



ACADÉMIE
DES SCIENCES
INSTITUT DE FRANCE

Comptes Rendus

Mathématique


Cyril Letrouit

Unstable optimal transport maps

Volume 364 (2026), p. 333-344

Online since: 18 May 2026

<https://doi.org/10.5802/crmath.834>

 This article is licensed under the
CREATIVE COMMONS ATTRIBUTION 4.0 INTERNATIONAL LICENSE.
<http://creativecommons.org/licenses/by/4.0/>



*The Comptes Rendus. Mathématique are a member of the
Mersenne Center for open scientific publishing*
www.centre-mersenne.org — e-ISSN : 1778-3569



Research article
Partial differential equations

Unstable optimal transport maps

Cyril Letrouit ^a

^a Université Paris-Saclay, CNRS, Laboratoire de mathématiques d'Orsay, 91405 Orsay, France

E-mail: cyril.letrouit@universite-paris-saclay.fr

Abstract. The stability of optimal transport maps with respect to perturbations of the marginals is a question of interest for several reasons, ranging from numerical analysis and statistics to the justification of the linearized optimal transport framework. Under various assumptions on the source measure, it is known that optimal transport maps are stable with respect to variations of the target measure.

In this note, we focus on the mechanisms that can, on the contrary, lead to *instability*. We identify two of them. We first show that instability may arise from the unboundedness of the density: we exhibit a source density on the unit ball of \mathbb{R}^d which blows up at two points of the boundary and for which optimal transport maps are highly unstable. Then we prove that even for uniform densities on bounded open sets, optimal transport maps can be rather unstable sufficiently close to configurations where uniqueness of optimal plans is lost.

Keywords. Optimal transport, stability, uniqueness, Wasserstein distance.

2020 Mathematics Subject Classification. 49Q22.

Funding. The author acknowledges the support of the Agence nationale de la recherche, through the PEPR PDE-AI project (ANR-23-PEIA-0004).

Manuscript received 23 October 2025, revised 11 March 2026, accepted 3 April 2026, online since 18 May 2026.

1. Introduction

1.1. Main results

Brenier's theorem is one of the foundational results of optimal transport theory [4,5]. It provides simple conditions guaranteeing the existence of an optimal map transporting a given “source” probability measure ρ on \mathbb{R}^d to another “target” probability measure μ on \mathbb{R}^d , when the cost of transporting mass is the squared Euclidean distance. If $\#$ denotes the pushforward operation on measures, it asserts the ρ -a.e. existence and uniqueness of a solution to the optimization problem (in T)

$$\inf \left\{ \int_{\mathbb{R}^d} |x - T(x)|^2 d\rho(x) \mid T_{\#}\rho = \mu \right\}$$

as soon as ρ, μ belong to the set $\mathcal{P}_2(\mathbb{R}^d)$ of probability measures on \mathbb{R}^d with finite second moment, and ρ is absolutely continuous with respect to the Lebesgue measure. This solution is denoted by $T_{\mu} \in L^2(\rho)$ in the sequel; the dependence on ρ is omitted in the notation since ρ will be fixed throughout.

After existence and uniqueness, the final ingredient for a well-posed problem is *stability*. Brenier [5] proved that his construction is stable with respect to perturbations of the target measure, for a fixed source measure ρ : he showed that the map $\mu \mapsto T_\mu$ is continuous from $(\mathcal{P}_2(\mathbb{R}^d), W_2)$ to $L^2(\rho)$, where W_p denotes the p -Wasserstein distance.

In recent years, several people have tried to quantify this stability, i.e., to provide conditions on ρ that guarantee quantitative bounds on the L^2 -distance between T_μ and T_ν in terms of Wasserstein distances between μ and ν . These quantitative stability bounds have been established under various assumptions on ρ (see the literature review in Section 1.3). They are of the form

$$\forall \mu, \nu \in \mathcal{P}_2(\mathcal{Y}), \quad \|T_\mu - T_\nu\|_{L^2(\rho)} \leq CW_p(\mu, \nu)^\alpha \quad (1)$$

for $C, \alpha > 0$, $p \geq 1$ and $\mathcal{Y} \subset \mathbb{R}^d$. There are multiple motivations for looking at such estimates. For instance, they provide quantitative rates of convergence of numerical approximations of the optimal transport map; they serve as a foundational ground for the statistical estimation of optimal transport maps (see [6, Chapter 3.2] and [2]); and finally they justify theoretically the linearized optimal transport framework (see Section 1.3).

However, it has never been proven that optimal transport maps could be *unstable*. In this note, we fill this gap by elucidating two mechanisms leading to instability, or partial loss of stability, of optimal transport maps. Our results reveal limitations on the generality under which quantitative stability results may hold.

We first construct an absolutely continuous source measure ρ for which optimal transport maps are *highly unstable*. More precisely, in Theorem 1, we construct ρ whose density is bounded below on the (closed) unit ball $B_{\mathbb{R}^d}(0, 1)$ such that for any ball $\mathcal{Y} = B_{\mathbb{R}^d}(0, R)$, any $C, \alpha > 0$ and $p \geq 1$, the bound (1) fails. In what follows, $\mathcal{P}(\mathcal{X})$ denotes the set of probability measures on $\mathcal{X} \subset \mathbb{R}^d$.

Theorem 1. *Let $d \geq 2$. There exists an absolutely continuous $\rho \in \mathcal{P}(B_{\mathbb{R}^d}(0, 1))$ with density bounded below, such that for any ball $\mathcal{Y} = B_{\mathbb{R}^d}(0, R)$ with $R > 0$, any $C, \alpha > 0$ and $p \geq 1$, the following inequality fails:*

$$\forall \mu, \nu \in \mathcal{P}(\mathcal{Y}), \quad \|T_\mu - T_\nu\|_{L^2(\rho)} \leq CW_p(\mu, \nu)^\alpha. \quad (2)$$

Behind this result, the mechanism leading to instability is the blow-up of the density ρ . The density ρ constructed to prove Theorem 1 actually blows up at two points A and A' of the boundary of $B_{\mathbb{R}^d}(0, 1)$, at a carefully chosen rate. It may be expected that such source measure leads to some form of instability, since the extreme case of ρ being of the form $\frac{1}{2}(\delta_A + \delta_{A'})$ generates extreme instabilities — but it is not absolutely continuous and lacks existence of an optimal transport map for generic target measures, whereas in Theorem 1 we look for an absolutely continuous ρ . The results of [10] show that the blow-up rate that we choose is optimal, in the sense that if ρ blows up slightly slower at the boundary of $B_{\mathbb{R}^d}(0, 1)$, then (2) holds for some $C, \alpha > 0$ and $p = 1$ (see Remark 5).

Apart from unbounded densities, if we focus on source densities bounded above and below on bounded open sets, some loss of stability can nevertheless occur close enough to configurations where optimal plans are non-unique. It is actually natural to expect loss of stability just before loss of uniqueness. To illustrate this idea, we show the following result — whose relevance and numerology is explained in Section 1.2 — using as a source measure the *uniform probability measure* ρ on a carefully chosen *bounded open set* $\mathcal{X} \subset \mathbb{R}^d$.

Theorem 2. *Let $d \geq 2$. There exists a bounded open set $\mathcal{X} \subset \mathbb{R}^d$ and a compact set \mathcal{Y} such that if ρ denotes the uniform probability measure on \mathcal{X} , then*

(1) for any $C > 0$, $p \geq 1$ and any $\alpha > \frac{p}{2(p+1)}$ (in particular for $\alpha = \frac{1}{2}$), the following inequality fails:

$$\forall \mu, \nu \in \mathcal{P}(\mathcal{Y}), \quad \|T_\mu - T_\nu\|_{L^2(\rho)} \leq CW_p(\mu, \nu)^\alpha; \tag{3}$$

(2) for any $p \geq 1$ and $\alpha < 1/6$ the inequality (3) holds for some $C > 0$.

The bounded set \mathcal{X} used to prove Theorem 2 is described in Section 3.1. It repeats at smaller and smaller scales a same structure, getting closer at smaller scales to a configuration where uniqueness of optimal transport maps is lost. Nevertheless \mathcal{X} is open, so ρ remains in a setting where Brenier’s theorem applies, and uniqueness is guaranteed. Let us also mention that our results hold not only for the quadratic cost, but also for arbitrary costs of the form $h(|x - y|)$ with $h: \mathbb{R}_+ \rightarrow \mathbb{R}$ increasing and strictly convex, in particular for $h(r) = r^q$, $q > 1$ (see Remark 8).

1.2. Regularity of the map $\mu \mapsto T_\mu$

Our results shed light on the regularity of the map $\mu \mapsto T_\mu$ from $(\mathcal{P}(\mathcal{Y}), W_2)$ to $L^2(\rho)$, which reflects the stability properties of the optimal transport maps with source measure ρ with respect to perturbations of the marginals. Gigli [8, Theorem 5.1] proved in 2011 that in some situations, this map is not better than $\frac{1}{2}$ -Hölder. Since then, it remains an open question to determine under which conditions this regularity is actually achieved: for instance, Gigli [8, Corollary 3.4] (see also [11, Theorem 2.3]) showed that $\frac{1}{2}$ -Hölder regularity is achieved at any μ such that T_μ is Lipschitz. This result plays a key role in statistical optimal transport (see [6, Chapter 3.2] and [2]). It is believed that $\frac{1}{2}$ -Hölder regularity actually holds in greater generality, and in particular the following conjecture has emerged for a few years.

Conjecture 3. *Let ρ be uniform on a compact convex set of \mathbb{R}^d , and let \mathcal{Y} be compact. Then there exists $C > 0$ such that*

$$\forall \mu, \nu \in \mathcal{P}(\mathcal{Y}), \quad \|T_\mu - T_\nu\|_{L^2(\rho)} \leq CW_2(\mu, \nu)^{1/2}.$$

This inequality is known with an exponent 1/6 in the right-hand side (see [7] and Theorem 4 below), and the difficulty is therefore to prove it with the sharp exponent 1/2. There is some hope that the above conjecture holds under weaker assumptions on the support of ρ , still assuming ρ bounded above and below on its support.

Before this note, no result had shown that the map $\mu \mapsto T_\mu$ (from $(\mathcal{P}(\mathcal{Y}), W_2)$ to $L^2(\rho)$) could fail to be $\frac{1}{2}$ -Hölder, even for other choices of an absolutely continuous probability measure ρ . In the proof of Theorem 2 (with $p = 2$), we show that for ρ uniform on a well-chosen bounded open set, this map is in fact not better than $\frac{1}{3}$ -Hölder at some (explicit) μ . And Theorem 1 shows that if ρ is allowed to have unbounded density, then it can happen that for any $\alpha > 0$ the map is not α -Hölder.

1.3. Previous works

The literature has been focused so far on quantitative *stability* inequalities, of the form (2). They have been established under various assumptions on ρ . After works by Gigli [8], Berman [3], Mérigot, Delalande & Chazal [11], and Delalande & Mérigot [7], the following result has been achieved in [10].

Theorem 4 ([10, Theorem 1.7]). *Let $\mathcal{X} \subset \mathbb{R}^d$ be a John domain with rectifiable boundary, and let ρ be a probability density on \mathcal{X} , bounded from above and below by positive constants. Then, for any compact set \mathcal{Y} , there exists $C_{\rho, \mathcal{Y}} > 0$ such that for any probability measures μ, ν supported in \mathcal{Y} ,*

$$\|T_\mu - T_\nu\|_{L^2(\rho)} \leq C_{\rho, \mathcal{Y}} W_1(\mu, \nu)^{1/6}. \tag{4}$$

Recall that any bounded and connected Lipschitz domain is a John domain, and therefore Theorem 4 applies for instance in this case. The paper [10] also establishes similar stability inequalities for log-concave ρ , and for ρ blowing-up (or decaying) at some controlled rate at the boundary of a smooth compact set. Quantitative stability inequalities have also been proved for optimal transport maps with respect to p -costs in \mathbb{R}^d [12], and the squared distance cost on Riemannian manifolds [9].

Quantitative stability results were primarily motivated by numerical analysis questions: if μ is known only through an approximation $\nu = \widehat{\mu}$ (for instance through samples), is it true that the optimal transport map $T_{\widehat{\mu}}$, which one may compute for instance through semi-discrete optimal transport, is not far from T_{μ} ?

Quantitative stability inequalities, when they hold, also serve as a justification for the linearized optimal transport framework introduced in [14] (and used later in several applications): the mapping $\mu \mapsto T_{\mu}$ provides an embedding of $(\mathcal{P}(\mathcal{Y}), W_2)$ into the Hilbert space $L^2(\rho, \mathbb{R}^d)$, and this embedding allows one to apply the standard “Hilbertian” statistical toolbox to measure-valued data. This embedding is distance-increasing, meaning that $\|T_{\mu} - T_{\nu}\|_{L^2(\rho)} \geq W_2(\mu, \nu)$, and stability estimates such as (4) (combined with the fact that $W_1 \leq W_2$) show that it is bi-Hölder continuous when ρ satisfies the assumptions of Theorem 4. In other words, the distance $d(\mu, \nu) = \|T_{\mu} - T_{\nu}\|_{L^2(\rho)}$ preserves in a rough way the geometry associated to the Wasserstein distance. However, recall that some results show the impossibility of embedding (in a very coarse sense — in particular in a bi-Hölder way) Wasserstein spaces over \mathbb{R}^d ($d \geq 3$) into Banach spaces of non-trivial type such as Hilbert spaces (see e.g. [1]); this is why the literature about quantitative stability inequality has mostly focused on the case where targets are taken over a compact set \mathcal{Y} .

On the side of instability, the only known results hold merely for *Kantorovich potentials*. Recall that given $\rho, \mu \in \mathcal{P}_2(\mathbb{R}^d)$ with ρ absolutely continuous, a Kantorovich potential is a convex function whose gradient is equal to the Brenier map T_{μ} . If the support of ρ is connected, then there exists for each $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ a unique Kantorovich potential ϕ_{μ} satisfying $\int \phi_{\mu} d\rho = 0$. The paper [10] provides examples of uniform source measures ρ on non-John domains \mathcal{X} (but still bounded and connected) for which Kantorovich potentials are highly unstable in the same sense as above: the inequality

$$\forall \mu, \nu \in \mathcal{P}(\mathcal{Y}), \quad \|\phi_{\mu} - \phi_{\nu}\|_{L^2(\rho)} \leq CW_p(\mu, \nu)^{\alpha}$$

fails for any $C, \alpha > 0$ and $p \geq 1$. However it does not provide examples of unstable transport maps.

2. Proof of Theorem 1

We construct ρ an absolutely continuous probability measure over the (closed) unit ball $B_{\mathbb{R}^d}(0, 1)$. Its density is defined as a function of the distance to two points A and A' , and blows up at A and A' while being uniformly bounded elsewhere. In the sequel, absolutely continuous measures with respect to the Lebesgue measure are identified with their density.

Let $A = (1, 0, \dots, 0) \in \mathbb{R}^d$, $A' = (-1, 0, \dots, 0) \in \mathbb{R}^d$, and $\mathcal{E} = \{A, A'\}$, and let

$$f: r \mapsto r^{-d} \min(1, (\log r)^{-2}) \tag{5}$$

for $r > 0$. Denote by $\text{dist}(x, \mathcal{E})$ the Euclidean distance from x to the set \mathcal{E} . We choose $c_0 > 0$ in a way that the density

$$\rho(x) = c_0 f(\text{dist}(x, \mathcal{E}))$$

on $B_{\mathbb{R}^d}(0, 1)$ is a probability density.

To see that this is possible, let us verify that the function $g: x \mapsto f(\text{dist}(x, \mathcal{E}))$ is integrable on $B_{\mathbb{R}^d}(0, 1)$. We only need to check integrability close to \mathcal{E} . Taking polar coordinates around A , $\varepsilon > 0$ small, and denoting by σ_{d-1} the area of the $(d-1)$ -dimensional unit sphere, we have

$$\int_{B_{\mathbb{R}^d}(A, \varepsilon) \cap B_{\mathbb{R}^d}(0, 1)} g(x) \, dx \leq \sigma_{d-1} \int_0^\varepsilon r^{-d} (\log r)^{-2} r^{d-1} \, dr = \sigma_{d-1} \int_0^\varepsilon \frac{1}{r (\log r)^2} \, dr < +\infty.$$

The same computation holds replacing A by A' . This shows the existence of $c_0 > 0$ as above.

We assume $R = 1$ in the sequel, i.e., target measures are supported in $B_{\mathbb{R}^d}(0, 1)$. Theorem 1 follows for arbitrary R by a scaling argument. Let us consider $B_\theta = (\sin(\theta), \cos(\theta), 0, \dots, 0)$ and $B'_\theta = (-\sin(\theta), -\cos(\theta), 0, \dots, 0)$ for $\theta \in \mathbb{R}$. We set

$$\mu_\theta = \frac{1}{2}(\delta_{B_\theta} + \delta_{B'_\theta}).$$

We study the inequality (2) close to $\theta = 0$. Since ρ is invariant under the transformation $x \mapsto -x$, the optimal transport map from ρ to μ_θ is the map

$$T_{\mu_\theta}: x \mapsto \begin{cases} B_\theta & \text{if } \langle x, B_\theta \rangle > 0, \\ B'_\theta & \text{if } \langle x, B_\theta \rangle < 0. \end{cases}$$

In other words T_{μ_θ} sends each point of the source to the closest point in the support of the target.

For small θ , most points that T_{μ_0} sends to B_0 are sent to B_θ under T_{μ_θ} , but some of them, those satisfying $\langle x, B_0 \rangle > 0 > \langle x, B_\theta \rangle$, are sent to B'_θ . We will bound from below the measure of the latter points in order to bound from below $\|T_{\mu_\theta} - T_{\mu_0}\|_{L^2(\rho)}$. Let us show that for any $\theta \geq 0$ small enough (so that $\sin \theta \geq \theta/2$),

$$B(A', \theta/4) \cap \{x \in \mathbb{R}^d \mid x_2 > 0\} \subset \{x \in \mathbb{R}^d \mid \langle x, B_0 \rangle > 0 > \langle x, B_\theta \rangle\}. \quad (6)$$

Let $x = (x_1, \dots, x_d) \in \mathbb{R}^d$ be an element of the left-hand side. Then $\langle x, B_0 \rangle > 0$ since $x_2 > 0$. Moreover,

$$\langle x, B_\theta \rangle = x_1 \sin \theta + x_2 \cos \theta \leq x_1 \sin \theta + x_2 \leq \left(-1 + \frac{\theta}{4}\right) \frac{\theta}{2} + \frac{\theta}{4} < 0$$

which concludes the proof of (6). Let us now observe that there exists $c_1 > 0$ (independent of θ) such that the left-hand side of (6) has measure at least $c_1 \rho(B(A', \theta/4))$. Indeed, for θ small enough the support of ρ contains the spherical sector

$$B(A', \theta/4) \cap \mathcal{C} \quad (7)$$

(see Figure 1) where \mathcal{C} is the cone

$$\mathcal{C} = \left\{ x \in \mathbb{R}^d \mid x_2 > 0, \frac{x_1 + 1}{|x - A'|} > \frac{1}{2} \right\}.$$

Denote by $c_1 > 0$ the angular aperture of this cone relatively to the full solid angle of the $(d-1)$ -dimensional unit sphere, i.e., the area of the intersection of \mathcal{C} with the unit sphere $\partial B(A', 1)$, divided by the area of the full unit sphere $\partial B(A', 1)$. Since ρ has a radial density close to the center A' , the ρ -measure of $B(A', \theta/4) \cap \mathcal{C}$ is at least equal to $c_1 \rho(B(A', \theta/4))$.

We deduce that for θ small enough, since $|B'_\theta - B_0| \geq 1$,

$$\begin{aligned} \|T_{\mu_\theta} - T_{\mu_0}\|_{L^2(\rho)}^2 &\geq |B'_\theta - B_0|^2 \rho(\{x \mid \langle x, B_0 \rangle > 0 > \langle x, B_\theta \rangle\}) \\ &\geq c_1 \rho(B(A', \theta/4)) \\ &\geq c_0 c_1 \sigma_{d-1} \int_0^{\theta/4} \frac{1}{r (\log r)^2} \, dr \\ &= c_0 c_1 \sigma_{d-1} \frac{1}{|\log(\theta/4)|} \end{aligned}$$

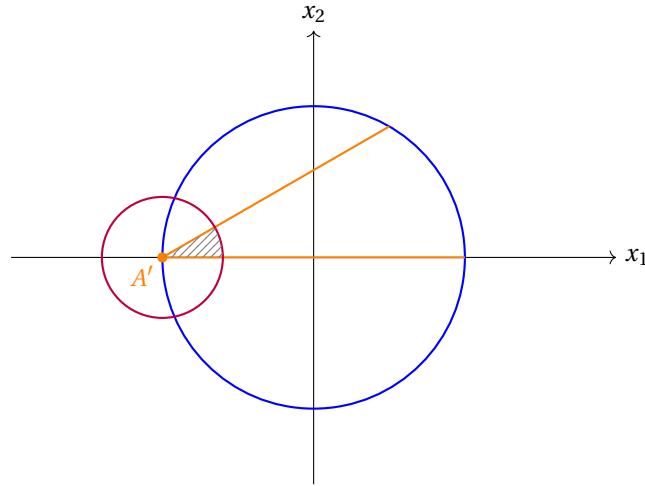


Figure 1. Illustration of (7): in purple the ball $B(A', \theta/4)$, in orange the cone in (7), in grey dashed lines the set (7), and in blue the boundary of the support of ρ .

which decays to 0 as $\theta \rightarrow 0$ slower than any θ^α with $\alpha > 0$. Moreover, for θ small enough, $W_p(\mu_0, \mu_\theta) = \sin(\theta/2) \sim \theta/2$. Put together, these estimates conclude the proof of Theorem 1.

Remark 5. The density ρ used in the proof of Theorem 1 blows up at rate $r^{-d}(\log r)^{-2}$. Let us instead consider $\rho(x) = c_\delta h(\text{dist}(x, \mathcal{E}))$ where

$$h: r \mapsto r^{-d+\delta}$$

for some $\delta > 0$, i.e., a density which blows up slightly slower at A and A' . The assumption $\delta > 0$ is necessary to ensure that ρ is integrable, and c_δ is chosen to make ρ a probability measure. Following the proof strategy of [10], notably the proof of Theorem 1.10 in loc. cit., one can show that for any compact $\mathcal{Y} \subset \mathbb{R}^d$, there exists $C > 0$ such that

$$\|T_\mu - T_\nu\|_{L^2(\rho)} \leq CW_2(\mu, \nu)^{\frac{\delta}{6d}}.$$

Conversely, the same argument as in the proof of Theorem 1 (with h replacing f) shows that if

$$\|T_\mu - T_\nu\|_{L^2(\rho)} \leq CW_p(\mu, \nu)^\alpha \tag{8}$$

holds for some $C, \alpha > 0$ and $p \geq 1$, then necessarily $\alpha \leq \delta$.

It is therefore no surprise that for densities which blow up faster than $r^{-d+\delta}$ for any $\delta > 0$, any quantitative stability inequality of the form (8) breaks down. For these densities, the techniques of [10] do not apply anymore, and we are not aware of any quantitative stability inequality, even with a larger right-hand side than $W_p(\mu, \nu)^\alpha$ (e.g., $(1 + |\log W_p(\mu, \nu)|)^{-1}$).

Actually, weaker quantitative estimates (than (8)) should then take into account quite finely the blowup rate of the density. For instance, switching f defined in (5) to a Bertrand integrand

$$r \mapsto r^{-d} \min\left(1, \frac{1}{|\log r| \cdots |\log|^{\circ k}(r) (|\log|^{\circ k+1} r)^2}\right),$$

where $g^{\circ k}$ denotes the k times composition of a function g with itself, will give arbitrarily slow convergence (in log scale) to 0.

Remark 6. For the choice of ρ used in the proof of Theorem 1, the strong instability of optimal transport maps naturally raises the following question: how can one statistically estimate optimal transport maps with this source measure?

3. Proof of Theorem 2

3.1. Idea of the construction

To prove Theorem 2, we build upon the idea that stability of optimal transport maps can be lost just before uniqueness is lost. To explain our construction of ρ , let us start with a simple and well-known example where optimal transport plans are not unique: consider the vertices A, B, A', B' of a square, for instance $A = (1, 0), B = (0, 1), A' = (-1, 0)$ and $B' = (0, -1)$, and consider the optimal transport problem from $\frac{1}{2}(\delta_A + \delta_{A'})$ to $\frac{1}{2}(\delta_B + \delta_{B'})$.

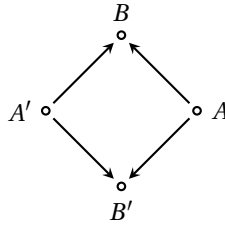


Figure 2. Any transport plan between $\frac{1}{2}(\delta_A + \delta_{A'})$ and $\frac{1}{2}(\delta_B + \delta_{B'})$ is optimal.

It is not difficult to see that all transport plans have the same cost, in particular there exist infinitely many optimal transport plans. Moreover, if we keep A and A' unchanged, but we move a little bit B and B' , then we can recover unique optimal transport plans/maps. For instance, if we slightly move B horizontally to the right, and B' symmetrically slightly to the left, then we recover a unique optimal transport map, where the mass at A is sent to B , and the mass at A' is sent to B' . Symmetrically, if we move B to the left and B' to the right, then we also recover a unique optimal transport map, but the mass at A is sent to B' and the mass at A' is sent to B . The configuration displayed in Figure 2 therefore generates a strong instability of optimal transport maps (and plans). However, this example is quite specific in the sense that it deals with discrete measures, and Brenier’s theorem does not apply.

Nonetheless, we can take inspiration from this example to construct absolutely continuous (and compactly supported) source measures ρ for which optimal transport maps are highly unstable. First, we slightly smooth the source measure by replacing it by $\rho_r = \frac{1}{2\pi r^2}(\delta_{B(A,r)} + \delta_{B(A',r)})$ for some small $r > 0$. We denote by $T_\mu^{(r)}$ the optimal transport map from ρ_r to a measure μ . We know by the results of [10] that stability holds for any $r > 0$: for any compact set $\mathcal{Y} \subset \mathbb{R}^d$,

$$\forall \mu, \nu \in \mathcal{P}(\mathcal{Y}), \quad \|T_\mu^{(r)} - T_\nu^{(r)}\|_{L^2(\rho)} \leq C_r W_1(\mu, \nu)^{1/6}.$$

It can be checked that the constant C_r tends to $+\infty$ as $r \rightarrow 0$, which is a sign of instability (in the limit $r \rightarrow 0$). To prove Theorem 1 we repeat the above construction “at all scales”. In other words, we take ρ uniform over an infinite union of well-chosen pairs of balls $B(A_i, r_i), B(A'_i, r_i)$ of various sizes. Actually, we do not take balls but rectangular parallelepipeds because it is helpful at some point to have more parameters than just the radius.

3.2. The construction

We first construct the support of ρ , denoted by \mathcal{X} in what follows (and depicted in Figure 3). In what follows, we consider the rectangular parallelepiped

$$\mathcal{Q}(\ell, r) = \left\{ (x_1, \dots, x_d) \in \mathbb{R}^d \mid 0 < x_1 < \ell, |x_2| < \frac{r}{2}, \text{ and } |x_i| < \frac{1}{2} \text{ for } 3 \leq i \leq d \right\}$$

and consider for $A \in \mathbb{R}^d$ the translates

$$\mathcal{F}^+(A, \ell, r) = A + \mathcal{Q}(\ell, r), \quad \mathcal{F}^-(A, \ell, r) = A - \mathcal{Q}(\ell, r)$$

(in the sense of the Minkowski sum and difference).

In the sequel, $\mathbb{N} = \{1, 2, \dots\}$ denotes the set of positive integers. Let $(\ell_i), (r_i), (w_i) \in (\mathbb{R}_+)^{\mathbb{N}}$ and $(u_i) \in \mathbb{R}^{\mathbb{N}}$. We will make several assumptions on these sequences in Sections 3.3 and 3.4. Let

$$\mathcal{S}_i = \mathcal{F}^+(A_i^+, \ell_i, r_i) \cup \mathcal{F}^-(A_i^-, \ell_i, r_i) \tag{9}$$

where $A_i^+ = (u_i + w_i, 0, \dots, 0)$ and $A_i^- = (u_i - w_i, 0, \dots, 0)$, and let

$$\mathcal{X} = \bigcup_{i=1}^{+\infty} \mathcal{S}_i.$$

In the sequel, the sequences will always be chosen in a way that when traveling along the x_1 -axis in the increasing x_1 direction, the parallelepipeds defined above do not meet, and are in the order $\mathcal{S}_1, \mathcal{S}_2, \dots$ as in Figure 3.

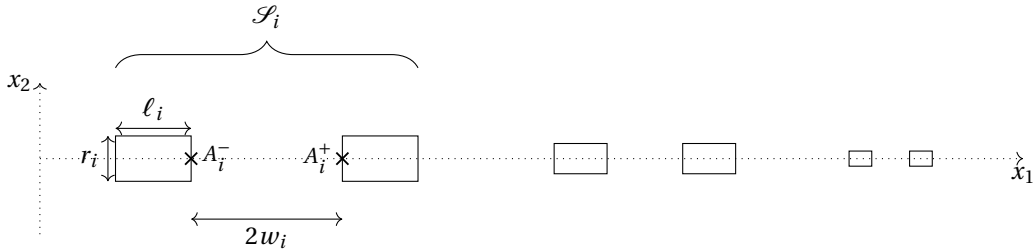


Figure 3. Part of the support of ρ , projected on the (x_1, x_2) -plane.

We consider ρ an absolutely continuous probability measure whose support is \mathcal{X} and which, for any i , is uniform on \mathcal{S}_i . Finally, we let for any i

$$\sigma_i = \rho(\mathcal{F}^+(A_i^+, \ell_i, r_i)) = \rho(\mathcal{F}^-(A_i^-, \ell_i, r_i)).$$

With a slight abuse of notation, the density of ρ with respect to the Lebesgue measure is also denoted by ρ in what follows.

Let $B_i^+ = (u_i, w_i, 0, \dots, 0)$ and $B_i^- = (u_i, -w_i, 0, \dots, 0)$. Let

$$\mu = \sum_{i=1}^{+\infty} \sigma_i (\delta_{B_i^+} + \delta_{B_i^-}). \tag{10}$$

It is immediate to check that this is a probability measure.

For $i \in \mathbb{N}$, let $C_i^+ = (u_i + r_i, w_i, 0, \dots, 0)$, $C_i^- = (u_i - r_i, -w_i, 0, \dots, 0)$, and

$$\nu_i = \mu + \sigma_i (\delta_{C_i^+} + \delta_{C_i^-} - \delta_{B_i^+} - \delta_{B_i^-}).$$

It is of course a probability measure. The only difference between ν_i and μ is that in the sum (10), the i -th term has been replaced by $\sigma_i (\delta_{C_i^+} + \delta_{C_i^-})$, while all other terms are left unchanged. With the choices made in the next section, C_i^+ and C_i^- are seen as perturbations of B_i^+ and B_i^- .

In the sequel we call “the i -th cell” the set

$$\pi_{12}(\mathcal{S}_i \cup \{B_i^+, B_i^-, C_i^+, C_i^-\})$$

where π_{12} denotes the projection onto the first two coordinates.

3.3. Preliminary computations

In what follows we take the convention $u_0 = -\infty$. We make the following assumptions on the sequences (roughly illustrated in Figure 4): for any $i \in \mathbb{N}$,

$$\min(u_i - u_{i-1}, u_{i+1} - u_i) \geq 100 \max(\ell_i, r_i, w_i), \tag{11}$$

$$w_i \geq 100r_i. \tag{12}$$

The first inequality means that the cells are well-separated: the distances between points inside a given cell are much smaller than the distance to the closest other cells. The second inequality means B_i^+ is much closer to C_i^+ than to B_i^- (and B_i^- is much closer to C_i^- than to B_i^+). We will choose specific sequences (ℓ_i) , (r_i) , (w_i) and (u_i) in Section 3.4. In any case of application, these sequences are bounded (which implies that the probability measure ρ which we construct has compact support).

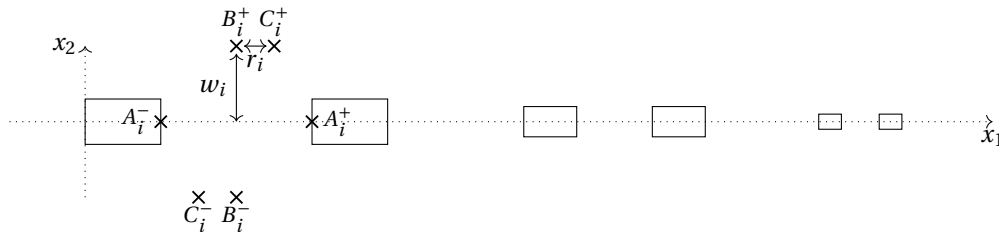


Figure 4. The support of the measures ρ , μ and ν_i .

Under assumptions (11)–(12), we show that for any choice of C , $\alpha > 0$ and $p \geq 1$, the inequality

$$\|T_\mu - T_{\nu_i}\|_{L^2(\rho)} \leq CW_p(\mu, \nu_i)^\alpha \tag{13}$$

cannot hold for all ν_i simultaneously. For this, we first compute T_μ and T_{ν_i} for any i .

For $x = (x_1, \dots, x_d) \in \mathbb{R}^d$, let us check that

$$x \mapsto \begin{cases} B_i^+ & \text{if } x \in \mathcal{S}_i \text{ and } x_2 \geq 0, \\ B_i^- & \text{if } x \in \mathcal{S}_i \text{ and } x_2 < 0, \end{cases} \tag{14}$$

coincides ρ -a.e. with T_μ . It is immediate to verify that this application transports ρ to μ . Let us check that each point of the support of ρ is sent by this application to the closest point in the support of μ . Let $x \in \mathcal{S}_i$. We may assume that $x_2 \geq 0$ (the case $x_2 < 0$ is symmetric) and that $x_3 = \dots = x_d = 0$ because all points in the support of μ have these coordinates equal to 0. Then x is closer to B_i^+ than to B_i^- , and x is at distance at most $(w_i^2 + (\ell_i + w_i)^2)^{1/2}$ from B_i^+ . Let y be another point in the support of μ , different from B_i^- and B_i^+ . The triangle inequality and then (11) yield

$$|x - y| \geq \min(u_i - u_{i-1}, u_{i+1} - u_i) - w_i - \ell_i > (w_i^2 + (\ell_i + w_i)^2)^{1/2} \geq |x - B_i^+| \tag{15}$$

i.e., x is closer to B_i^+ than to y . Therefore, the transport is made at smallest possible cost, and (14) is the optimal transport map.

Regarding T_{ν_i} , let us show that it coincides ρ -a.e. with the application

$$x \mapsto \begin{cases} T_\mu(x) & \text{if } x \notin \mathcal{S}_i, \\ C_i^+ & \text{if } x \in \mathcal{F}^+(A_i^+, \ell_i, r_i), \\ C_i^- & \text{if } x \in \mathcal{F}^-(A_i^-, \ell_i, r_i). \end{cases} \tag{16}$$

It is clear that this application defines a transport map from ρ to ν_i . We only need to show the following claim, which implies that (16) is the optimal transport map from ρ to ν_i .

Claim. *The application (16) sends each point x in the support of ρ to its closest point in the support of ν_i .*

This claim is straightforward to check for $x \in \mathcal{S}_j$ ($j \neq i$) with a similar argument as in (15). It is also immediate that the closest point to $x \in \mathcal{T}^+(A_i^+, \ell_i, r_i)$ is either C_i^+ or C_i^- . Let us show that any $x \in \mathcal{T}^+(A_i^+, \ell_i, r_i)$ is closer to C_i^+ than to C_i^- . We let $x = A_i^+ + (x_1, \dots, x_d)$ and observe that

$$\begin{aligned} |x - C_i^-|^2 - |x - C_i^+|^2 &= (x_1 + w_i + r_i)^2 + (w_i + x_2)^2 - (x_1 + w_i - r_i)^2 - (w_i - x_2)^2 \\ &= 4r_i(w_i + x_1) + 4w_i x_2 \\ &\geq 4r_i w_i - 2r_i w_i \\ &> 0. \end{aligned}$$

Similarly if $x \in \mathcal{T}^-(A_i^-, \ell_i, r_i)$, then x is closer to C_i^- than to C_i^+ . Hence the transport is made at smallest possible cost, and (16) is the optimal transport map.

From the explicit expressions of T_μ and T_{ν_i} obtained above, we deduce

$$\|T_\mu - T_{\nu_i}\|_{L^2(\rho)}^2 = \frac{1}{2}\rho(\mathcal{S}_i)|B_i^+ - C_i^+|^2 + \frac{1}{2}\rho(\mathcal{S}_i)|B_i^- - C_i^-|^2 = \sigma_i(2r_i^2 + 4w_i^2) \geq 4\sigma_i w_i^2. \tag{17}$$

Besides, due to (12), we have $r_i \leq w_i/100$, hence the optimal coupling between μ and ν_i is a coupling where the mass at B_j^+ (resp. B_j^-) does not move if $j \neq i$, and is sent to C_j^+ (resp. C_j^-) if $j = i$. Indeed, in this way, each piece of mass in the support of μ is sent to the closest point in the support of ν_i . Therefore, for any $p \geq 1$,

$$W_p(\mu, \nu_i) = r_i \sigma_i^{1/p}. \tag{18}$$

Combining (17) and (18) we deduce that for any $\alpha > 0$,

$$\frac{\|T_\mu - T_{\nu_i}\|_{L^2(\rho)}^2}{W_p(\mu, \nu_i)^{2\alpha}} \geq 4w_i^2 r_i^{-2\alpha} \sigma_i^{1 - \frac{2\alpha}{p}}. \tag{19}$$

3.4. End of the proof of Theorem 2(1)

Let $p \geq 1$ and $\alpha > \frac{p}{2(p+1)}$. We choose the sequences $(\ell_i)_{i \in \mathbb{N}}$, $(r_i)_{i \in \mathbb{N}}$, $(w_i)_{i \in \mathbb{N}}$ and $(u_i)_{i \in \mathbb{N}}$ as follows. We let $r_i = c_0 2^{-i}$ (again, any sequence with superpolynomial decay would work), $\ell_i = w_i = c_0 c_1 i^{-2}$ and $u_1 \in \mathbb{R}$, $u_{i+1} - u_i = c_0 c_2 i^{-2}$ where $c_1, c_2 > 0$ are chosen in a way that (11) and (12) hold. Choosing appropriately $c_0 > 0$, we fix

$$\sum_{i=1}^{+\infty} \ell_i r_i = \frac{1}{2}$$

and let $\sigma_i = \ell_i r_i$, which makes ρ uniform over its support. We get

$$\frac{\|T_\mu - T_{\nu_i}\|_{L^2(\rho)}^2}{W_p(\mu, \nu_i)^{2\alpha}} \geq 4w_i^2 r_i^{-2\alpha} \sigma_i^{1 - \frac{2\alpha}{p}} = C w_i^{3 - \frac{2\alpha}{p}} r_i^{1 - \frac{2\alpha}{p} - 2\alpha}. \tag{20}$$

Since $\alpha > \frac{p}{2(p+1)}$, this quantity tends to $+\infty$ as $i \rightarrow +\infty$.

Remark 7. We did not obtain any improvement of our results by introducing other parameters in the geometric picture, for instance by making the distance between A_i^- and A_i^+ different from the distance between B_i^- and B_i^+ .

Remark 8. Theorems 1 and 2(1) hold, without any single word changed, for any cost of the form $h(|x - y|)$ with $h: \mathbb{R}_+ \rightarrow \mathbb{R}$ increasing and strictly convex (see for instance [13, Theorem 1.17] for the existence of optimal transport maps in this context). The families of optimal transport maps used to prove Theorems 1 and 2(1) are indeed optimal with respect to any cost of this form, since they are obtained by sending each point in the support of the source measure into the closest point (in Euclidean distance) in the support of the target measure.

3.5. Proof of Theorem 2(2)

We use a truncation argument similar to the one in [10, Section 2.2]. For any $i \in \mathbb{N}$, we consider ρ_i^\pm the restriction of ρ to $\mathcal{T}^\pm(A_i^\pm, \ell_i, r_i)$. We denote by $\tilde{\rho}_i^\pm = \rho_i^\pm / \rho_i^\pm(\mathcal{T}^\pm(A_i^\pm, \ell_i, r_i))$ the associated probability density. According to [10, Theorem 2.1], for any $\psi_0, \psi_1 \in \mathcal{C}^0(\mathcal{Y})$,

$$\text{Var}_{\tilde{\rho}_i^\pm}(\psi_1^* - \psi_0^*) \leq C_{\mathcal{Y}} \langle \psi_1 - \psi_0 \mid \nabla \psi_{0\#}^* \tilde{\rho}_i^\pm - \nabla \psi_{1\#}^* \tilde{\rho}_i^\pm \rangle.$$

Importantly, $C_{\mathcal{Y}}$ does not depend on i . Let ϕ_μ (resp. ϕ_ν) denote a Brenier potential for the quadratic optimal transport problem from ρ to μ (resp. ρ to ν). Applying the above inequality to $\psi_0 = \phi_\mu^*$ and $\psi_1 = \phi_\nu^*$, and multiplying by $\rho_i^\pm(\mathcal{T}^\pm(A_i^\pm, \ell_i, r_i))$ on both sides, we obtain

$$\rho_i^\pm(\mathcal{T}^\pm(A_i^\pm, \ell_i, r_i)) \text{Var}_{\tilde{\rho}_i^\pm}(\phi_\mu - \phi_\nu) \leq C_{\mathcal{Y}} \langle \phi_\nu^* - \phi_\mu^* \mid \nabla \phi_{\mu\#} \rho_i^\pm - \nabla \phi_{\nu\#} \rho_i^\pm \rangle. \quad (21)$$

Let $c_i^\pm = \int_{\mathcal{X}} (\phi_\mu - \phi_\nu) d\tilde{\rho}_i^\pm$. For $k \in \mathbb{N}$, let also ρ_k be the probability density whose density is proportional to $\sum_{i=1}^k \rho_i^+ + \rho_i^-$, and write $b_k \rho_k = \sum_{i=1}^k \rho_i^+ + \rho_i^-$ for some $b_k \in [0, 1]$. Finally, let $\mu_k = \nabla \phi_{\mu\#} \rho_k$ and $\nu_k = \nabla \phi_{\nu\#} \rho_k$. In the sequel we assume that k is large enough so that $b_k \geq 1/2$. Then summing the inequality (21) over i and over $\{+, -\}$ we deduce

$$\sum_{i=1}^k \sum_{\{+, -\}} \|\phi_\mu - \phi_\nu - c_i^\pm\|_{L^2(\rho_i^\pm)}^2 \leq C_{\mathcal{Y}} \langle \phi_\nu^* - \phi_\mu^* \mid \mu_k - \nu_k \rangle \leq C_{\rho, \mathcal{Y}} W_1(\mu_k, \nu_k) \quad (22)$$

where in the last inequality we used Kantorovich–Rubinstein duality and the fact that ϕ_μ^* and ϕ_ν^* are Lipschitz, since \mathcal{X} (the support of ρ) is bounded. To get stability of maps we apply [7, Proposition 4.1] in $\mathcal{T}^\pm(A_i^\pm, \ell_i, r_i)$, which gives

$$\|T_\mu - T_\nu\|_{L^2(\rho_i^\pm)}^2 \leq C_{\rho, \mathcal{Y}} \|\phi_\mu - \phi_\nu - c_i^\pm\|_{L^2(\rho_i^\pm)}^{2/3}$$

where again $C_{\rho, \mathcal{Y}}$ does not depend on i . Summing over i and $\{+, -\}$ we obtain

$$\begin{aligned} \|T_\mu - T_\nu\|_{L^2(\rho_k)}^2 &\leq C_{\rho, \mathcal{Y}} \sum_{i=1}^k \sum_{\{+, -\}} \|\phi_\mu - \phi_\nu - c_i^\pm\|_{L^2(\rho_i)}^{2/3} \\ &\leq C_{\rho, \mathcal{Y}} k^{2/3} W_1(\mu_k, \nu_k)^{1/3} \end{aligned} \quad (23)$$

where we used that $\sum_{i=1}^k a_i \leq k^{2/3} (\sum_{i=1}^k a_i^3)^{1/3}$ for any sequence (a_i) of non-negative numbers, together with (22).

To find an upper bound on $W_1(\mu_k, \nu_k)$ we proceed as follows. We write $\mu = b_k \mu_k + \mu'_k$ where b_k is introduced below (21). Thanks to the Kantorovich–Rubinstein duality formula we know there exists a 1-Lipschitz function g such that $W_1(\mu_k, \mu) = \int_{\mathcal{Y}} g d(\mu_k - \mu)$, and we may assume that $g(0) = 0$. Hence $|g(x)| \leq R_{\mathcal{Y}} := \sup_{y \in \mathcal{Y}} |y|$ for any $x \in \mathcal{Y}$. Since $\mu'_k(\mathcal{Y}) = 1 - b_k$ we deduce

$$W_1(\mu_k, \mu) = (1 - b_k) \int_{\mathcal{Y}} g d\mu_k - \int_{\mathcal{Y}} g d\mu'_k \leq 2R_{\mathcal{Y}}(1 - b_k).$$

Similarly, $W_1(\nu_k, \nu) \leq 2R_{\mathcal{Y}}(1 - b_k)$. Thus

$$W_1(\mu_k, \nu_k) \leq W_1(\mu, \nu) + W_1(\mu, \mu_k) + W_1(\nu, \nu_k) \leq W_1(\mu, \nu) + 4R_{\mathcal{Y}}(1 - b_k).$$

By construction, there holds $b_k \leq C_\rho 2^{-k}$ for some $C_\rho > 0$ independent of k . Plugging into (23) and using that T_μ and T_ν are essentially bounded by $R_{\mathcal{Y}}$ on \mathcal{X} , we get

$$\|T_\mu - T_\nu\|_{L^2(\rho)}^2 \leq C_{\rho, \mathcal{Y}} k^{2/3} (W_1(\mu, \nu) + 2^{-k})^{1/3} + C_{\rho, \mathcal{Y}} 2^{-k}.$$

To establish (4) we may assume that $W_1(\mu, \nu) \leq 1/10$, since otherwise the boundedness of T_μ, T_ν allows to conclude. Choosing $k = \lfloor -\log_2 W_1(\mu, \nu) \rfloor$, we get

$$\|T_\mu - T_\nu\|_{L^2(\rho)} \leq C_{\rho, \mathcal{Y}} W_1(\mu, \nu)^{1/6} |\log W_1(\mu, \nu)|^{1/3}$$

which implies (4).

Remark 9. It is unlikely that strong instability as in Theorem 1 can be obtained for some ρ bounded above and below on a well-chosen bounded open set \mathcal{X} : due to [10, Remark 4.2], \mathcal{X} would need to have infinitely many connected components, and the above computations show that even in this case (3) holds for some $C, \alpha > 0$ and $p = 1$, as soon as the size of the connected components decays fast enough.

Acknowledgments

I am very thankful to an anonymous referee for a suggestion that led to part (2) of Theorem 2.

Declaration of interests

The author does not work for, advise, own shares in, or receive funds from any organization that could benefit from this article, and has declared no affiliations other than their research organizations.

References

- [1] A. Andoni, A. Naor and O. Neiman, “Snowflake universality of Wasserstein spaces”, *Ann. Sci. Éc. Norm. Supér. (4)* **51** (2018), no. 3, pp. 657–700.
- [2] S. Balakrishnan, T. Manole and L. Wasserman, “Statistical inference for optimal transport maps: recent advances and perspectives”, 2025. Online at <https://arxiv.org/abs/2506.19025>.
- [3] R. J. Berman, “Convergence rates for discretized Monge–Ampère equations and quantitative stability of optimal transport”, *Found. Comput. Math.* **21** (2021), no. 4, pp. 1099–1140.
- [4] Y. Brenier, “Décomposition polaire et réarrangement monotone des champs de vecteurs”, *Comptes Rendus. Mathématique* **305** (1987), no. 19, pp. 805–808.
- [5] Y. Brenier, “Polar factorization and monotone rearrangement of vector-valued functions”, *Commun. Pure Appl. Math.* **44** (1991), no. 4, pp. 375–417.
- [6] S. Chewi, J. Niles-Weed and P. Rigollet, “Statistical optimal transport”, 2024. Online at <https://arxiv.org/abs/2407.18163>.
- [7] A. Delalande and Q. Mérigot, “Quantitative stability of optimal transport maps under variations of the target measure”, *Duke Math. J.* **172** (2023), no. 17, pp. 3321–3357.
- [8] N. Gigli, “On Hölder continuity-in-time of the optimal transport map towards measures along a curve”, *Proc. Edinb. Math. Soc., II. Ser.* **54** (2011), no. 2, pp. 401–409.
- [9] J. Kitagawa, C. Letrouit and Q. Mérigot, “Stability of optimal transport maps on Riemannian manifolds”, 2025. Online at <https://arxiv.org/abs/2504.05412>.
- [10] C. Letrouit and Q. Mérigot, “Gluing methods for quantitative stability of optimal transport maps”, 2024. Online at <https://arxiv.org/abs/2411.04908>.
- [11] Q. Mérigot, A. Delalande and F. Chazal, “Quantitative stability of optimal transport maps and linearization of the 2-Wasserstein space”, in *Proceedings of the Twenty Third International Conference on Artificial Intelligence and Statistics* (S. Chiappa and R. Calandra, eds.), Proceedings of Machine Learning Research, vol. 108, PMLR, 2020, pp. 3186–3196.
- [12] O. Mischler and D. Trevisan, “Quantitative stability in optimal transport for general power costs”, 2024. Online at <https://arxiv.org/abs/2407.19337>.
- [13] F. Santambrogio, *Optimal transport for applied mathematicians*, Progress in Nonlinear Differential Equations and their Applications, vol. 87, Birkhäuser/Springer, 2015.
- [14] W. Wang, D. Slepčev, S. Basu, J. A. Ozolek and G. K. Rohde, “A linear optimal transportation framework for quantifying and visualizing variations in sets of images”, *Int. J. Comput. Vis.* **101** (2013), no. 2, pp. 254–269.