PARTIAL REGULARITY FOR OPTIMAL TRANSPORT WITH P-COST AWAY FROM FIXED POINTS

MICHAEL GOLDMAN AND LUKAS KOCH

ABSTRACT. We consider maps T solving the optimal transport problem with a cost $c(x-y)$ modeled on the p-cost. For Hölder continuous marginals, we prove a $C^{1,\alpha}$-partial regularity result for T in the set $\{|T(x) - x| > 0\}$.

1. INTRODUCTION

In this paper we are concerned with the optimal transportation problem

$$(1.1) \qquad \min_{\pi \in \Pi(\rho_0, \rho_1)} \int c(x - y) \mathrm{d}\pi,$$

where $\Pi(\rho_0, \rho_1)$ denotes the set of plans in $\mathbb{R}^d \times \mathbb{R}^d$ with marginals ρ_0 and ρ_1. Under mild assumptions, see e.g. [16, Theorem 2.12], it is known that the minimiser is unique and of Monge-form, that is $\pi = (x, T(x))_{\#\rho_0}$ for some map $T \colon \mathbb{R}^d \to \mathbb{R}^d$. We are concerned with obtaining a partial regularity result for T in the setting where c is modeled on the cost $|x - y|^p$ for some $p > 1$.

In order to state our main theorem, let us make our assumptions on the cost function precise.

Assumption 1. *Let $p > 1$ and $\alpha \in (0, 1]$. We consider a convex cost function $c \in C^{2,\alpha}(\mathbb{R}^d \setminus \{0\}) \cap C^1(\mathbb{R}^d)$ satisfying $c(0) = 0$, $\nabla c(0) = 0$ and for some $\Lambda > 0$,*

$$(1.2) \qquad \Lambda^{-1}(|x| + |y|)^{p-2}|x - y|^2 \leq \langle \nabla c(x) - \nabla c(y), x - y \rangle, \qquad \forall x, y \in \mathbb{R}^d$$

$$(1.3) \qquad |\nabla c(x) - \nabla c(y)| \leq \Lambda(|x| + |y|)^{p-2}|x - y|, \qquad \forall x, y \in \mathbb{R}^d.$$

Let

$$X = \{x \colon |T(x) - x| > 0\} \qquad \text{and} \qquad Y = \{y \colon |y - T^{-1}(y)| > 0\}.$$

Our main result then reads as follows:

Theorem 1. *Let $p > 1$. Suppose c satisfies Assumption (1) and let $\rho_0, \rho_1 \in C^\alpha(\mathbb{R}^d)$. Suppose $\pi = (x, T(x))_{\#\rho_0}$ solves (1.1). Then, there exist open sets X' and Y' with $|X \setminus X'| = |Y \setminus Y'| = 0$ and that T is a $C^{1,\alpha}$-diffeomorphism between the sets X' and Y'.*

Partial regularity theorems for (1.1) were obtained by [6] in the case $|x - y|^2$ and for general non-singular cost c in [4] using CAFFARELLI'S viscosity approach to solutions of the Monge-Ampère equation, which is the Euler-Lagrange equation for (1.1). In [9, 7] the result of [6] was obtained using a variational approach based on the geometric linearisation of the Monge-Ampère equation, see also [8]. This was later generalised to reprove the results of

[4] in [15]. Regarding singular cost functions of $p-$type, to the best of our knowledge, the only known result is the global regularity for densities defined on convex domains which are either sufficiently far apart or under the condition that $p \geq 2$ with $|p - 2| \ll 1$, see [3]. In this note we thus obtain partial regularity when the domains are neither convex nor disjoint.

The main (and very natural) idea behind our proof is that away from fixed points, that is for $x \neq 0$, c locally looks like a quadratic cost function. In order to localize and apply the existing partial regularity theory for non-singular cost-functions from [4, 15] we need a localized L^∞ bound on the displacement. When $p \geq 2$, since the Kantorovich potentials are semi-convex, this follows from Alexandrov theorem (see [4, 15]) while the situation is more delicate for $p < 2$. Our second main result is the following L^∞-bound.

Theorem 2. *Let $p > 1$ and π be a minimiser for the c-optimal transport problem between ρ_0 and ρ_1. Given $A \in \mathbb{R}^{d \times d}$, $b \in \mathbb{R}^d$ and $m > 0$, denote*

$$E_m = \int_{B_2} |T(x) - x - b - Ax|^m \mathrm{d}\rho_0, \quad D = \inf_{B_r \Subset B_2} \frac{\rho_0(B_r)}{r^d}.$$

Assuming $D > 0$, it holds that for $x \in B_{1/2}$,

$$|T(x) - x - b - Ax| \lesssim \begin{cases} E_m^{\frac{1}{m}} & \text{if } E_m \geq 1 \\ E_m^{\frac{1}{m+d}} & \text{if } E_m \leq 1. \end{cases}$$

The implicit constant depends on D, p, m, d and $|A|$, but not on b.

In order to prove Theorem 1, we will apply Theorem 2 with $m = p$, in which case E_p is a standard excess quantity.

Theorem 2 is of interest in its own right and might actually be seen as the main contribution of the note. Global versions of this bound with $m = p$, $b = 0$ and $A = 0$ first appeared in [2]. As far as local bounds are concerned, in the case $p = 2$, where by affine invariance of the cost it is possible to reduce to the case $b = 0$ and $A = 0$, such bounds were obtained in [9, 15, 8]. These bounds played a major role in the variational approach to regularity theory for optimal transport maps. Still when $b = 0$ and $A = 0$ but for $p \neq 2$, Theorem 2 recovers earlier results of [14, 10]. In the case $p \geq 2$ (and for homogeneous costs), Theorem 2 has been proven in [12] (see also the first Arxiv version of [11]) with a much more complicated proof.

As alluded to, when $p \neq 2$, we lose the affine invariance of the cost and Theorem 2 cannot be reduced to the case $b = 0$ and $A = 0$ anymore. Moreover, in this case, the approach from [9, 15] with the geometric interpretation from [14] seems difficult to implement. Our argument is much closer in spirit to the original proof in [2]. As in [2], the idea is to use the c-monotonicity optimality condition for many points which lie on the ray $T(x) - x - b$.

We anticipate that Theorem 2 might have applications for the derivation of a full partial regularity result (also including fixed points). This motivated us to include the case $A \neq 0$

in our statement. Let us however point out that an $\varepsilon-$regularity theorem around fixed points seems to be a very challenging problem for multiple reasons. First, even though an 'harmonic-approximation' result has been obtained in this context in [14], the approach of [9] works at the level of $C^{1,\alpha}$ regularity for the map T which requires $C^{2,\alpha}$ regularity for the solutions of the geometrically linearized equation (here the $p'-$Laplace equation). However such regularity is known to fail for $p > 2$ and is only known qualitatively for $d = 2$ and $p < 2$, see [13]. Second, even if we ignore this issue, due to the lack of affine invariance of the cost, after one-step of the Campanato iteration we lose the structure of having a cost of the form $c(x - y)$ which is essential for the 'harmonic-approximation' results of [9, 8, 14]. Last, in the case $p < 2$, where one could hope to get the needed $C^{2,\alpha}$ regularity for the $p'-$Laplace equation, the Kantorovich potentials are not semi-convex anymore and we cannot appeal to Alexandrov theorem to prove that around most fixed point the excess is small. See however [5] where a related partial regularity results has been obtained for the geometrically linearized problem. We thus leave this problem for future research.

The outline of our paper is as follows. In Section 2 we collect some preliminary comments from convex analysis and regarding certain quantities related to p-cost. In Section 3 we prove Theorem 2, while we prove Theorem 1 in Section 4

Notation

The symbols $\simeq, \gtrsim, \lesssim$ indicate estimates that hold up to a global constant C, which only depends on the dimension d, the exponent p and the Hölder exponent α (if applicable). For instance, $f \lesssim g$ means that there exists such a constant with $f \leq Cg$, $f \simeq g$ means $f \lesssim g$ and $g \lesssim f$. An assumption of the form $f \ll 1$ means that there exists $\varepsilon > 0$, only depending on p, d and α, such that if $f \leq \varepsilon$, then the conclusion holds.

2. Preliminaries

We collect some well-known facts from convex analysis and regarding certain quantities naturally related to the type of cost functions we consider.

Let $c \in C^1(\mathbb{R}^d)$ be a convex function. Then for any $x, y \in \mathbb{R}^d$,

$$(2.1) \qquad c(x) - c(y) \leq \langle \nabla c(x), x - y \rangle.$$

Furthermore if c is strictly convex, C^1 and superlinear, then ∇c is a homeomorphism of $\mathbb{R}^d$.

(1.2) is equivalent to the assumption that for some $\Lambda > 0$ and all $x, y \in \mathbb{R}^d$,

$$(2.2) \qquad c(x) \geq c(y) + \langle \nabla c(y), x - y \rangle + \Lambda^{-1}(|x| + |y|)^{p-2}|x - y|^2.$$

A further equivalent formulation is that for some $\Lambda > 0$ and all $x, y \in \mathbb{R}^d$, $\lambda \in [0, 1]$,

$$(2.3) \qquad c(\lambda x + (1 - \lambda)y) \leq \lambda c(x) + (1 - \lambda)c(y) - \Lambda^{-1}\lambda(1 - \lambda)(|x| + |y|)^{p-2}|x - y|^2.$$

Note that (1.3) with $y = 0$ combined with (1.2) for $y = 0$ imply $|\nabla c(x)| \sim |x|^{p-1}$. Then applying (2.2) with $x = 0$ or $y = 0$ in combination with this fact, we obtain that

$$(2.4) \qquad\qquad\qquad c(x) \sim |x|^p.$$

Notice also that (1.2) together with (1.3) and $c \in C^2(\mathbb{R}^d \backslash \{0\})$ implies

$$(2.5) \qquad\qquad\qquad |\nabla^2 c(z)| \simeq |z|^{p-2}.$$

Lemma 1. *Let $p > 1$ and $\beta \geq 1$. Then for any $\mu \geq 0$, the function*

$$t \to (\mu + t)^{p-2} t^\beta$$

is monotonic increasing in t. In fact, for every $C_0 > 0$ there exists $C_1 > 0$ such that for every $s, t, \mu \geq 0$,

$$(2.6) \qquad (\mu + t)^{p-2} t^\beta \leq C_0 (\mu + s)^{p-2} s^\beta \quad \Rightarrow \quad t \leq C_1 s.$$

Finally, a Young-type inequality holds: For any $\varepsilon > 0$ there is $C_\varepsilon > 0$ such that for any $\mu \geq 0$, $x, y \in \mathbb{R}^d$,

$$(2.7) \qquad (\mu + |x|)^{p-2}|x||y| \leq \varepsilon(\mu + |x|)^{p-2}|x|^2 + C_\varepsilon(\mu + |y|)^{p-2}|y|^2.$$

Proof. Monotonicity of $t \mapsto (\mu + t)^{p-2} t^\alpha$ follows directly from checking that the derivative of the function is non-negative. Let $\gamma = \min(\beta, p-2+\beta) > 0$ and for $C_0 > 0$ set $C_1 = C_0^{1/\beta}$. If $s, t \geq 0$ are such that

$$(\mu + t)^{p-2} t^\beta \leq C_0 (\mu + s)^{p-2} s^\beta,$$

assume for the sake of contradiction that $t > C_1 s$. Using the monotonicity we find

$$C_0(\mu + s)^{p-2} s^\beta \geq (\mu + t)^{p-2} t^\beta > (\mu + C_1 s)^{p-2}(C_1 s)^\beta$$
$$\geq C_1^\gamma (\mu + s)^{p-2} s^\beta = C_0 (\mu + s)^{p-2} s^\beta.$$

This gives the desired contradiction.

A proof of (2.7) can be found in [1, Lemma 2.3]. $\square$

3. L^∞-BOUNDS

In this section we prove Theorem 2.

Proof of Theorem 2. **Step 1: Cyclical monotonicity** Let $\widetilde{x} \in B_{1/2}$ and write $\widetilde{y} = T(\widetilde{x})$. Making the change of variables

$$x = \widetilde{x} + x', \quad y = \widetilde{x} + y', \quad \rho_0'(x') = \rho_0(x) \quad \text{and} \quad \rho_1'(y') = \rho_1(y)$$

we have by translation invariance of the cost that $T'(x') = T(x) - \widetilde{x}$ is the optimal transport map between ρ_0' and ρ_1'. Setting $\widetilde{b} = b + A\widetilde{x}$ and

$$\widetilde{E} = \int_{B_1} |T' - x' - \widetilde{b} - Ax'|^m d\rho_0',$$

it is enough to prove that

$$(3.1) \qquad |T'(0) - \widetilde{b}| \lesssim \begin{cases} \widetilde{E}^{\frac{1}{m}} & \text{if } \widetilde{E} \geq 1 \\ \widetilde{E}^{\frac{1}{m+d}} & \text{if } \widetilde{E} \leq 1. \end{cases}$$

To lighten notation we write T for T' and E for $\widetilde{E}$. Moreover we denote $y = \widetilde{y} - \widetilde{x}$. Given $(x', y') = (x', T(x')) \in B_1 \times \mathbb{R}^d$, by cyclical monotonicity we have

$$(3.2) \qquad c(y) + c(y' - x') \leq c(y') + c(y - x').$$

Let us point out that this is the only form in which we use the minimality of T.

Step 2: Preliminary estimates Note that for

$$(3.3) \qquad \varepsilon < \frac{1}{4}|A|^{-1},$$

we may define v_ε uniquely as a solution of the equation

$$(3.4) \qquad v_\varepsilon = \varepsilon\left(y - (Av_\varepsilon + \widetilde{b})\right) \quad \Leftrightarrow \quad v_\varepsilon = (\mathrm{Id} + \varepsilon A)^{-1}\varepsilon(y - \widetilde{b})$$

Further note that, if (3.3) holds and we additionally ensure

$$(3.5) \qquad \varepsilon|y - \widetilde{b}| \leq \frac{3}{8},$$

then

$$|v_\varepsilon| \leq \|\mathrm{Id} + \varepsilon A\|^{-1}|\varepsilon(y - \widetilde{b})| \leq \frac{1}{2}.$$

We will precise our choice of $\varepsilon \in (0, 1/2)$ at a later stage, but for now assume that (3.3) and (3.5) are satisfied. Then for $r \in (0, |v_\varepsilon|/2)$ to be fixed at a later stage, we may find $(x', y') = (x', T(x')) \in B_1 \times \mathbb{R}^d$ such that

$$(3.6) \qquad |x' - v_\varepsilon| \leq r, \qquad |y' - x' - b'|^m \leq r^{-d}\frac{E}{D}.$$

where

$$(3.7) \qquad b' = \widetilde{b} + Ax'.$$

Indeed, suppose that for all $(x', T(x')) \in B_r(v_\varepsilon) \times \mathbb{R}^d \Subset B_2 \times \mathbb{R}^d$ and some $M > D^{-1}$,

$$|y' - x' - b'|^m \geq \frac{ME}{r^d}.$$

Then

$$E \geq \frac{ME}{r^d} \int_{B_r(v_\varepsilon)} \mathrm{d}\rho_0 \geq MED,$$

giving a contradiction.

Note that since $r \in (0, |v_\varepsilon|/2)$, definition (3.4) of v_ε and (3.3), and definition (3.6) of x',

$$(3.8) \qquad |x'| + |v_\varepsilon| \simeq |v_\varepsilon| \simeq \varepsilon|y - \widetilde{b}|.$$

Hence it also holds that

$$(3.9) \qquad |y - \widetilde{b} - Av_\varepsilon| = \varepsilon^{-1}|v_\varepsilon| \simeq |y - \widetilde{b}|.$$

Finally, using (3.8) and the triangle inequality, we find

$$(3.10) \qquad |\widetilde{b} + Av_\varepsilon + x'| + |y - x'| \simeq |\widetilde{b}| + |y|.$$

Step 3: Applying cyclical monotonicity

We now apply (3.2) with the choice of (x', y') given in (3.6) and re-write (3.2) as

$$c(y) + c(\widetilde{b} + Av_\varepsilon) - c(\varepsilon y + (1 - \varepsilon)(\widetilde{b} + Av_\varepsilon)) - c(\varepsilon(\widetilde{b} + Av_\varepsilon) + (1 - \varepsilon)y)$$

$$\leq \left[\left(c(\widetilde{b} + Av_\varepsilon + x') - c(\widetilde{b} + Av_\varepsilon + v_\varepsilon) \right) + (c(y - x') - c(y - v_\varepsilon)) \right]$$

$$+ \left[\left(c(\widetilde{b} + Av_\varepsilon) - c(\widetilde{b} + Av_\varepsilon + x') \right) + (c(y') - c(y' - x')) \right]$$

$$(3.11) \qquad = I + II.$$

Using convexity in the form of (2.1) followed by (1.3), we find

$$I \lesssim \langle \nabla c(\widetilde{b} + Av_\varepsilon + x') - \nabla c(y - x'), x' - v_\varepsilon \rangle$$

$$\lesssim (|\widetilde{b} + Av_\varepsilon + x'| + |y - x'|)^{p-2}|\widetilde{b} + Av_\varepsilon - y + 2x'||x' - v_\varepsilon|$$

$$\overset{(3.6)}{\leq} (|\widetilde{b} + Av_\varepsilon + x'| + |y - x'|)^{p-2}|\widetilde{b} + Av_\varepsilon - y + 2x'|r.$$

Using (3.8) and (3.10), we may further simplify the estimate above to

$$(3.12) \qquad I \lesssim (|\widetilde{b}| + |y|)^{p-2}|\widetilde{b} - y|r.$$

Similarly,

$$II \lesssim \langle \nabla c(y') - \nabla c(\widetilde{b} + Av_\varepsilon), x' \rangle$$

$$\lesssim (|y'| + |\widetilde{b} + Av_\varepsilon|)^{p-2}|y' - \widetilde{b} - Av_\varepsilon||x'|$$

$$(3.13) \qquad \lesssim (|y' - \widetilde{b} - Av_\varepsilon| + |\widetilde{b} + Av_\varepsilon|)^{p-2}|y' - \widetilde{b} - Av_\varepsilon||x'| = II'.$$

On the one hand, if $|y' - \widetilde{b} - Av_\varepsilon| \lesssim \varepsilon|y - \widetilde{b} - Av_\varepsilon|$, then we deduce using the monotonicity from Lemma 1, as well as (3.8) and (3.9),

$$II' \lesssim \varepsilon^{\min(2,p)}(|y - \widetilde{b} - Av_\varepsilon| + |\widetilde{b}|)^{p-2}|y - \widetilde{b} - Av_\varepsilon||y - \widetilde{b}|$$

$$\lesssim \varepsilon^{\min(2,p)}(|y - \widetilde{b} - Av_\varepsilon| + |\widetilde{b}|)^{p-2}|y - \widetilde{b}|^2$$

$$\overset{(3.6)}{\lesssim} \varepsilon^{\min(2,p)}(|y - \widetilde{b}| + |\widetilde{b}|)^{p-2}|y - \widetilde{b}|^2.$$

On the other hand, if $|y' - \widetilde{b} - Av_\varepsilon| \gg \varepsilon|y - \widetilde{b} - Av_\varepsilon|$, using (3.8) and (3.9),

$$(3.14) \qquad |x'| + |Av_\varepsilon| \lesssim \varepsilon|y - \widetilde{b}| \lesssim \varepsilon|y - \widetilde{b} - Av_\varepsilon| \ll |y' - \widetilde{b} - Av_\varepsilon|.$$

Recalling the definition (3.7) of b' we have by triangle inequality,

$$|y' - \widetilde{b} - Av_\varepsilon| \leq |y' - b' - x'| + |Av_\varepsilon| + |x'| + |Ax'|.$$

Using (3.14) to absorb the last three terms on the right-hand side, we have in this case,

$$(3.15) \qquad |y' - \widetilde{b} - Av_\varepsilon| \lesssim |y' - b' - x'|.$$

Using the monotonicity of Lemma 1, this allows to estimate in this case

$$II' \lesssim (|y' - b' - x'| + |\widetilde{b} + Av_\varepsilon|)^{p-2}|y' - b' - x'||x'|.$$

Using now (3.14) and (3.15) we find

$$|y' - b' - x'| + |\widetilde{b} + Av_\varepsilon| \simeq |y' - b' - x'| + |\widetilde{b}|$$

so that we can post-process it to

$$II' \lesssim (|y' - b' - x'| + |\widetilde{b}|)^{p-2}|y' - b' - x'||x'|.$$

Combining both cases we conclude that

$$(3.16) \qquad II \lesssim \varepsilon^{\min(2,p)}(|y - \widetilde{b}| + |\widetilde{b}|)^{p-2}|y - \widetilde{b}|^2 + (|y' - b' - x'| + |\widetilde{b}|)^{p-2}|y' - b' - x'||x'|.$$

Putting (3.12) and (3.16) together yields

$$(3.17) \quad I + II \lesssim \varepsilon^{\min(2,p)}(|y - \widetilde{b}| + |\widetilde{b}|)^{p-2}|y - \widetilde{b}|^2$$
$$+ (|\widetilde{b}| + |y|)^{p-2}|\widetilde{b} - y|r + (|y' - b' - x'| + |\widetilde{b}|)^{p-2}|y' - b' - x'||x'|.$$

We now estimate $I + II$ from below. As $\varepsilon \in (0, 1)$, (2.3) applied to the third and fourth term on the left-hand side in (3.11) allows us to obtain the following lower bound (up to a constant depending on p only):

$$\varepsilon(1 - \varepsilon)(|\widetilde{b} + Av_\varepsilon| + |y|)^{p-2}|y - \widetilde{b} - Av_\varepsilon|^2 \lesssim I + II.$$

Choosing $\varepsilon \ll 1$ and employing (3.9), we may replace this by

$$(3.18) \qquad \varepsilon(|\widetilde{b}| + |y - \widetilde{b}|)^{p-2}|y - \widetilde{b}|^2 \lesssim I + II.$$

Collecting estimates (3.17) and (3.18) together and choosing $\varepsilon \ll 1$ to absorb the first right-hand side term in (3.17), we have shown

$$\varepsilon(|\widetilde{b}| + |y - \widetilde{b}|)^{p-2}|y - \widetilde{b}|^2 \lesssim (|\widetilde{b}| + |y|)^{p-2}|\widetilde{b} - y|r$$
$$+ (|y' - b' - x'| + |\widetilde{b}|)^{p-2}|y' - b' - x'||x'|$$

where the implicit constant depends on p, d and $|A|$. By (3.6) and the monotonicity from Lemma 1, this implies

$$(3.19) \qquad \varepsilon(|\widetilde{b}| + |y - \widetilde{b}|)^{p-2}|y - \widetilde{b}|^2 \lesssim (|\widetilde{b}| + |y - \widetilde{b}|)^{p-2}|\widetilde{b} - y|r$$
$$+ ((r^{-d}E)^{\frac{1}{m}} + |\widetilde{b}|)^{p-2}(r^{-d}E)^{\frac{1}{m}}|x'|.$$

Step 4: Proof of (3.1) for $E \geq 1$: As $E \geq 1$ it suffices to show

$$(3.20) \qquad |y - \widetilde{b}| \lesssim 1 + E^{\frac{1}{m}}.$$

Thus, we may assume without loss of generality that $|y - \widetilde{b}| \geq 1$. We choose for $0 < \delta \ll 1$ to be determined $\varepsilon = \frac{\delta}{|y-\widetilde{b}|}$. If δ is sufficiently small, then (3.3) and (3.5) are satisfied and $\varepsilon \ll 1$. Choosing $r \ll \delta \simeq |v_\varepsilon|$, but $r \sim \delta$, using that $|x'| \leq 1$, (3.19) simplifies to

$$(3.21) \qquad (|\widetilde{b}| + |y - \widetilde{b}|)^{p-2}|y - \widetilde{b}| \lesssim (E^{\frac{1}{m}} + |\widetilde{b}|)^{p-2} E^{\frac{1}{m}}.$$

Using (2.6) with $\beta = 1$, we deduce (3.20).

Step 5: Proof of (3.1) **for** $E \leq 1$: Note that in order to show (3.20) we did not use that $E \geq 1$. Thus, also if $E \leq 1$, (3.20) holds and we may assume that $|y - \widetilde{b}| \lesssim 1$. In particular, it suffices to ensure $\varepsilon \ll 1$ in order for (3.3) and (3.5) to hold. We begin by using (2.7) to estimate the right-hand side of (3.19) (up to constant) by

$$(3.22)$$
$$\varepsilon^2(|y-\widetilde{b}| + |\widetilde{b}|)^{p-2}|y-\widetilde{b}|^2 + C_\varepsilon(r + |\widetilde{b}|)^{p-2}r^2 + (|x'| + |\widetilde{b}|)^{p-2}|x'|^2 + ((r^{-d}E)^{\frac{1}{m}} + |\widetilde{b}|)^{p-2}(r^{-d}E)^{\frac{2}{m}}$$
$$\lesssim \varepsilon^2(|y-\widetilde{b}| + |\widetilde{b}|)^{p-2}|y-\widetilde{b}|^2 + C_\varepsilon(r + |\widetilde{b}|)^{p-2}r^2 + ((r^{-d}E)^{\frac{1}{m}} + |\widetilde{b}|)^{p-2}(r^{-d}E)^{\frac{2}{m}}.$$

In order to obtain the second line, we used (3.8) together with the monotonicity from Lemma 1. Note that for $\varepsilon \ll 1$, the first term may be absorbed on the left-hand side of (3.19). Choosing $r = E^{\frac{1}{m+d}}$ we have shown

$$(|y - \widetilde{b}| + |\widetilde{b}|)^{p-2}|y - \widetilde{b}|^2 \lesssim (E^{\frac{1}{m+d}} + |\widetilde{b}|)^{p-2} E^{\frac{2}{m+d}}.$$

Now (2.6) implies the desired inequality and concludes the proof. $\qquad \square$

4. Main argument

This section is devoted to proving Theorem 1.

Proof of Theorem 1. Let T be a minimizer for the c-optimal transport problem between ρ_0 and ρ_1. Let then $u(x) = T(x) - x$, $v(y) = y - T^{-1}(y)$. Recall that we have set

$$X = \{x \ : \ |u(x)| > 0\} \qquad Y = \{y \ : \ |v(y)| > 0\}.$$

We need to show that there exists X', and Y' open sets such that $|X' \backslash X| = 0 = |Y' \backslash Y|$ and such that T is a $C^{1,\alpha}$ diffeomorphism between X' and Y'. Let $\eta > 0$ and set

$$X^\eta = \{x \ : \ |u(x)| > \eta \text{ and } x \text{ is a Lebesgue point of } u\}$$
$$Y^\eta = \{y \ : \ |v(y)| > \eta \text{ and } y \text{ is a Lebesgue point of } v\}.$$

Since up to a set of Lebesgue measure 0, $X = \cup_{\eta>0} X^\eta$ and $Y = \cup_{\eta>0} Y^\eta$ it is enough to prove the statement with X^η and Y^η instead of X and Y. Let (ϕ, ψ) be optimal Kantorovich potentials so that using the same convention as in [4, 15]

$$(4.1) \qquad \psi(x) = \sup_y -\phi(y) - c(y - x).$$

We claim that ψ is twice differentiable a.e. in X^η and that ϕ is twice differentiable a.e. in Y^η. Notice that for $p \geq 2$, hypothesis (2.5) implies that ϕ and ψ are semi-convex so that the claim follows by Alexandrov Theorem. We may thus focus on $p \leq 2$ (although

the argument works also for $p \geq 2$). Since the argument for ϕ and ψ are the same we only prove the claim for ψ. We claim that for every $\bar{x} \in X^\eta$, there exists $R(\bar{x}, \eta) > 0$ such that

$$(4.2) \qquad\qquad |u(x)| \geq \eta/2 \qquad \text{for all } x \in B_{R(\bar{x},\eta)}(\bar{x}).$$

By translation we assume that $\bar{x} = 0$ and set $b = u(0)$. Notice that by definition of X^η we have $|b| > \eta$. Let

$$E = \frac{1}{R^{p+d}} \int_{B_{2R}} |u(x) - b|^p.$$

Note that we may assume without loss of generality that $\rho_0(0) > 0$ and hence if R is sufficiently small, $\inf_{B_r \in B_R} \frac{\rho_0(B_r)}{r^d} > 0$ as $\rho_0 \in C^0$. By Theorem 2 and scaling we then have the L^∞ bound,

$$\sup_{B_{R/4}} |u(x) - b| \lesssim R \left(E^{\frac{1}{p+d}} + E^{\frac{1}{p}} \right).$$

Thus if $E \leq 1$ (so that $E^{\frac{1}{p}} \leq E^{\frac{1}{p+d}}$) and R is small enough,

$$\sup_{B_{R/4}} |u(x) - b| \lesssim R \leq \eta/2$$

whereas if $E \geq 1$ (so that $E^{\frac{1}{p+d}} \leq E^{\frac{1}{p}}$) and R is small enough,

$$\sup_{B_{R/4}} |u(x) - b| \lesssim R E^{\frac{1}{p}} = \left(\frac{1}{R^d} \int_{B_R} |u - b|^p \right)^{1/p} \leq \eta/2$$

by definition of Lebesgue points. Using triangle inequality this concludes the proof of (4.2). Let

$$\widetilde{X}^\eta = \cup_{\bar{x} \in X^\eta} B_{R(\bar{x},\eta)}(\bar{x})$$

so that $\widetilde{X}^\eta$ is an open set with $X^\eta \subset \widetilde{X}^\eta$. Recalling that (ϕ, ψ) are optimal Kantorovich potentials we notice that for every x the supremum in (4.1) is attained at $y = T(x)$. In particular if $x \in \widetilde{X}^\eta$ we may restrict the supremum to y such that $|y - x| \geq \eta/2$. We finally claim that in $\widetilde{X}^\eta$, ψ is $C(\eta)-$semi-concave. Indeed, let $\hat{x} \in \widetilde{X}^\eta$. By definition of $\widetilde{X}^\eta$, there exists $\bar{x} \in X^\eta$ such that $\hat{x} \in B_{R(\bar{x},\eta)}(\bar{x})$. Let $r \ll \eta$ be such that $B_r(\hat{x}) \subset B_{R(\bar{x},\eta)}(\bar{x})$. Then for every $x \in B_r(\hat{x})$, if y is such that $|y - \hat{x}| \leq \eta/4$, we have $|y - x| \leq r + \eta/4 < \eta/2$ and thus since $x \in B_{R(\bar{x},\eta)}(\bar{x})$, $|u(x)| \geq \eta/2 > |y - x|$. Therefore,

$$\psi(x) = \sup_{y \in B_{\eta/4}(\hat{x})^c} -\phi(y) - c(y - x).$$

Now for every $(x, y) \in B_r(\hat{x}) \times B_{\eta/4}(\hat{x})^c$, the function $f_y(x) = -\phi(y) - c(y - x)$ satisfies by (2.5),

$$|D^2 f_y(x)| \lesssim |x - y|^{p-2} \lesssim \eta^{p-2}$$

and is thus semi-convex with a semi-convexity constant of the order of η^{p-2}. This concludes the proof of the claim.

By Alexandrov Theorem, ψ is twice differentiable a.e. in $\widetilde{X}^\eta$ and thus also in X^η as claimed.

We may now closely follow the argument from [4, 15]. Let $X_1^\eta \subset X^\eta$ be the set of points $\bar{x}$ such that ψ is twice differentiable at $\bar{x}$ and define similarly Y_1^η. We let then

$$X_\eta' = X_1^\eta \cap T^{-1}(Y_1^\eta) \qquad \text{and} \qquad Y_\eta' = Y_1^\eta \cap T(X_1^\eta).$$

Let $(\bar{x}, \bar{y}) \in X_\eta' \times Y_\eta'$ be such that $T(\bar{x}) = \bar{y} =: b$. We claim that for $r \ll 1$, $T \in C^{1,\alpha}(B_r(\bar{x}))$. By translation invariance of c we may assume that $\bar{x} = 0$. Since by hypothesis $0 \in X_\eta' \subset X^\eta$ we have $|b| > \eta$. Writing $y = b + y'$ (and thus $T(x) = b + T'(x)$), $c_b(x, y') = c(y' + b - x)$, $\rho_{1,b}(y') = \rho_1(y)$ and $\psi_b(x) = \psi(x)$ we see that ψ_b is a c_b−convex function and T' is the optimal transport map between ρ_0 and $\rho_{1,b}$ with $T'(0) = 0$. We then set

$$\bar{\psi}(x) = \psi_b(x) - \psi_b(0) + c_b(x, 0) - c_b(0, 0) = \psi(x) - \psi(0) + c(b - x) - c(b)$$

and

$$\bar{c}(x, y) = c_b(x, y) - c_b(x, 0) - c_b(0, y) + c_b(0, 0)$$
$$= c(y + b - x) - c(b - x) - c(y + b) + c(b).$$

Since $\bar{\psi}$ is a $\bar{c}$−convex function we see that T' is the $\bar{c}$-optimal transport map from ρ_0 to $\rho_{1,b}$. Moreover, since ψ and $c(b - \cdot)$ are twice differentiable at 0, $\bar{\psi}$ is also twice differentiable at 0 so that for some symmetric matrix A,

$$\nabla \bar{\psi}(x) = \nabla \bar{\psi}(0) + Ax + o(|x|).$$

Since

$$\nabla \bar{\psi}(0) = -\nabla_x \bar{c}(0, 0) = 0,$$

this reduces further to

$$\nabla \bar{\psi}(x) = Ax + o(|x|).$$

Letting $M = -\nabla_{xy} \bar{c}(0, 0) = \nabla^2 c(b)$ (which is non-degenerate by (2.5) and $|b| > \eta$) we have

(4.3) $$\qquad\qquad\qquad T'(x) = M^{-1} Ax + o(|x|).$$

In particular, if r is small enough then for $x \in B_r$, $|T'(x)| \ll \eta \le |b|$. Arguing similarly we have that for $y \in B_r$ also $|(T')^{-1}(y)| \ll \eta \le |b|$. Let r_0 be a small radius such that these two properties hold and let

$$\bar{\rho}_0 = \rho_0 \chi_{B_{r_0}} \qquad \text{and} \qquad \bar{\rho}_{1,b} = T' \# \bar{\rho}_0.$$

By the previous considerations, if $r \ll r_0$ then

(4.4) $$\qquad\qquad\qquad \bar{\rho}_{1,b} = \rho_{1,b} \qquad \text{in } B_r.$$

Moreover, by (2.5),

(4.5) $$\qquad -\nabla_{xy} \bar{c}(x, y) = \nabla^2 c(y + b - x) \sim_b \text{Id} \qquad \text{in } B_{r_0} \times B_{r_0}.$$

We now make the change of variables (arguing as in [15, Proof of Corollary 1.4] we have that A is positive definite) $\widetilde{x} = A^{\frac{1}{2}} x$ and $\widetilde{y} = A^{-\frac{1}{2}} My$ and thus

$$\widetilde{T}(\widetilde{x}) = A^{-\frac{1}{2}} M T'(A^{-\frac{1}{2}} \widetilde{x}), \qquad \widetilde{c}(\widetilde{x}, \widetilde{y}) = \bar{c}(A^{-\frac{1}{2}} \widetilde{x}, M^{-1} A^{\frac{1}{2}} \widetilde{y}).$$

Setting

$$\widetilde{\rho}_0(\widetilde{x}) = \frac{\rho_0(A^{-\frac{1}{2}}\widetilde{x})}{\rho_0(0)} \qquad \text{and} \qquad \widetilde{\rho}_1(\widetilde{y}) = \frac{\bar{\rho}_{1,b}(M^{-1}A^{\frac{1}{2}}\widetilde{y})}{\bar{\rho}_{1,b}(0)}$$

we may check as in [15, Proof of Corollary 1.4] that $\widetilde{T}$ is the optimal transport map between $\widetilde{\rho}_0$ and $\widetilde{\rho}_1$ for the cost $\widetilde{c}$. By (4.4) we have for $r \ll r_0$ that $\widetilde{\rho}_i \in C^\alpha(B_r)$. Moreover, by definition, $\widetilde{\rho}_0(0) = \widetilde{\rho}_1(0) = 1$. With this change of variables (4.3) becomes

$$\widetilde{T}(\widetilde{x}) = \widetilde{x} + o(|\widetilde{x}|)$$

so that

$$\lim_{R \to 0} \frac{1}{R^{2+d}} \int_{B_R} |\widetilde{T} - \widetilde{x}|^2 \widetilde{\rho}_0 = 0.$$

Finally, by (4.5), if $r \ll r_0$,

- $\widetilde{c} \in C^{2,\alpha}(B_r \times B_r)$;
- for $x \in B_r$, the map $y \mapsto \nabla_x \widetilde{c}(x,y)$ is one-to-one from B_r into $\mathbb{R}^d$;
- for $y \in B_r$, the map $x \mapsto \nabla_y \widetilde{c}(x,y)$ is one-to-one from B_r into $\mathbb{R}^d$;
- $\det \nabla_{xy}\widetilde{c} \neq 0$ for $(x,y) \in B_r \times B_r$.

We may thus apply [15, Theorem 1.1] and conclude that $T' \in C^{1,\alpha}(B_R)$ for $R \ll r$ small enough. Returning back to the original variables we find that $T \in C^{1,\alpha}(B_R)$ for a (non relabelled) R small enough. Moreover, if R is sufficiently small, $B_R \subset X'_\eta$. Arguing similarly for T^{-1} we conclude that T is a diffeomorphism between B_R and $T(B_R)$. In particular $B_R \times T(B_R) \subset X'_\eta \times Y'_\eta$ which is open and it follows that T is a global $C^{1,\alpha}$ diffeomorphism between X'_η and Y'_η. $\qquad \square$

ACKNOWLEDGMENTS

REFERENCES

[1] Acerbi, E. and Mingione, G., Regularity Results for Stationary Electro-Rheological Fluids *Arch. Rational Mech. Anal.* 164, (2002).

[2] Bouchitté, G., Jimenez, C. and Rajesh, M. A new L^∞ estimate in optimal mass transport *Proc. Amer. Math. Soc.* 135(11), 3525–3535 (2007).

[3] Caffarelli, L., González, M. D. M. and Nguyen, T. A perturbation argument for a Monge–Ampère type equation arising in optimal transportation. *Archive for rational mechanics and analysis*, 212, 359-414 (2014).

[4] De Philippis, G. and Figalli, A. Partial regularity for optimal transport maps *Publ. Math. Inst. Hautes Études Sci.* 121, 81–112 (2015).

[5] Duzaar, F. and Mingione, G. Regularity for degenerate elliptic problems via p-harmonic approximation. *Annales de l'Institut Henri Poincaré C, Analyse non linéaire*, Vol. 21, No. 5, 735-766 (2004).

[6] Figalli, A. and Kim, Y.-H. Partial regularity of Brenier solutions of the Monge-Ampère equation *Discrete Contin. Dyn. Syst.* 28(2), 559–565 (2010).

[7] Goldman, M. An ε-regularity result for optimal transport maps between continuous densities. *Rendiconti Lincei. Matematica e Applicazioni*, (2020).

[8] Goldman, M., Huesmann, M. and Otto, F. Quantitative Linearization Results for the Monge-Ampère Equation. *Communications on Pure and Applied Mathematics*, 74(12), 2483-2560 (2021).

[9] Goldman, M. and Otto, F. A variational proof of partial regularity for optimal transportation maps. *Ann. Sci. Ec. Norm. Super.* 53(5), 1209–1233(2020).

[10] Gutiérrez, C.E. and Montanari, A. L^∞-estimates in optimal transport for non quadratic costs. *Calc. Var.* 61,163 (2022).

[11] Gutiérrez, C.E. and Montanari, A. Fine properties of monotone maps arising in optimal transport for non-quadratic costs. *arXiv* 2208.00193, (2022).

[12] Gutiérrez, C.E. and Montanari, A. Differentiability of monotone maps related to non-quadratic costs. In preparation.

[13] Iwaniec, T. and Manfredi, J. Regularity of p-harmonic functions on the plane *Rev. Math. Iberoam.* 5(1-2), 1-10 (1989).

[14] Koch, L. Geometric linearisation for optimal transport with strongly p-convex cost. *Calc. Var.* 63, 87 (2024).

[15] F. Otto, M. Prod'homme, and T. Ried Variational Approach to Regularity of Optimal Transport Maps: General Cost Functions. *Ann. PDE* 7(17), 2021.

[16] Villani, C. Topics in optimal transportation *Amer. Math. Soc., Providence, RI* Graduate Studies in Mathematics 58(2003).

CMAP, CNRS, École polytechnique, Institut Polytechnique de Paris, 91120 Palaiseau, France
Email address: `michael.goldman@cnrs.fr`

MPI, MIS Inselstrasse 22, 04103 Leipzig, Germany
Email address: `lukas.koch@mis.mpg.de`