CONVEXITY OF THE SUPPORT OF THE DISPLACEMENT INTERPOLATION: COUNTEREXAMPLES

FILIPPO SANTAMBROGIO, XU-JIA WANG

ABSTRACT. Given two smooth and positive densities ρ_0, ρ_1 on two compact convex sets K_0, K_1, respectively, we consider the question whether the support of the measure ρ_t obtained as the geodesic interpolant of ρ_0 and ρ_1 in the Wasserstein space $\mathbb{W}_2(\mathbb{R}^d)$ is necessarily convex or not. We prove that this is not the case, even when ρ_0 and ρ_1 are uniform measures.

1. Introduction

One of the main features of optimal transport theory (we refer to [16] and [15] for a general presentation) is the fact that it provides an original and efficient way to define interpolations between probability measures.

Given a domain $\Omega \subset \mathbb{R}^d$ (that we take compact for simplicity and convex for the sake of the interpolation), we define the space $\mathbb{W}_2(\Omega)$ as the space of probabilities on Ω endowed with the distance W_2, defined through

$$W_2^2(\mu,\nu) = \min\left\{ \int_{\Omega\times\Omega} |x-y|^2 \, \mathrm{d}\gamma \,:\, \gamma \in \Pi(\mu,\nu) \right\},$$

where $\Pi(\mu,\nu)$ is the set of the so-called *transport plans*, i.e.

$$\Pi(\mu,\nu) = \{\gamma \in \mathcal{P}(\Omega\times\Omega) : \, (\pi_x)_\#\gamma = \mu, \, (\pi_y)_\#\gamma = \nu, \},$$

where $\pi_x(x,y) := x$ and $\pi_y(x,y) = y$ are the standard projections on the two factors of $\Omega \times \Omega$.

It is possible to prove that the above minimization problem has a solution, which is unique and of the form $\gamma = (id, T)_\#\mu$ (i.e. it is concentrated on the graph of a map $T : \Omega \to \Omega$, called *optimal transport map*) in many situations (in particular if $\mu \ll \mathcal{L}^d$) and that W_2, i.e. the square root of the minimal value above, is indeed a distance on $\mathcal{P}(\Omega)$.

The space $\mathbb{W}_2(\Omega)$ can be checked to be a geodesic space, where for $\mu, \nu \in \mathcal{P}(\Omega)$ the (unique) geodesic curve connecting them is obtained through

$$\rho_t := ((1-t)\,\mathrm{id} + tT)_\#\mu, \tag{1.1}$$

where T is the optimal transport map from μ to ν. This provides a useful interpolation between μ and ν which is very different from the linear interpolation $(1-t)\mu + t\nu$. In particular, it is useful in many applications, for instance in image processing: this is the case when doing histograms interpolations (roughly speaking, the average between a distribution of pixel colors which are almost white and another were they are almost black should be a a distribution with intermediate grey pixels, and not one with half pixels which are white and half which are black), or when artificially creating intermediate images between two pictures representing a same object which ahs moved (and the goal is to find the same object in the middle, instead of two half objects at the starting and arrival spots). But this very interpolation has also a lot of mathematical applications, as it was first pointed out by McCann in [14]. Indeed, McCann found a class of functionals $F : \mathcal{P}(\Omega) \to \mathbb{R}$ which are convex along these geodesic lines (but not necessarily convex in the usual sense, think for instance at $\mu \mapsto \int\int |x-y|^2 \mathrm{d}\mu(x)\mathrm{d}\mu(y)$), thus making possible to obtain uniqueness results or sufficient optimality conditions (see [4], for instance) for variational problems involving them. Also, this notion of convexity, called *displacement convexity* (the above interpolation is also called *displacement interpolation*) is an important notion in the study of gradient flows of these functionals (see [2]). We also recall that the interpolation ρ_t can be found numerically via one of the most classical algorithm for optimal transport, the so-called Benamou-Brenier method [3]. By solving the kinetic energy minimization problem

$$\min\left\{\int_0^1\int_\Omega |v_t|^2 \mathrm{d}\rho_t \, \mathrm{d}t \ : \ \partial_t\rho_t + \nabla\cdot(v_t\rho_t) = 0\right\},$$

among curves of measures with given initial and final data, one recovers the above interpolation, and the optimal velocity field v_t allows to find T (as we have $v_t(x) = (T - \mathrm{id}) \circ (T_t)^{-1}$ with $T_t := (1-t)\,\mathrm{id} + tT$).

Finally, let us remark that the interpolated measures ρ_t can also be considered as weighted barycenters between ρ_0 and ρ_1, as they solve the minimization problem

$$\min\left\{(1-t)W_2^2(\rho,\rho_0) + tW_2^2(\rho,\rho_1) \ : \ \rho\in\mathcal{P}(\Omega)\right\},$$

a minimization problem which has ρ_t as unique solution (provided one of the measures ρ_0, ρ_1 is absolutely continuous). This problem has also been considered when more than two measures are given, in order to find the weighted barycenter of many of them, solving (see [1])

$$\min\left\{\sum_{i=1}^n \lambda_i W_2^2(\rho,\rho_i) \ : \ \rho\in\mathcal{P}(\Omega)\right\} \tag{1.2}$$

(a minimization problem that can be also recast in the setting of multi-marginal problems, as in [11]). Note that when $n \geq 3$ this variational definition of the barycenter is the only possible one, differentely from the case $n = 2$ where one can simply use the geodesic in $\mathbb{W}_2$.

In optimal transport theory, and in particular for regularity issues, the convexity of the support of a measure is an important fact, and some results are only available under this assumption. In particular, the optimal transport map between two smooth densities ρ_0, ρ_1 is smooth itself, provided their supports are convex (this theory was first developed by Caffarelli, [7, 9, 8], see [10] for a survey). Since it is possible to write optimality conditions for the minimizers of (1.2) in terms of the optimal maps T_i sending the optimal ρ onto the various ρ_i (which actually amounts to a system of Monge-Ampère equations), one could imagine some boothstrap strategy to study the regularity of ρ using that of the T_i. Yet, this requires as a preliminary fact to check the convexity of $\mathrm{spt}(\rho)$! Note that this is not needed for $n = 2$, since formula (1.1) allows to deduce the regularity of ρ_t from that of T, which is the optimal map from ρ_0 to ρ_1: in particular, only the convexity of the supports of these two measures is needed. Anyway, the question whether the support of the barycenter is guaranteed to be convex whenever the supports of the ρ_i are so is a very natural question. The case $n = 2$ is the easiest to study and if the answer is no in such a case there is no hope to get yes in the case $n \geq 3$.

Moreover, independently of the regularity motivation, this question is very natural from a geometric point of view, as a question on convex bodies. One can for instance insist on the case where ρ_0 and ρ_1 are uniform measures on convex sets and look at $\mathrm{spt}(\rho_t)$ as an interpolation between these sets. More generally, one could wonder if the interpolation between log-concave densities (of the form $\rho = e^{-V}$ for a convex function $V : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$) is still log-concave. This is true in dimension 1 (and, contrarily to the question on the supports, it is non-trivial[1]).

Unfortunately, the present paper shows that the answer to these questions is no. More precisely, we will see (Theorem 1) that on any smooth convex domain on $\mathbb{R}^2$ we can find smooth and positive densities ρ_0, ρ_1 such that the support of $\rho_{1/2}$ is not convex, and that this can also be done with uniform measures on (different, of course) convex sets.

Acknowledgments This work has been performed during a visit of the first author to MSI at ANU. The support of MSI is warmly acknowledged.

2. The case of two segments or two curves

Consider two Lipschitz curves $\omega_0, \omega_1 : [a, b] \to \mathbb{R}^d$, and two measures $\nu_0, \nu_1 \in \mathcal{P}([a, b])$. We need to discuss the optimal transport between the measures ρ_0 and ρ_1, defined through $\mu_i := (\omega_i)_\# \nu_i$ for $i = 0, 1$, in connection with the optimal transport between ν_0 and ν_1. Note that, if $d > 1$, these measures fall out of the usual assumptions for the existence of an optimal transport map between them, as it is not true that μ_0 does not give mass to "small" (i.e. $(d-1)-$dimensional, or smaller) sets (see [5, 13] and, more recently, [12], for the exact conditions for this existence result).

Yet, the fact that *both* measures are one-dimensional, allows to act differently (by the way, let us mention a recent preprint, [6], on the case where optimal transport between measures in different dimensions is considered). We will suppose that ω_0 and ω_1 are such that $\dot\omega_0(t) \cdot \dot\omega_1(s) > 0$ for a.e. $(t, s) \in [a, b]^2$. In this case what happens is that optimal transport plans $\gamma \in \Pi(\mu_0, \mu_1)$ for the quadratic cost $|x - y|^2$ are just obtained as the image of optimal transport plans in $\Pi(\nu_0, \nu_1)$ through the map $(t, s) \mapsto (\omega_0(t), \omega_1(s))$.

Indeed, it is well-known that optimal transport plans for the quadratic cost are such that their support is monotone, in the sense that $(x, y), (x', y') \in \operatorname{spt}(\gamma)$ implies $(x - x') \cdot (y - y') \geq 0$. Yet, for $(x, y) = (\omega_0(t), \omega_1(s))$ and $x' = (\omega_0(t'), \omega_1(s'))$, we have

$$(x - x') \cdot (y - y') = \left(\int_t^{t'} \dot\omega_0(\tau)\mathrm{d}\tau\right) \cdot \left(\int_s^{s'} \dot\omega_1(\sigma)\mathrm{d}\sigma\right) = \int_t^{t'} \int_s^{s'} \dot\omega_0(\tau) \cdot \dot\omega_1(\sigma)\, \mathrm{d}\sigma\mathrm{d}\tau,$$

and, since the integrand in the last expression is positive, the sign of the integral is the same as the sign of $(t'-t)(s'-s)$. This proves that, if we write $\gamma = (\omega_0 \times \omega_1)_\# \tilde\gamma$, the monotonicity of the support of γ implies the 1D monotonicity of that of $\tilde\gamma$. In particular, when $\nu_0 \ll \mathcal{L}^1$, the optimal transport plan in $\Pi(\nu_0, \nu_1)$ is unique and easy to find, and it is induced by a map, the unique monotone increasing map T sending ν_0 onto ν_1; as a consequence, there is a unique optimal transport plan from μ_0 to μ_1, and it is induced by the map $\omega_1 \circ T \circ \omega_0^{-1}$.

We just point out that the situation is even simpler in case ω_0 and ω_1 are affine functions, i.e. their images are two segments S_0 and S_1. Up to transations, we can suppose that S_0 and S_1 are of the form $S_i = \{tv_i, t \in [0, 1]\}$, for some vectors $v_i \in \mathbb{R}^d$ $(i = 0, 1)$. Translating the segments is not restrictive as we know that optimal transport and translations commute.

Then, we distinguish three cases: either $v_0 \cdot v_1 = 0$, or $v_0 \cdot v_1 > 0$, or $v_0 \cdot v_1 < 0$. In the first case it is easy to check that, for every $x \in S_0$ and $y \in S_1$, we have $|x - y|^2 = |x|^2 + |y|^2$. This implies $\int |x - y|^2\, \mathrm{d}\gamma = \int |x|^2 \mathrm{d}\mu_0 + \int |y|^2 \mathrm{d}\mu_1$ for every $\gamma \in \Pi(\mu_0, \mu_1)$, and hence every transport plan between these two measures is optimal (in particular, there is no uniqueness). In the case $v_0 \cdot v_1 > 0$, we have, for $x = tv_0$, $y = sv_1$ and $c = v_0 \cdot v_1$,

$$|x - y|^2 = c|t - s|^2 + (|v_0|^2 - c)t^2 + (|v_1|^2 - c)s^2,$$

[1] This result has been proven by Young-Heon Kim via un-published computations, in the hope of generalizing to higher dimensions.

where the first term in the right hand side is a positive multiple of the quadratic cost in the variables (t,s) and the second term is separable. The case $v_0 \cdot v_1 < 0$ can be reduced to the previous one up to replacing v_0 with $-v_0$ and translating again. This means that, whenever v_0 and v_1 are not orthogonal, the optimal transport between the two measures can be computed as if we were in dimension one, i.e. it is the monotone increasing map. The orientation of the segments (which is crucial to define the monotonicity of the map) has to be chosen so that the scalar product is positive.

With these considerations in mind, we consider the following esample.

Let $S_0 = \{(t, at) : t \in [-1, 1]\}$ and $S_1 = \{(t, -at) : t \in [-1, 1]\}$ be two segments in $\mathbb{R}^2$ with middle point at the origin, and connecting $(-1, -a)$ to $(1, a)$ and $(-1, a)$ to $(1, -a)$, respectively. We select $a \in]0, 1[$ so that the correct orientation of the two segments is the one based on the abscissas. Then we consider a probability measure μ_0 on S_0 with linear density, proportional to $1 - t$, and μ_1 a probability measure on S_1 with density proportional to $t + 1$. A simple calculation provides the formula for the optimal transport map

$$T(t, at) = (-1 + \sqrt{4 - (1 - t)^2}, a(1 - \sqrt{4 - (1 - t)^2}))$$

and the shape of the interpolated measure is represented in the following figure. The fact that this interpolation is supported on a curve which is not a segment (it is actually part of the boundary of an ellipse) will be the key point for our counter-examples.

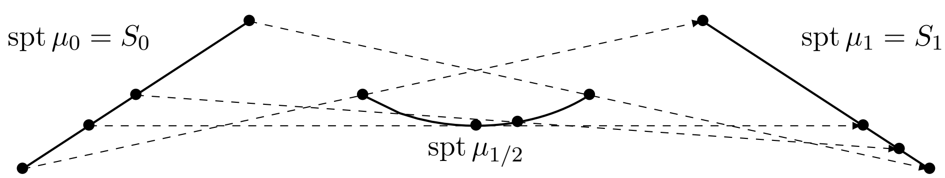

FIGURE 1. Support of $\mu_{1/2}$ when μ_0 and μ_1 have linear densities on segments S_0 and S_1.

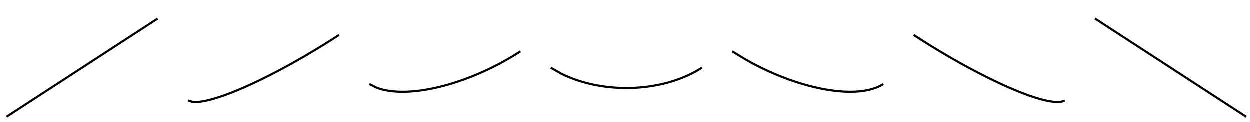

FIGURE 2. The supports of the interpolations μ_t for $t = k/6$.

3. Counterexamples with convex bodies

In this section we use the previous considerations to provide a wide class of cases where two measures ρ_0 and ρ_1 have smooth and positive densities on their supports, their supports are convex sets, but the support of the interpolant $\rho_{1/2}$ cannot be convex.

Geometric setting. We consider now a convex set $\Omega \subset \mathbb{R}^2$ and two portions Γ_0 and Γ_1 of its boundary, and we suppose that, in suitable coordinates, we have $\Gamma_0 = \{(t, f(t)) : t \in I\}$ and $\Gamma_1 = \{(t, g(t)) : t \in J\}$ for some disjoint intervals $I = [a, b]$ and $J = [c, d]$ and $f : I \to \mathbb{R}$, $g : J \to \mathbb{R}$ Lipschitz functions, with $0 < f' < \lambda$ and $0 > g' > -\lambda$ for a suitable constant $\lambda \in (0, 1)$. We suppose $f(b) = g(c)$ and $f(a) = g(d)$. Let us call $x = (a, f(a))$ and $x' = (b, f(b))$ the endpoints of Γ_0, and $y = (c, g(c))$ and $y' = (d, g(d))$ those of Γ_1. We also take two points $x'' \in \Gamma_0$ and $y'' \in \Gamma_1$, with the same ordinate, smaller than $(f(a) + f(b))/2 = (g(c) + g(d))/2$. We consider two measures

$\mu_0 \in \mathcal{P}(\Gamma_0)$ and $\mu_1 \in \mathcal{P}(\Gamma_1)$ such that the optimal transport plan between them exists, is unique, and is induced by a map T with $T(x) = y$, $T(x') = y'$ and $T(x'') = y''$.

The situation is the one sketched in the picture below.

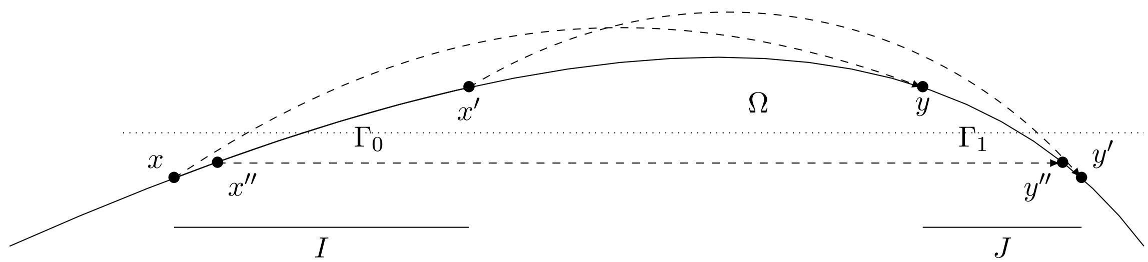

FIGURE 3. The configuration of Ω, Γ_0, Γ_1.

We now consider two sequence of densities ρ_0^n and ρ_1^n weakly converging to μ_0 and μ_1, respectively, and concentrated on Ω (we will see different examples later where their support can be, or not, the whole domain Ω). Let us call $\rho_{1/2}^n$ the interpolated density between them at time $t = 1/2$. We will prove that, for n large enough, it is not possible that the support of $\rho_{1/2}^n$ is convex.

Lemma 1. *Let Ω be a convex domain and μ_0 and μ_1 two measures on the curves Γ_0 and Γ_1 as in the geometric setting above. Let ρ_0^n and ρ_1^n be smooth densities weakly converging to μ_0 and μ_1, respectively, and concentrated on Ω. Let $\rho_{1/2}^n$ be the middle point of the geodesic in $\mathbb{W}_2(\Omega)$ between them. Then, for n large enough, the support of $\rho_{1/2}^n$ is not convex.*

Proof. Consider the optimal transport plan $\gamma^n \in \Pi(\rho_0^n, \rho_1^n)$ for the quadratic cost. By uniqueness of the optimal plan γ from μ_0 to μ_1, it is clear that $\gamma^n \rightharpoonup \gamma$. From the fact that the support of γ is included in the Hausdorff limit of the supports of γ^n, we deduce the existence of points $(x_n, y_n), (x'_n, y'_n), (x''_n, y''_n) \in \mathrm{spt}(\gamma^n)$ converging to $(x, y), (x', y')$ and (x'', y''), respectively. Since $((x_n + y_n)/2, (x'_n + y'_n)/2)$ belong to the support of $\rho_{1/2}^n$, if this support were convex, it should also contain $p_n := (x_n + y_n + x'_n + y'_n)/4$.

For simplicity, and without loss of generality, we will suppose that $(x + y + x' + y')/4 = 0$ (this is possible up to a translation).

Now, suppose $p_n \in \mathrm{spt}(\rho_{1/2}^n)$. This means that there exist $z_n, w_n \in \Omega$ such that $(z_n + w_n)/2 = p_n$ and $(z_n, w_n) \in \mathrm{spt}(\gamma^n)$. Note that the monotonicity of $\mathrm{spt}(\gamma^n)$ implies the inequality

$$(w_n - y''_n) \cdot (z_n - x''_n) \geq 0.$$

Up to subsequences, we can pass to the limit and obtain the existence of two points $z, w \in \bar{\Omega}$ such that $z + w = 0$ and $(w - y'') \cdot (z - x'') \geq 0$.

Using $w = -z$ this translates into $z \in \bar{\Omega} \cap (-\bar{\Omega})$ and

$$\left| z - \frac{x'' - y''}{2} \right|^2 \leq \left| \frac{x'' + y''}{2} \right|^2.$$

This means that z belongs to the intersection of $\bar{\Omega}$, $-\bar{\Omega}$, and a ball centered at $q := (x'' - y'')/2$ with radius $|(x'' + y'')/2|$. But it is possible to see that this intersection is empty, as one can observe from the next figures.

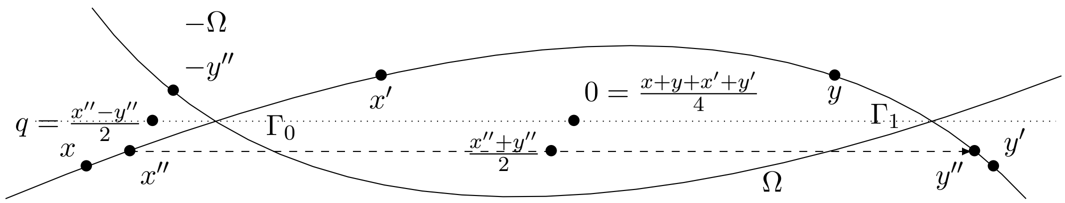

FIGURE 4. Ω, $-\Omega$, and the reflected points.

Indeed, the closest point of $\bar\Omega \cap (-\bar\Omega)$ to q is the point $\tilde q$ on the intersection of horizontal line through 0 with $\partial\Omega$, yet it does not belong to the required ball, as the segment $[x'', -y'']$ is a diameter of this ball, and $\tilde q$ sees both segments $[x'', q]$ and $[q, -y'']$ under an angle smaller than 45° because of the derivative condition on f and g. □

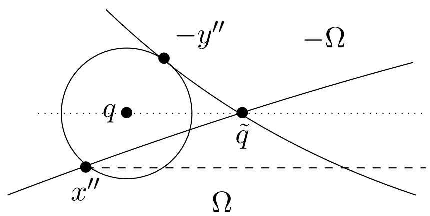

FIGURE 5. A zoom around q and $\tilde q$.

We can now provide our main counter-examples.

Theorem 1. *For every smooth convex body $\Omega \subset \mathbb{R}^2$ there exist two smooth probability densities ρ_0 and ρ_1 in $\mathcal{P}(\Omega)$, strictly positive on Ω, such that the support of their interpolation $\rho_{1/2}$ is not convex.*

Proof. It is enough to notice that it is always possible to find two portions of boundary $\Gamma_0, \Gamma_1 \subset \partial\Omega$ such that the assumptions of Lemma 1 (i.e. those of the above geometrical setting) are satisfied (it is enough to choose Γ_0 and Γ_1 close enough to an extremal point of Ω, and orientate the coordinate system so that this point is the unique point of Ω with maximal ordinate and to choose suitable non-uniform probabilities μ_0 and μ_1 on Γ_0 and Γ_1). Then, we take any smooth and strictly positive approximations ρ_0^n and ρ_1^n and choose $\rho_0 = \rho_0^n$ and $\rho_1 = \rho_1^n$, for n large enough. □

Remark 1. The same procedure can be applied to any convex domain Ω, provided that there exist portions of its boundary with tangent vectors forming angle strictly smaller than 90°, which allows to choose a suitable coordinate system and fit the geometrical setting. The only convex domains in $\mathbb{R}^2$ where this is not possible are the rectangles and the acute (or right) triangles.

Theorem 2. *There exist convex sets $A, B \subset \mathbb{R}^2$ such that, setting $\rho_0 = |A|^{-1}\mathcal{L}^2 \llcorner A$ and $\rho_1 = |B|^{-1}\mathcal{L}^2 \llcorner B$, the support of the middle point interpolation $\rho_{1/2}$ is not convex.*

Proof. It is enough to choose $a \in (0,1)$ and use, in Lemma 1, $\Omega = \{(t,s) \in \mathbb{R}^2 : s \le -a|t|, s \ge -1, |t| \le 1\}$, $\Gamma_0 = [(-1,-a),(0,0)]$, $\Gamma_1 = [(1,-a),(0,0)]$, and measures μ_0 and μ_1 supported on Γ_0 and Γ_1 and constructed as in the example at the end of the previous section. Then we take A_n to be the triangle with vertices $(0,0)$, $(-1,-a)$ and $(-1,-a-\frac{1}{n})$, B_n the triangle with vertices $(0,0)$, $(1,-a)$ and $(1,-a-\frac{1}{n})$, $\rho_0^n = |A_n|^{-1}\mathcal{L}^2 \llcorner A_n$ and $\rho_1^n = |B_n|^{-1}\mathcal{L}^2 \llcorner B_n$. These measures converge to μ_0 and μ_1, respectively. Hence, for n large enough we can take $A = A_n$ and $B = B_n$. □

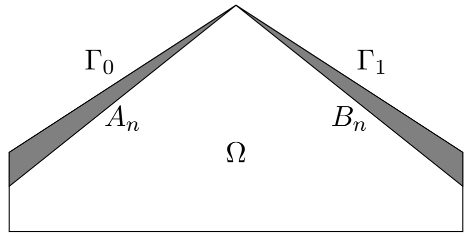

Figure 6. Counter-example with uniform measures.

Remark 2. Note that the domains A_n and B_n in the above proof can be smoothened, and made strictly convex, with no extra difficulty.

Acknowledgment The first author acknowledges the hospitality of the MSI of ANU and the partial support of the ANR through the contract ANR- 12-BS01-0014-01 GEOMETRYA.

References

[1] M. Agueh and G. Carlier, Barycenters in the Wasserstein space. *SIAM J. Math. An.*, 43 (2), 904–924, 2011.
[2] L. Ambrosio, N. Gigli and G. Savaré, *Gradient Flows in Metric Spaces and in the Space of Probability Measures*, Lectures in Mathematics, Birkhäuser, 2005.
[3] J.-D. Benamou and Y. Brenier, A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem, *Numer. Math.*, 84, 375–393, 2000.
[4] A. Blanchet, P. Mossay and F. Santambrogio, Existence and uniqueness of equilibrium for a spatial model of social interactions, to appear in *Int. Econ. Rev.*
[5] Y. Brenier, Décomposition polaire et réarrangement monotone des champs de vecteurs. (French) *C. R. Acad. Sci. Paris Sér. I Math.* 305 (19), 805–808, 1987.
[6] P.-A. Chiappori, R. J. McCann and B. Pass Multi- to one-dimensional transportation, preprint, 2015.
[7] L. Caffarelli, A localization property of viscosity solutions to the Monge-Ampère equation and their strict convexity. *Ann. of Math.* 131 (1), 129–134, 1990.
[8] L. Caffarelli, Interior $W^{2,p}$ estimates for solutions of the Monge-Ampère equation. *Ann. of Math.* 131 (1), 135–150, 1990.
[9] L. Caffarelli, Some regularity properties of solutions of Monge Ampère equation. *Comm. Pure Appl. Math.* 44 (8-9), 965–969, 1991.
[10] G De Philippis and A Figalli, The Monge–Ampère equation and its link to optimal transportation *Bulletin of the American Mathematical Society*, 2014.
[11] W. Gangbo and A. Święch, Optimal maps for the multidimensional Monge- Kantorovich problem, *Comm. Pure Appl. Math.*, 51 (1), 23–45, 1998.
[12] N. Gigli On the inverse implication of Brenier-McCann theorems and the structure of $(P_2(M), W_2)$, *Meth. Appl. Anal.*, 18 (2), 127–158, 2011.
[13] R. J. McCann, Existence and uniqueness of monotone measure preserving maps, *Duke Math. J.*, 80, 309–323, 1995.
[14] R. J. McCann, A convexity principle for interacting gases. *Adv. Math.* 128 (1), 153–159, 1997.
[15] F. Santambrogio, Optimal Transport for Applied Mathematicians. Birkäuser, NY, to be published in November 2015.
[16] C. Villani, *Optimal Transport. Old and New.* Grundlehren der Mathematischen Wissenschaften, 338. Springer-Verlag, Berlin, 2009.

F. S. Laboratoire de Mathématiques d'Orsay, Univ. Paris-Sud, CNRS, Université Paris-Saclay, 91405 Orsay Cedex, France, filippo.santambrogio@math.u-psud.fr — **X.-J. W.** Centre for Mathematics and its Applications, Australian National University, ACT 0200, Australia, Xu-Jia.Wang@anu.edu.au