On the Total Variation Wasserstein Gradient Flow

Guillaume Carlier
Université Paris-Dauphine

Clarice Poon
University of Cambridge

May 2, 2017

Total variation gradient flows

Rudin, Osher and Fatemi (1992): Given $u_0 \in L^2(\Omega) \cap BV(\Omega)$, let

$$u_1 \in \operatorname*{argmin}_{u \in BV(\Omega)} |Du|(\Omega) + \frac{1}{2\tau} \|u - u_0\|_{L^2}^2.$$

The Euler-Lagrange equation is

$$\frac{u_1 - u_0}{\tau} \in -\partial |Du_1| = \operatorname{div}\left(\frac{Du_1}{|Du_1|}\right).$$

Total variation gradient flows

Rudin, Osher and Fatemi (1992): Given $u_0 \in L^2(\Omega) \cap BV(\Omega)$, let

$$u_1 \in \operatorname*{argmin}_{u \in BV(\Omega)} |Du|(\Omega) + \frac{1}{2\tau}\|u - u_0\|_{L^2}^2.$$

The Euler-Lagrange equation is

$$\frac{u_1 - u_0}{\tau} \in -\partial |Du_1| = \operatorname{div}\left(\frac{Du_1}{|Du_1|}\right).$$

The PDE associated with the L^2 gradient descent of the functional $u \mapsto |Du|$ is

$$\partial_t u = \operatorname{div}\left(\frac{Du}{|Du|}\right), \quad \text{on } (0,T) \times \Omega, \qquad u|_{t=0} = u_0$$

with boundary condition

$$\nabla u \cdot \nu = 0 \text{ on } \partial\Omega.$$

Studied by Andreu, Ballester, Caselles and Mazón (2001), Bellettini, Caselles and Novaga (2002), Evans and Spruck (1992), ...

Total variation gradient flows

Rudin, Osher and Fatemi (1992): Given $u_0 \in L^2(\Omega) \cap BV(\Omega)$, let

$$u_1 \in \operatorname*{argmin}_{u \in BV(\Omega)} |Du|(\Omega) + \frac{1}{2\tau}\|u - u_0\|_H^2.$$

The Euler-Lagrange equation is

$$\frac{u_1 - u_0}{\tau} \in -\partial|Du_1| = \operatorname{div}\left(\frac{Du_1}{|Du_1|}\right).$$

The PDE associated with the L^2 gradient descent of the functional $u \mapsto |Du|$ is

$$\partial_t u = \operatorname{div}\left(\frac{Du}{|Du|}\right), \quad \text{on } (0, T) \times \Omega, \qquad u|_{t=0} = u_0$$

with boundary condition

$$\nabla u \cdot \nu = 0 \text{ on } \partial\Omega.$$

Studied by Andreu, Ballester, Caselles and Mazón (2001), Bellettini, Caselles and Novaga (2002), Evans and Spruck (1992), ...

Studies of the TV gradient flow with other Hilbertian norms have been popular, in particular, the H^{-1} norm (Giga and Giga, 2010).

The total variation Wasserstein gradient flow

We consider the fourth-order nonlinear evolution equation

$$\partial_t \rho + \operatorname{div}\Big(\rho\, \nabla \operatorname{div}\Big(\frac{\nabla \rho}{|\nabla \rho|}\Big)\Big) = 0, \text{ on } (0, T) \times \Omega, \qquad \rho_{|_{t=0}} = \rho_0,$$

supplemented by the zero-flux boundary condition

$$\rho \nabla \operatorname{div}\Big(\frac{\nabla \rho}{|\nabla \rho|}\Big) \cdot \nu = 0 \text{ on } \partial\Omega,$$

The total variation Wasserstein gradient flow

We consider the fourth-order nonlinear evolution equation

$$\partial_t \rho + \operatorname{div}\Big(\rho \, \nabla \operatorname{div}\Big(\frac{\nabla \rho}{|\nabla \rho|}\Big)\Big) = 0, \text{ on } (0, T) \times \Omega, \qquad \rho_{|_{t=0}} = \rho_0,$$

supplemented by the zero-flux boundary condition

$$\rho \nabla \operatorname{div}\Big(\frac{\nabla \rho}{|\nabla \rho|}\Big) \cdot \nu = 0 \text{ on } \partial\Omega,$$

Applications and numerical solvers:

- Burger, Franek & Schönlieb. Regularized regression and density estimation based on optimal transport. 2012.
- Düring & Schönlieb. A high-contrast 4th order PDE from imaging: numerical solution by ADI splitting. 2012.
- Benning & Calatroni & Düring & Schönlieb. A primal-dual approach for a total variation Wasserstein flow. 2013.

Relation to the TV Wasserstein variational problem: JKO scheme

Due to the work of Jordan, Kinderlehrer and Otto for the Fokker-Planck equation,

$$\partial_t \rho + \operatorname{div}\Big(\rho\, \nabla(-E'(\rho))\Big) = 0, \text{ on } (0,T)\times\Omega, \quad \rho_{|_{t=0}} = \rho_0,$$

with zero-flux boundary condition can be obtained, at the limit $\tau \to 0^+$, of the JKO Euler implicit scheme:

$$\rho_0^\tau = \rho_0,\ \rho_{k+1}^\tau \in \operatorname{argmin}\Big\{\frac{1}{2\tau} W_2^2(\rho_k^\tau, \rho) + E(\rho),\ \rho \in \mathrm{BV}(\Omega)\cap \mathcal{P}_2(\overline{\Omega})\Big\}$$

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:

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

$\Pi(\rho_0,\rho_1)$ denoting the set of transport plans between ρ_0 and ρ_1 i.e. the set of probability measures on $\mathbb{R}^d \times \mathbb{R}^d$ having ρ_0 and ρ_1 as marginals.

Relation to the TV Wasserstein variational problem: JKO scheme

Due to the work of Jordan, Kinderlehrer and Otto for the Fokker-Planck equation,

$$\partial_t \rho + \mathrm{div}\Big(\rho\, \nabla \mathrm{div}\Big(\frac{\nabla \rho}{|\nabla \rho|}\Big)\Big) = 0, \text{ on } (0,T)\times\Omega, \quad \rho_{|_{t=0}} = \rho_0,$$

with zero-flux boundary condition can be obtained, at the limit $\tau \to 0^+$, of the JKO Euler implicit scheme:

$$\rho_0^\tau = \rho_0,\ \rho_{k+1}^\tau \in \mathrm{argmin}\Big\{\frac{1}{2\tau} W_2^2(\rho_k^\tau, \rho) + |D\rho|(\Omega),\ \rho \in \mathrm{BV}(\Omega)\cap \mathcal{P}_2(\overline{\Omega})\Big\}$$

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:

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

$\Pi(\rho_0,\rho_1)$ denoting the set of transport plans between ρ_0 and ρ_1 i.e. the set of probability measures on $\mathbb{R}^d \times \mathbb{R}^d$ having ρ_0 and ρ_1 as marginals.

This talk

In this talk, we present a study of solutions to the variational problem and its gradient flow.

1. Properties of the JKO iterates:
   1. a maximum principle.
   2. establish the optimality conditions.
   3. regularity of the level sets.
   4. analysis for step function initial data.
2. Convergence of the JKO scheme as $\tau \to 0$ in 1D, for strictly positive initial density ρ_0.

Some notation and definitions

Given an open subset Ω of $\mathbb{R}^d$ and $\rho \in L^1(\Omega)$,

- the total variation of ρ is given by

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

  and $\mathrm{BV}(\Omega)$ is by definition the subspace of $L^1(\Omega)$ consisting of those ρ's in $L^1(\Omega)$ such that $J(\rho)$ is finite.

- By defining

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

  we have that Γ_d is closed and convex in $L^d(\Omega)$ and $J : L^{d/(d-1)} \to [0, \infty)$ is its support function:

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

- the subdifferential of J at ρ is

$$\partial J(\rho) = \left\{ \xi \in \Gamma_d \ : \ \int_\Omega \xi\rho = J(\rho) \right\}.$$

Some examples in 1D

Let

$$\Phi(\rho) := \frac{1}{2\tau} W_2^2(\rho_0, \rho) + J(\rho), \qquad \forall \rho \in \mathrm{BV}(\mathbb{R}^d) \cap \mathcal{P}_2(\mathbb{R}^d).$$

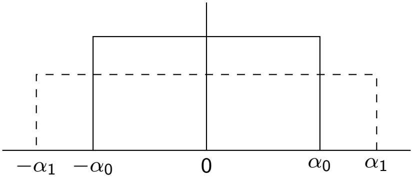
A sufficient optimality condition: ρ_1 is the minimizing solution if there exists $z \in C^1(\mathbb{R}^d; \mathbb{R}^d)$ with $\|z\|_{L^\infty} \leq 1$, $\mathrm{div}(z) \in L^d$ and $J(\rho_1) = \int_{\mathbb{R}^d} \mathrm{div}(z)\rho_1$ such that

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

where φ is a Kantorovich potential from ρ_1 to ρ_0.

Characteristic function

Let $\rho_0 := \frac{1}{2\alpha_0}\chi_{[-\alpha_0,\alpha_0]}$. Then, $\rho_1 := \frac{1}{2\alpha_1}\chi_{[-\alpha_1,\alpha_1]}$ where $\frac{\alpha_1^2(\alpha_1-\alpha_0)}{\tau} = 3$.



Setting $\rho^\tau(t) = \rho^\tau_{k+1}$ for $t \in (k\tau, (k+1)\tau]$,

- ρ^τ converges to $\rho(t,.) = \frac{1}{2\alpha(t)}\chi_{[-\alpha(t),\alpha(t)]}$ with $\alpha(t) = (\alpha_0^3 + 9t)^{1/3}$ in $L^\infty((0,T),(\mathcal{P}_2(\mathbb{R}),W_2))$ and in $L^p((0,T)\times\mathbb{R})$ for any $p \in (1,\infty)$.
- Moreover, ρ solves the continuity equation $\partial_t\rho + (-\rho z_{xx})_x = 0$ where

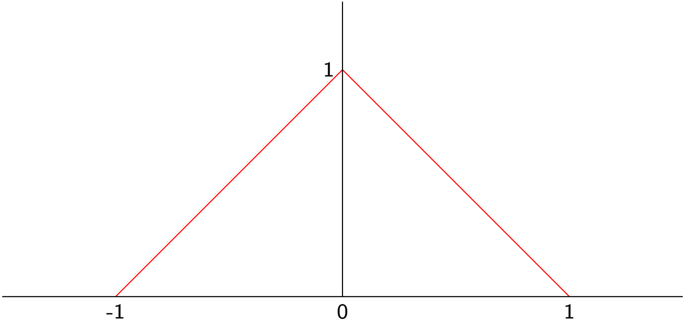
$$z(t,x) = -\frac{\alpha'(t)}{6\alpha(t)}x^3 + \frac{3}{2\alpha(t)}x, \quad x \in [-\alpha(t),\alpha(t)].$$

Instantaneous creation of discontinuities

Let $\rho_0 = (1 - |x|)_+$. Then one can show that

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

for $\beta \in (0, 1)$.



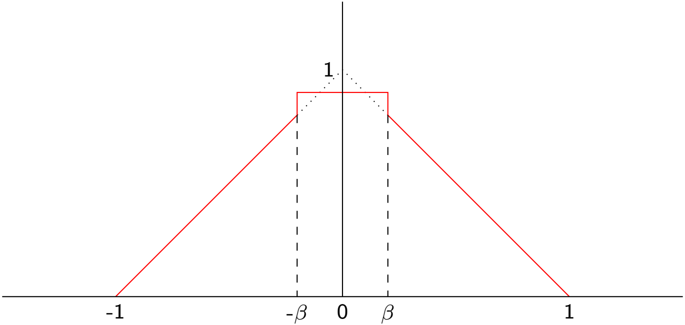
This example shows that discontinuities may appear at the very first iteration of the TV-JKO scheme.

Instantaneous creation of discontinuities

Let $\rho_0 = (1 - |x|)_+$. Then one can show that

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

for $\beta \in (0, 1)$.



This example shows that discontinuities may appear at the very first iteration of the TV-JKO scheme.

One step of the JKO flow

- Let $\Omega \subset \mathbb{R}^d$ be a convex open bounded subset.
- $\mathcal{P}_{\mathrm{ac}}(\Omega)$ the set of Borel probability measures on Ω that are absolutely continuous with respect to the Lebesgue measure

Given $\rho_0 \in \mathcal{P}_{\mathrm{ac}}(\Omega)$ and $\tau > 0$, we consider one step of the TV-JKO scheme:

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

- Existence of solutions follows by the direct method of the calculus of variations.
- Since J is convex and $\rho \mapsto W_2^2(\rho, \rho_0)$ is strictly convex whenever $\rho_0 \in \mathcal{P}_{\mathrm{ac}}(\Omega)$ (Santambrogio, 2015), the minimizer is in fact unique, and in the sequel we denote it by ρ_1.

Maximum principle

Theorem (Carlier & P. 2017)

Let $\rho_0 \in \mathcal{P}_{\mathrm{ac}}(\Omega) \cap L^\infty(\Omega)$ *and let* ρ_1 *be the solution of* $(\mathcal{P}_\tau)$*, then* $\rho_1 \in L^\infty(\Omega)$ *with*

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

Maximum principle

Theorem (Carlier & P. 2017)

Let $\rho_0 \in \mathcal{P}_{\mathrm{ac}}(\Omega) \cap L^\infty(\Omega)$ and let ρ_1 be the solution of $(\mathcal{P}_\tau)$, then $\rho_1 \in L^\infty(\Omega)$ with

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

TOOL 1 – BV estimate by De Philippis, Mészáros, Santambrogio & Velichkov (2016): given $\mu \in \mathcal{P}_{\mathrm{ac}}(\Omega) \cap \mathrm{BV}(\Omega)$, and $G : \mathbb{R}_+ \to \mathbb{R} \cup \{+\infty\}$, proper convex l.s.c., the solution of

$$\hat{\rho} \in \mathrm{argmin}_{\rho \in \mathcal{P}_{\mathrm{ac}}(\Omega)} \left\{ \frac{1}{2} W_2^2(\mu, \rho) + \int_\Omega G(\rho(x)) \mathrm{d}x \right\}$$

is BV with the bound $J(\hat{\rho}) \leq J(\mu)$.

Maximum principle

Theorem (Carlier & P. 2017)

Let $\rho_0 \in \mathcal{P}_{\rm ac}(\Omega) \cap L^\infty(\Omega)$ *and let* ρ_1 *be the solution of* $(\mathcal{P}_\tau)$*, then* $\rho_1 \in L^\infty(\Omega)$ *with*

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

TOOL 1 – BV estimate by De Philippis, Mészáros, Santambrogio & Velichkov (2016): given $\mu \in \mathcal{P}_{\rm ac}(\Omega) \cap {\rm BV}(\Omega)$, and $G : \mathbb{R}_+ \to \mathbb{R} \cup \{+\infty\}$, proper convex l.s.c., the solution of

$$\hat\rho \in {\rm argmin}_{\rho \in \mathcal{P}_{\rm ac}(\Omega)} \left\{ \frac{1}{2} W_2^2(\mu, \rho) + \int_\Omega G(\rho(x)){\rm d}x \right\}$$

is BV with the bound $J(\hat\rho) \leq J(\mu)$.

Choose

$$G(\rho) := \begin{cases} 0 & \text{if } \rho \in K, \\ +\infty & \text{otherwise,} \end{cases} \qquad K := \{\rho \in \mathcal{P}_{\rm ac}(\Omega) \ : \ \rho \leq \|\rho_0\|_{L^\infty(\Omega)}\}$$

Let $\hat\rho_1 = {\rm argmin}_{\rho \in K} W_2^2(\rho_1, \rho)$. *Then,* $J(\hat\rho_1) \leq J(\rho_1)$. *Is* $W_2(\hat\rho_1, \rho_0) \leq W_2(\rho_1, \rho_0)$*?*

Maximum principle

Tool II: generalized geodesics

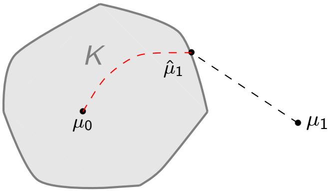
Given $\overline{\mu}$, μ_0 and μ_1 in $\mathcal{P}_{\rm ac}(\Omega)$, and denoting by T_0 (respectively T_1) the optimal transport (Brenier) map between $\overline{\mu}$ and μ_0 (respectively μ_1), the generalized geodesic with base $\overline{\mu}$ joining μ_0 to μ_1 is by definition the curve of measures:

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

Lemma (Ambrosio, Gigli & Savaré (2008))

Suppose that K is a nonempty subset of $\mathcal{P}_{\rm ac}(\Omega)$ such that: for $\mu_0 \in K$, $\mu_1 \in \mathcal{P}_{\rm ac}(\Omega)$, $\hat{\mu}_1 \in \operatorname{argmin}_{\mu\in K} W_2^2(\mu_1, \mu)$ implies that the generalized geodesic joining μ_0 to $\hat{\mu}_1$ with base μ_1 remains in K.
Then,

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



Maximum principle

Tool II: generalized geodesics

Given $\overline{\mu}$, μ_0 and μ_1 in $\mathcal{P}_{\rm ac}(\Omega)$, and denoting by T_0 (respectively T_1) the optimal transport (Brenier) map between $\overline{\mu}$ and μ_0 (respectively μ_1), the generalized geodesic with base $\overline{\mu}$ joining μ_0 to μ_1 is by definition the curve of measures:

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

Lemma (Ambrosio, Gigli & Savaré (2008))

Suppose that K is a nonempty subset of $\mathcal{P}_{\rm ac}(\Omega)$ such that: for $\mu_0 \in K$, $\mu_1 \in \mathcal{P}_{\rm ac}(\Omega)$, $\hat{\mu}_1 \in \mathrm{argmin}_{\mu \in K} W_2^2(\mu_1, \mu)$ implies that the generalized geodesic joining μ_0 to $\hat{\mu}_1$ with base μ_1 remains in K.
Then,

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

Valid choices:

- $K = \{\mu;\ \|\mu\|_{L^p} \leq C\}$ for any $p \in [1, \infty]$.

Maximum principle

Tool II: generalized geodesics

Given $\overline{\mu}$, μ_0 and μ_1 in $\mathcal{P}_{\rm ac}(\Omega)$, and denoting by T_0 (respectively T_1) the optimal transport (Brenier) map between $\overline{\mu}$ and μ_0 (respectively μ_1), the generalized geodesic with base $\overline{\mu}$ joining μ_0 to μ_1 is by definition the curve of measures:

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

Lemma (Ambrosio, Gigli & Savaré (2008))

Suppose that K is a nonempty subset of $\mathcal{P}_{\rm ac}(\Omega)$ such that: for $\mu_0 \in K$, $\mu_1 \in \mathcal{P}_{\rm ac}(\Omega)$, $\hat{\mu}_1 \in \operatorname{argmin}_{\mu \in K} W_2^2(\mu_1, \mu)$ implies that the generalized geodesic joining μ_0 to $\hat{\mu}_1$ with base μ_1 remains in K.
Then,

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

Valid choices:

- $K = \{\mu;\ \|\mu\|_{L^p} \leq C\}$ for any $p \in [1, \infty]$.
- In 1D, $K := \{\rho \in \mathcal{P}_{\rm ac}(\Omega)\ :\ \rho \geq \alpha\}$.

Proposition (Minimum principle in 1D, Carlier & P. 2017)

Assume that $d = 1$, that Ω is a bounded interval and that $\rho_0 \geq \alpha > 0$ a.e. on Ω then the solution ρ_1 of $(\mathcal{P}_\tau)$ also satifies $\rho_1 \geq \alpha > 0$ a.e. on Ω.

Optimality condition

By considering directly the first order condition of Φ, one can see that there exists a Kantorovich potential φ from ρ_1 to ρ_0 such that

$$\tau^{-1} \int \varphi\mu \leq J(\mu)$$

for all $\mu \in \mathcal{M}(X)$ such that $\mathrm{Supp}(\mu) \subset \mathrm{Supp}(\rho_1) =: \Omega_1$, and with equality when $\mu = \rho_1$.

Problems:

- Assuming that $\partial\Omega_1$ is Lipschitz, one deduce that on Ω_1, $\tau^{-1}\varphi = \mathrm{div}(z)$ with $\|z\|_{L^\infty} \leq 1$.
- How should we define $\mathrm{div}(z)$ outside of Ω_1? Which function space does it belong to?

Entropic approximation

Given $h > 0$ we consider the following approximation of $(\mathcal{P}_\tau)$:

$$\inf_{\rho \in \mathcal{P}(\Omega)} \left\{ \frac{1}{2\tau} W_2^2(\rho_0, \rho) + J(\rho) + hE(\rho) \right\} \qquad (\mathcal{E}_h)$$

where

$$E(\rho) := \int_\Omega \rho(x) \log(\rho(x)) \mathrm{d}x.$$

$(\mathcal{E}_h)$ admits a unique solution ρ_h and since $J(\rho_h)$ is bounded, up to a subsequence, ρ_h converges as $h \to 0$ a.e. and strongly in $L^p(\Omega)$ for every $p \in [1, \frac{d}{d-1})$ to ρ_1 the solution of $(\mathcal{P}_\tau)$.

Entropic approximation

Given $h > 0$ we consider the following approximation of $(\mathcal{P}_\tau)$:

$$\inf_{\rho \in \mathcal{P}(\Omega)} \left\{ \frac{1}{2\tau} W_2^2(\rho_0, \rho) + J(\rho) + hE(\rho) \right\} \qquad (\mathcal{E}_h)$$

where

$$E(\rho) := \int_\Omega \rho(x) \log(\rho(x)) \mathrm{d}x.$$

$(\mathcal{E}_h)$ admits a unique solution ρ_h and since $J(\rho_h)$ is bounded, up to a subsequence, ρ_h converges as $h \to 0$ a.e. and strongly in $L^p(\Omega)$ for every $p \in [1, \frac{d}{d-1})$ to ρ_1 the solution of $(\mathcal{P}_\tau)$.

Lemma

There is an $\alpha_h > 0$ such that $\rho_h \geq \alpha_h$ a.e.. In particular, $\beta_h := h \log(\rho_h)$ is uniformly bounded from below and is bounded in $L^p(\Omega)$ for any $p \geq 1$. Moreover, $\max(0, \beta_h)$ converges to 0 strongly in L^p.

The optimality condition

We then have the following characterization for ρ_h:

Proposition (Carlier & P. 2017)

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

$$\frac{\varphi_h}{\tau} + \mathrm{div}(z_h) = -h \log(\rho_h), \ \textit{a.e. in } \Omega$$

where φ_h *is the Kantorovich potential between* ρ_h *and* ρ_0.

The optimality condition

We then have the following characterization for ρ_h:

Proposition (Carlier & P. 2017)

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

$$\frac{\varphi_h}{\tau} + \mathrm{div}(z_h) = -h \log(\rho_h), \ \textit{a.e. in } \Omega$$

where φ_h is the Kantorovich potential between ρ_h and ρ_0.

By taking the limit as $h \to 0$:

Theorem (Carlier & P. 2017)

If ρ_1 solves $(\mathcal{P}_\tau)$, there exists φ a Kantorovich potential between ρ_0 and ρ_1 (in particular $\mathrm{id} - \nabla\varphi$ is the optimal transport between ρ_1 and ρ_0), $\beta \in L^\infty(\Omega)$, $\beta \geq 0$ and $z \in L^\infty(\Omega, \mathbb{R}^d)$ such that

$$\frac{\varphi}{\tau} + \mathrm{div}(z) = \beta,$$

and

$$\beta\rho_1 = 0, \ \ \|z\|_{L^\infty} \leq 1, \ J(\rho_1) = \int_\Omega \mathrm{div}(z)\rho_1, \ z \cdot \nu = 0 \text{ on } \partial\Omega.$$

Variational curvatures and regularity of level sets

$\mathrm{div}(z) \in L^\infty$ implies boundary regularity of the level sets of ρ_1:

Variational curvatures and regularity of level sets

$\mathrm{div}(z) \in L^\infty$ implies boundary regularity of the level sets of ρ_1:

- Anzellotti (1983): We can interpret $z \cdot D\rho_1$ as a Radon measure and $|D\rho_1| = -z \cdot D\rho_1$.

Variational curvatures and regularity of level sets

$\mathrm{div}(z) \in L^\infty$ implies boundary regularity of the level sets of ρ_1:

- Anzellotti (1983): We can interpret $z \cdot D\rho_1$ as a Radon measure and $|D\rho_1| = -z \cdot D\rho_1$.
- Barozzi, Massari, Tamanini (1975-1995): A set of finite perimeter $E \subset \Omega \subset \mathbb{R}^d$ is said to have variational mean curvature $g \in L^1(\Omega)$ precisely when E minimizes

$$\min_{F \subset \Omega} \mathrm{Per}(F) + \int_F g.$$

If, in addition, $g \in L^p(\Omega)$ with $p \in (d, +\infty]$, then the reduced boundary $\partial^* E$ is a $(d-1)$-dimensional manifold of class $C^{1,\alpha}$ with $\alpha \geq \frac{p-d}{2p}$ and $\mathcal{H}^s((\partial E \setminus \partial^* E) \cap \Omega) = 0$ for all $s > d - 8$.

Variational curvatures and regularity of level sets

$\mathrm{div}(z) \in L^\infty$ implies boundary regularity of the level sets of ρ_1:

- Anzellotti (1983): We can interpret $z \cdot D\rho_1$ as a Radon measure and $|D\rho_1| = -z \cdot D\rho_1$.
- Barozzi, Massari, Tamanini (1975-1995): A set of finite perimeter $E \subset \Omega \subset \mathbb{R}^d$ is said to have variational mean curvature $g \in L^1(\Omega)$ precisely when E minimizes

$$\min_{F \subset \Omega} \mathrm{Per}(F) + \int_F g.$$

  If, in addition, $g \in L^p(\Omega)$ with $p \in (d, +\infty]$, then the reduced boundary $\partial^* E$ is a $(d-1)$-dimensional manifold of class $C^{1,\alpha}$ with $\alpha \geq \frac{p-d}{2p}$ and $\mathcal{H}^s((\partial E \setminus \partial^* E) \cap \Omega) = 0$ for all $s > d - 8$.
- Chambolle, Goldman & Novaga (2015): pointwise geometric meaning to $z \cdot D\chi_E$. If $d = 2$ or $d = 3$, if $g = -\mathrm{div}(z) \in L^d(\Omega)$ is a variational mean curvature for the set E, then any point $x \in \partial^* E$ is a Lebesgue point of z and $z(x) = \nu_{\partial E}(x)$.

Regularity of the level set boundaries

For every level set $F_t = \{\rho_1 > t\}$ with $t \geq 0$,

$$\mathrm{Per}(F_t) = \int_{F_t} \mathrm{div}(z) \text{ and } F_t \in \underset{G \subset \Omega}{\mathrm{argmin}} \left\{ \mathrm{Per}(G) - \int_G \mathrm{div}(z) \right\}.$$

So, $-\mathrm{div}(z)$ is a variational mean curvature of F_t.

Theorem (Carlier & P. 2017)

If ρ_1 solves $(\mathcal{P}_\tau)$, then for every $t > 0$, the level set $F_t = \{\rho_1 > t\}$ has the property that its reduced boundary, $\partial^* F_t$ is a $C^{1,\frac{1}{2}}$ hypersurface and $(\partial F_t \setminus \partial^* F_t) \cap \Omega$ has Hausdorff dimension less than $d - 8$.

Step functions remain step functions

Theorem (Carlier & P. 2017)

Let $d = 1$, $\Omega = (a, b)$ and ρ_0 be a step function with at most N-discontinuities i.e.:

$$\rho_0 := \sum_{j=0}^{N} \alpha_j \chi_{[a_j,\, a_{j+1})},\ a_0 = a < a_1 \cdots < a_N < a_{N+1} = b,$$

then the solution ρ_1 of $(\mathcal{P}_\tau)$ is also a step function with at most N discontinuities.

Step functions remain step functions

Theorem (Carlier & P. 2017)

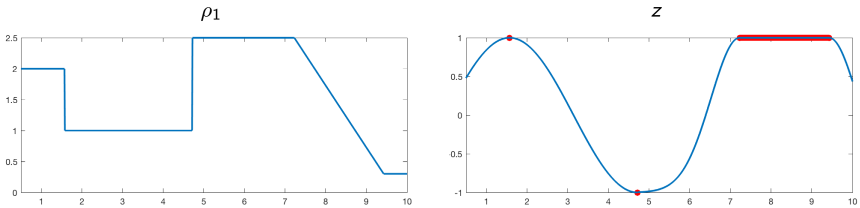
Let $d = 1$, $\Omega = (a, b)$ and ρ_0 be a step function with at most N-discontinuities i.e.:

$$\rho_0 := \sum_{j=0}^{N} \alpha_j \chi_{[a_j,\ a_{j+1})},\ a_0 = a < a_1 \cdots < a_N < a_{N+1} = b,$$

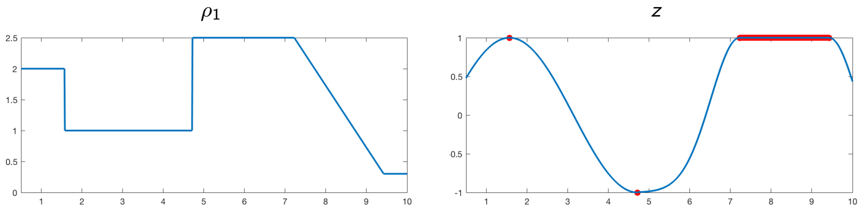
then the solution ρ_1 of $(\mathcal{P}_\tau)$ is also a step function with at most N discontinuities.

- Reduce the problem to the case $\rho_0 > \alpha > 0$ a.e.
- There exists $z \in W^{3,\infty}$ such that $|z| \leq 1$ such that $J(\rho_1) = \int_a^b z' \rho_1 = -\int_a^b z \cdot D\rho_1$.

Step functions remain step functions

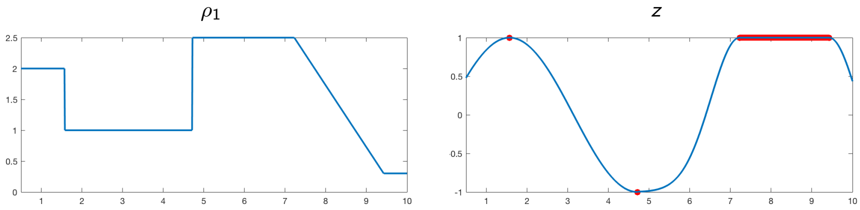


Step functions remain step functions



- $T(x) = x - \varphi'(x) = x + \tau z''(x)$ is the optimal transport from ρ_1 to ρ_0. So, $T(x) = x$ whenever $z''(x) = 0$. Since z achieves its extremal values on $A := \mathrm{Supp}(D\rho_1)$, $z' = 0$ on A. So, $z'' = 0$ on the limit points of A.

Step functions remain step functions



- $T(x) = x - \varphi'(x) = x + \tau z''(x)$ is the optimal transport from ρ_1 to ρ_0. So, $T(x) = x$ whenever $z''(x) = 0$. Since z achieves its extremal values on $A := \mathrm{Supp}(D\rho_1)$, $z' = 0$ on A. So, $z'' = 0$ on the limit points of A.
- Decompose $\mu = D\rho_1$ into its atomic and nonatomic parts, then

$$\mu = \sum_{x \in J} \mu(\{x\})\delta_x + \tilde{\mu}.$$

$\tilde{A} = \mathrm{Supp}(\tilde{\mu})$ is the limit points of $A \implies T(x) = x$ on $\tilde{A} \implies \rho_0 = \rho_1$ on $\tilde{A}$.

Convergence of the TV-JKO scheme in 1D

Let $\Omega = (0,1)$ and let $\rho_0 \in \mathcal{P}_{ac}(\Omega) \cap BV(\Omega)$ with $\rho_0 \geq \alpha > 0$ a.e. on Ω.

Fix T and for small τ, define

$$\rho_0^\tau = \rho_0,\ \rho_{k+1}^\tau \in \operatorname{argmin}\left\{ \frac{1}{2\tau} W_2^2(\rho_k^\tau, \rho) + J(\rho),\ \rho \in \mathrm{BV} \cap \mathcal{P}_{\mathrm{ac}}((0,1)) \right\}$$

for $k = 0, \dots N_\tau$ with $N_\tau := [\frac{T}{\tau}]$.

Since ρ_0 is uniformly bounded from above (as an element of BV) and away from 0,

$$M := \|\rho_0\|_{L^\infty} \geq \rho_k^\tau \geq \alpha.$$

Define piecewise constant interpolation:

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

Convergence of the TV-JKO scheme in 1D

Definition

A weak solution of

$$\partial_t \rho + \left(\rho\left(\frac{\rho_x}{|\rho_x|}\right)_{xx}\right)_x = 0,\ (t,x) \in (0,T) \times (0,1),\ \rho_{|_{t=0}} = \rho_0,$$

is a $\rho \in L^\infty((0,T), \mathrm{BV}((0,1))) \cap C^0((0,T),(\mathcal{P},W_2))$ *such that there exists* $z \in L^\infty((0,T) \times (0,1)) \cap L^2((0,T), H^2 \cap H^1_0((0,1)))$ *with*

$$\|z(t,.)\|_{L^\infty} \leq 1 \text{ and } J(\rho(t,.)) = \int_0^1 z_x(t,x)\rho(x)dx,\ \text{for a.e. } t \in (0,T),$$

and ρ *is a weak solution of*

$$\partial_t \rho - (\rho z_{xx})_x = 0,\ \rho_{|_{t=0}} = \rho_0,\ \rho z_{xx} = 0 \text{ on } (0,T) \times \{0,1\}.$$

i.e. for every $u \in C^1_c([0,T) \times [0,1])$

$$\int_0^T \int_0^1 (\partial_t u\ \rho - (\rho z_{xx})u_x)dxdt = -\int_0^1 u(0,x)\rho_0(x)dx.$$

Convergence of the TV-JKO scheme in 1D

Theorem (Carlier & P. 2017)

There exists a vanishing sequence of time steps $\tau_n \to 0$ *such that the sequence* ρ^{τ_n} *converges strongly in* $L^p((0,T)\times(0,1))$ *for any* $p\in[1,+\infty)$ *and in* $C^0((0,T),(\mathcal{P}([0,1]),W_2))$ *to* $\rho\in L^\infty((0,T),\mathrm{BV}((0,1)))\cap C^0((0,T),(\mathcal{P}([0,1]),W_2))$*, a weak solution of*

$$\partial_t\rho+\left(\rho\left(\frac{\rho_x}{|\rho_x|}\right)_{xx}\right)_x=0,\ (t,x)\in(0,T)\times(0,1),\ \rho_{|_{t=0}}=\rho_0,$$

The proof is fairly standard, let us make some remarks:

- By construction, one has

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

  Using the Aubin-Lions Compactness Theorem of Savaré and Rossi (2003), refinement of Arzèla Ascoli and fact that $BV(0,1)$ compactly embeds into $L^p((0,1))$ for all $p\in[1,\infty)$,

$$\rho^\tau\to\rho \text{ a.e. in } (0,T)\times(0,1) \text{ and in } L^p((0,T)\times(0,1)),\ \forall p\in[1,+\infty)$$

  and

$$\sup_{t\in[0,T]}W_2(\rho^\tau(t,.),\rho(t,.))\to 0 \text{ as } \tau\to 0,$$

  for some limit curve $\rho\in C^{0,\frac{1}{2}}((0,T),(\mathcal{P}([0,1]),W_2))\cap L^p((0,T)\times(0,1))$. By the uniform bounds on the JKO iterates, $M\ge\rho\ge\alpha$.

Convergence of the TV-JKO scheme in 1D

- for each $k = 0, \dots, N_\tau$, there exists $z_k^\tau \in W^{2,\infty}((0,1))$ such that

$$\|z_k^\tau\|_{L^\infty} \leq 1,\ z_k^\tau(0) = z_k^\tau(1) = 0,\ J(\rho_k^\tau) = \int_0^1 (z_k^\tau)_x \rho_k^\tau,$$

and $T_{k+1}^\tau = \mathrm{id} + \tau(z_{k+1}^\tau)_{xx}$ is the optimal transport T_{k+1}^τ from ρ_{k+1}^τ to ρ_k^τ.

$$\begin{aligned} W_2^2(\rho_k^\tau, \rho_{k+1}^\tau) &= \int_0^1 (x - T_{k+1}^\tau(x))^2 \rho_{k+1}^\tau(x) \mathrm{d}x \\ &= \tau^2 \int_0^1 (z_{k+1}^\tau)_{xx}^2 \rho_{k+1}^\tau(x) \mathrm{d}x \geq \alpha\tau^2 \int_0^1 (z_{k+1}^\tau)_{xx}^2 \mathrm{d}x \end{aligned}$$

Let z^τ be the piecewise constant interpolation of z_k^τ. We thus get an $L^2((0,T), H^2((0,1))$ bound

$$C \geq \|z^\tau\|_{L^2((0,T),H^2((0,1)))}.$$

Convergence of the TV-JKO scheme in 1D

- for each $k = 0, \dots, N_\tau$, there exists $z_k^\tau \in W^{2,\infty}((0,1))$ such that

$$\|z_k^\tau\|_{L^\infty} \le 1,\ z_k^\tau(0) = z_k^\tau(1) = 0,\ J(\rho_k^\tau) = \int_0^1 (z_k^\tau)_x \rho_k^\tau,$$

and $T_{k+1}^\tau = \mathrm{id} + \tau(z_{k+1}^\tau)_{xx}$ is the optimal transport T_{k+1}^τ from ρ_{k+1}^τ to ρ_k^τ.

$$W_2^2(\rho_k^\tau, \rho_{k+1}^\tau) = \int_0^1 (x - T_{k+1}^\tau(x))^2 \rho_{k+1}^\tau(x)\mathrm{d}x$$
$$= \tau^2 \int_0^1 (z_{k+1}^\tau)_{xx}^2 \rho_{k+1}^\tau(x)\mathrm{d}x \ge \alpha\tau^2 \int_0^1 (z_{k+1}^\tau)_{xx}^2 \mathrm{d}x$$

Let z^τ be the piecewise constant interpolation of z_k^τ. We thus get an $L^2((0,T), H^2((0,1))$ bound

$$C \ge \|z^\tau\|_{L^2((0,T),H^2((0,1)))}.$$

So, there exists $z \in L^\infty((0,T)\times(0,1)) \cap L^2((0,T), H^2((0,1)))$ with $\|z\|_{L^\infty} \le 1$ such that

- $\rho^\tau z_{xx}^\tau \rightharpoonup \rho z_{xx}$ in $L^1((0,T)\times(0,1))$.
- $z^\tau \rightharpoonup z$ in $L^2((0,T), H^2((0,1)))$, and weakly $*$ in $L^\infty((0,T)\times(0,1))$.
- $J(\rho(t,\cdot)) = \int_0^1 z_x(t,x)\rho(x)\mathrm{d}x$ for a.e. $t \in (0,T)$.

Convergence of the TV-JKO scheme in 1D

- for each $k = 0, \dots, N_\tau$, there exists $z_k^\tau \in W^{2,\infty}((0,1))$ such that

$$\|z_k^\tau\|_{L^\infty} \leq 1,\ z_k^\tau(0) = z_k^\tau(1) = 0,\ J(\rho_k^\tau) = \int_0^1 (z_k^\tau)_x \rho_k^\tau,$$

and $T_{k+1}^\tau = \mathrm{id} + \tau (z_{k+1}^\tau)_{xx}$ is the optimal transport T_{k+1}^τ from ρ_{k+1}^τ to ρ_k^τ.

$$W_2^2(\rho_k^\tau, \rho_{k+1}^\tau) = \int_0^1 (x - T_{k+1}^\tau(x))^2 \rho_{k+1}^\tau(x)\mathrm{d}x$$

$$= \tau^2 \int_0^1 (z_{k+1}^\tau)_{xx}^2 \rho_{k+1}^\tau(x)\mathrm{d}x \geq \alpha\tau^2 \int_0^1 (z_{k+1}^\tau)_{xx}^2 \mathrm{d}x$$

Let z^τ be the piecewise constant interpolation of z_k^τ. We thus get an $L^2((0,T), H^2((0,1))$ bound

$$C \geq \|z^\tau\|_{L^2((0,T),H^2((0,1)))}.$$

So, there exists $z \in L^\infty((0,T)\times(0,1)) \cap L^2((0,T), H^2((0,1)))$ with $\|z\|_{L^\infty} \leq 1$ such that

  - $\rho^\tau z_{xx}^\tau \rightharpoonup \rho z_{xx}$ in $L^1((0,T)\times(0,1))$.
  - $z^\tau \rightharpoonup z$ in $L^2((0,T), H^2((0,1)))$, and weakly $*$ in $L^\infty((0,T)\times(0,1))$.
  - $J(\rho(t,\cdot)) = \int_0^1 z_x(t,x)\rho(x)\mathrm{d}x$ for a.e. $t \in (0,T)$.

- By standard computations, let $u \in C_c^1([0,T)\times[0,1])$ then

$$\int_0^T \int_0^1 \partial_t u\, \rho^\tau - (\rho^\tau z_{xx}^\tau) u_x \mathrm{d}x\mathrm{d}t = -\int_0^1 u(0,x)\rho_0(x)\mathrm{d}x + R_\tau(u).$$

Taking the limit concludes this proof.

Summary

We discussed some properties of the JKO iterates:

1. a maximum principle (and a minimum principle in 1D).
2. establish the optimality conditions.
3. regularity of the level sets.
4. analysis for step function initial data.

Thanks to the minimum principle, we have convergence of the JKO scheme as $\tau \to 0$ in 1D, for strictly positive initial density ρ_0.

Preprint on arXiv: *On the total variation Wasserstein gradient flow and the TV-JKO scheme.*

Summary

We discussed some properties of the JKO iterates:

1. a maximum principle (and a minimum principle in 1D).
2. establish the optimality conditions.
3. regularity of the level sets.
4. analysis for step function initial data.

Thanks to the minimum principle, we have convergence of the JKO scheme as $\tau \to 0$ in 1D, for strictly positive initial density ρ_0.

Preprint on arXiv: *On the total variation Wasserstein gradient flow and the TV-JKO scheme.*

Thank you for your attention.