1}^\tau(x) dx \right),
\end{aligned}$$

where  $|\tilde{R}_\tau(x)| \leq C \|D^2 u(k\tau, \cdot)\|_{L^\infty} |T_{k+1}^\tau(x) - x|^2$ . Note also that for  $t \in (k\tau, (k+1)\tau]$ ,  $|\nabla u(k\tau, \cdot) - \nabla u(t, \cdot)| \leq \tau \|\partial_t \nabla u\|_{L^\infty}$ . Therefore,

$$\int_0^T \int_{\Omega} (\partial_t u \rho^\tau - \rho^\tau \nabla \operatorname{div}(z^\tau) \cdot \nabla u) dx dt = - \int_{\Omega} u(0, x) \rho_1^\tau(x) dx + R_\tau(u) \quad (5.17)$$



with

$$|R_\tau(u)| \leq C \max\{\|D^2u\|_{L^\infty}, \|\partial_t \nabla u\|_{L^\infty}\} \sum_{k=0}^{N_\tau} W_2^2(\rho_k^\tau, \rho_{k+1}^\tau) \leq C\tau. \quad (5.18)$$

Passing to the limit  $\tau$  to 0 in (5.17) yields that  $\rho$  is a weak solution to

$$\partial_t \rho - \operatorname{div}(\rho \nabla \operatorname{div}(z)) = 0, \quad \rho|_{t=0} = \rho_0, \quad \rho \nabla \operatorname{div}(z) \cdot \nu = 0 \text{ on } (0, T) \times \partial\Omega.$$

It remains to prove that  $J(\rho(t, \cdot)) = \int_\Omega \operatorname{div}(z(t, x)) \rho(t, x) dx$ , for a.e.  $t \in (0, T)$ . The inequality  $J(\rho(t, \cdot)) \geq \int_\Omega \operatorname{div}(z(t, x)) \rho(t, x) dx$  is obvious since  $z(t, \cdot) \in H_{\operatorname{div}}^1(\Omega)$ ,  $z(t, \cdot) \cdot \nu = 0$  on  $\partial\Omega$  and  $\|z(t, \cdot)\|_{L^\infty} \leq 1$ . To prove the converse inequality, we use Fatou's Lemma, the lower semi-continuity of  $J$ , (5.13) and the weak-convergence in  $L^1((0, T) \times \Omega)$  of  $\rho^\tau \operatorname{div}(z^\tau)$  to  $\rho \operatorname{div}(z)$ :

$$\begin{aligned} \int_0^T J(\rho(t, \cdot)) dt &\leq \int_0^T \liminf_\tau J(\rho^\tau(t, \cdot)) dt \\ &\leq \liminf_\tau \int_0^T J(\rho^\tau(t, \cdot)) dt \\ &= \liminf_\tau \int_0^T \int_\Omega \operatorname{div}(z^\tau(t, x)) \rho^\tau(t, x) dx dt \\ &= \int_0^T \int_\Omega \operatorname{div}(z(t, x)) \rho(t, x) dx dt \end{aligned}$$

which concludes the proof. □

**Acknowledgements:** The authors wish to thank Vincent Duval and Gabriel Peyré for suggesting the TV-Wasserstein problem to them as well as for fruitful discussions. They also thank Maxime Laborde and Filippo Santambrogio for helpful remarks in particular regarding the maximum principle.

## References

- [1] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré. *Gradient flows: in metric spaces and in the space of probability measures*. Springer Science & Business Media, 2008.
- [2] G. Anzellotti. Pairing between measures and bounded functions and compensated compactness. *Ann. di Matematica Pura ed Appl.*, IV(135):293–318, 1983.

- [3] Giovanni Bellettini, Vicent Caselles, and Matteo Novaga. The total variation flow in  $\mathbb{R}^n$ . *Journal of Differential Equations*, 184(2):475–525, 2002.
- [4] Martin Benning, Luca Calatroni, Bertram Düring, and Carola-Bibiane Schönlieb. A primal-dual approach for a total variation Wasserstein flow. In *Geometric Science of Information*, pages 413–421. Springer, 2013.
- [5] Yann Brenier. Polar factorization and monotone rearrangement of vector-valued functions. *Comm. Pure Appl. Math.*, 44(4):375–417, 1991.
- [6] Haïm Brezis. *Analyse fonctionnelle*. Collection Mathématiques Appliquées pour la Maîtrise. [Collection of Applied Mathematics for the Master’s Degree]. Masson, Paris, 1983. Théorie et applications. [Theory and applications].
- [7] Martin Burger, Marzena Franek, and Carola-Bibiane Schönlieb. Regularized regression and density estimation based on optimal transport. *Applied Mathematics Research eXpress*, 2012(2):209–253, 2012.
- [8] Guillaume Carlier and Maxime Laborde. A splitting method for non-linear diffusions with nonlocal, nonpotential drifts. *Nonlinear Anal.*, 150:1–18, 2017.
- [9] Antonin Chambolle, Vicent Caselles, Daniel Cremers, Matteo Novaga, and Thomas Pock. An introduction to total variation for image analysis. In *Theoretical foundations and numerical methods for sparse recovery*, volume 9 of *Radon Ser. Comput. Appl. Math.*, pages 263–340. Walter de Gruyter, Berlin, 2010.
- [10] Antonin Chambolle, Vincent Duval, Gabriel Peyré, and Clarice Poon. Geometric properties of solutions to the total variation denoising problem. *Inverse Problems*, 33(1):015002, 2016.
- [11] Guido De Philippis, Alpár Richárd Mészáros, Filippo Santambrogio, and Bozhidar Velichkov. BV estimates in optimal transportation and applications. *Archive for Rational Mechanics and Analysis*, 219(2):829–860, 2016.
- [12] Marco Di Francesco and Daniel Matthes. Curves of steepest descent are entropy solutions for a class of degenerate convection-diffusion equations. *Calc. Var. Partial Differential Equations*, 50(1-2):199–230, 2014.

- [13] Bertram Düring and Carola-Bibiane Schönlieb. A high-contrast fourth-order pde from imaging: numerical solution by ADI splitting. *Multi-scale and High-Contrast Partial Differential Equations*, H. Ammari et al.(eds.), pages 93–103, 2012.
- [14] L. C. Evans and J. Spruck. Motion of level sets by mean curvature. II. *Trans. Amer. Math. Soc.*, 330(1):321–332, 1992.
- [15] Mi-Ho Giga and Yoshikazu Giga. Very singular diffusion equations: second and fourth order problems. *Jpn. J. Ind. Appl. Math.*, 27(3):323–345, 2010.
- [16] Yoshikazu Giga, Hirotoshi Kuroda, and Hideki Matsuoka. Fourth-order total variation flow with Dirichlet condition: characterization of evolution and extinction time estimates. *Adv. Math. Sci. Appl.*, 24(2):499–534, 2014.
- [17] Richard Jordan, David Kinderlehrer, and Felix Otto. The variational formulation of the Fokker–Planck equation. *SIAM journal on mathematical analysis*, 29(1):1–17, 1998.
- [18] Daniel Loibl, Daniel Matthes, and Jonathan Zinsl. Existence of weak solutions to a class of fourth order partial differential equations with Wasserstein gradient structure. *Potential Anal.*, 45(4):755–776, 2016.
- [19] E Gonzalez-U Massari. Variational mean curvatures. *Rend. Sem. Mat. Univ. Pol. Torino*, 52:1–28, 1994.
- [20] Umberto Massari. Esistenza e regolarità delle ipersuperfici di curvatura media assegnata in  $\mathbb{R}^n$ . *Arch. Ration. Mech. Anal.*, 55:357–382, 1974.
- [21] Umberto Massari. Frontiere orientate di curvatura media assegnata in  $L^p$ . *Rendiconti del Seminario Matematico della Università di Padova*, 53:37–52, 1975.
- [22] Daniel Matthes, Robert J. McCann, and Giuseppe Savaré. A family of nonlinear fourth order equations of gradient flow type. *Comm. Partial Differential Equations*, 34(10-12):1352–1397, 2009.
- [23] Robert J. McCann. A convexity principle for interacting gases. *Adv. Math.*, 128(1):153–179, 1997.
- [24] Riccarda Rossi and Giuseppe Savaré. Tightness, integral equicontinuity and compactness for evolution problems in Banach spaces. *Annali della Scuola Normale Superiore di Pisa-Classe di Scienze*, 2(2):395–431, 2003.

- [25] Leonid I. Rudin, Stanley Osher, and Emad Fatemi. Nonlinear total variation based noise removal algorithms. *Phys. D*, 60(1-4):259–268, 1992. Experimental mathematics: computational issues in nonlinear science (Los Alamos, NM, 1991).
- [26] Filippo Santambrogio. *Optimal transport for applied mathematicians*, volume 87 of *Progress in Nonlinear Differential Equations and their Applications*. Birkhäuser/Springer, Cham, 2015. Calculus of variations, PDEs, and modeling.
- [27] Italo Tamanini. Boundaries of Caccioppoli sets with Hölder-continuous normal vector. *Journal für die reine und angewandte Mathematik*, 334:27–39, 1982.
- [28] Cédric Villani. *Topics in optimal transportation*, volume 58 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2003.