

Lipschitz changes of variables between perturbations of log-concave measures

MARIA COLOMBO, ALESSIO FIGALLI AND YASH JHAVERI

Abstract. Extending a result of Caffarelli, we provide global Lipschitz changes of variables between compactly supported perturbations of log-concave measures. The result is based on a combination of ideas from optimal transportation theory and a new Pogorelov-type estimate. In the case of radially symmetric measures, Lipschitz changes of variables are obtained for a much broader class of perturbations.

Mathematics Subject Classification (2010): 49Q20 (primary); 35J96, 26D10 (secondary).

1. Introduction

In [4], Caffarelli built Lipschitz changes of variables between log-concave probability measures. More precisely, he showed that if $V, W \in C_{\text{loc}}^{1,1}(\mathbb{R}^n)$ are convex functions with $D^2V(x) \leq \Lambda_V \text{Id}$ and $\lambda_W \text{Id} \leq D^2W(x)$ for a.e. $x \in \mathbb{R}^n$ with $0 < \Lambda_V, \lambda_W < \infty$, then there exists a Lipschitz map $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ such that $T_{\#}(e^{-V(x)} dx) = e^{-W(x)} dx$ ¹ and

$$\|\nabla T\|_{L^\infty(\mathbb{R}^n)} \leq \sqrt{\Lambda_V/\lambda_W}. \quad (1.1)$$

¹ Given two finite Borel measures μ and ν and a Borel map $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$, recall that $T_{\#}\mu = \nu$ if

$$\int_{\mathbb{R}^n} \varphi(y) d\nu(y) = \int_{\mathbb{R}^n} \varphi(T(x)) d\mu(x) \quad \forall \varphi \text{ Borel and bounded.}$$

M. Colombo acknowledges the support of the Gruppo Nazionale per l'Analisi Matematica, la Probabilità e le loro Applicazioni (GNAMPA) of the Istituto Nazionale di Alta Matematica (INdAM), of Dr. Max Rössler, of the Walter Haefner Foundation and of the ETH Zurich Foundation. A. Figalli has been partially supported by NSF Grant DMS-1262411 and NSF Grant DMS-1361122.

Received October 13, 2015; accepted September 27, 2016.

Published online December 2017.

The map T is obtained via optimal transportation. It is the unique solution of the Monge problem for quadratic cost

$$\min \left\{ \int_{\mathbb{R}^n} |x - T(x)|^2 e^{-V(x)} dx : T_{\#}(e^{-V(x)} dx) = e^{-W(x)} dx \right\}$$

(see Section 2 for more details and [12] for a completely different construction of a Lipschitz change of variables in this setting). We note that a particularly important feature of Caffarelli’s result is that the bound (1.1) is *independent* of the dimension n .

A consequence of Caffarelli’s result is the possible deduction of certain functional inequalities (such as log-Sobolev or Poincaré-type inequalities) for log-concave measures from their corresponding Gaussian versions. For instance, denoting the standard Gaussian measure on \mathbb{R}^n by γ_n , consider the Gaussian log-Sobolev inequality,

$$\int_{\mathbb{R}^n} f^2 \ln f d\gamma_n \leq \int_{\mathbb{R}^n} |\nabla f|^2 d\gamma_n + \left(\int_{\mathbb{R}^n} f^2 d\gamma_n \right) \ln \left(\int_{\mathbb{R}^n} f^2 d\gamma_n \right),$$

which holds for every function $f \in W^{1,2}(\mathbb{R}^n)$. For any measure ν such that there exists a Lipschitz change of variables between ν and the Gaussian measure, namely $\nu = T_{\#}\gamma_n$, we deduce, applying twice the change of variable formula, that

$$\begin{aligned} \int_{\mathbb{R}^n} f^2 \ln f d\nu &= \int_{\mathbb{R}^n} f(T)^2 \ln f(T) d\gamma_n \\ &\leq \int_{\mathbb{R}^n} |\nabla[f \circ T]|^2 d\gamma_n + \left(\int_{\mathbb{R}^n} f(T)^2 d\gamma_n \right) \ln \left(\int_{\mathbb{R}^n} f(T)^2 d\gamma_n \right) \\ &\leq \|\nabla T\|_{L^\infty(\mathbb{R}^n)}^2 \int_{\mathbb{R}^n} |\nabla f(T)|^2 d\gamma_n + \left(\int_{\mathbb{R}^n} f(T)^2 d\gamma_n \right) \ln \left(\int_{\mathbb{R}^n} f(T)^2 d\gamma_n \right) \\ &= \|\nabla T\|_{L^\infty(\mathbb{R}^n)}^2 \int_{\mathbb{R}^n} |\nabla f|^2 d\nu + \left(\int_{\mathbb{R}^n} f^2 d\nu \right) \ln \left(\int_{\mathbb{R}^n} f^2 d\nu \right). \end{aligned}$$

Therefore, ν enjoys a log-Sobolev inequality with constant $\|\nabla T\|_{L^\infty(\mathbb{R}^n)}^2$.

Besides the natural consequences described in [4] and above, Caffarelli’s Theorem has found numerous applications in various fields: indeed, it can be used to transfer isoperimetric inequalities, to obtain correlation inequalities, and more (see, for instance, [6, 7, 11, 13]). Some recent extensions and variations of Caffarelli’s Theorem can be found in, for example, [14, 15, 17].

In this paper, we extend the result of Caffarelli by building Lipschitz changes of variables between perturbations of V and W that are not necessarily convex. Perturbations of log-concave measures (in particular, perturbations of Gaussian measures) appear, for instance, in quantum physics as a means to help understanding solutions to physical theories with nonlinear equations of motion. In cases where an explicit

solution is unknown, perturbations of log-concave measures can be used to yield approximate solutions.

We let $\mathcal{P}(X)$ denote the space of probability measures on a metric space X . The main result of the paper is the following:

Theorem 1.1. *Let $V \in C_{\text{loc}}^{1,1}(\mathbb{R}^n)$ be such that $e^{-V(x)} dx \in \mathcal{P}(\mathbb{R}^n)$. Suppose that $V(0) = \inf_{\mathbb{R}^n} V$ and there exist constants $0 < \lambda, \Lambda < \infty$ for which $\lambda \text{Id} \leq D^2V(x) \leq \Lambda \text{Id}$ for a.e. $x \in \mathbb{R}^n$. Moreover, let $R > 0$, $q \in C_c^0(B_R)$, and $c_q \in \mathbb{R}$ be such that $e^{-V(x)+c_q-q(x)} dx \in \mathcal{P}(\mathbb{R}^n)$. Assume that $-\lambda_q \text{Id} \leq D^2q$ in the sense of distributions for some constant $\lambda_q \geq 0$. Then, there exists a constant $C = C(R, \lambda, \Lambda, \lambda_q) > 0$, independent of n , such that the optimal transport map T that takes $e^{-V(x)} dx$ to $e^{-V(x)+c_q-q(x)} dx$ satisfies*

$$\|\nabla T\|_{L^\infty(\mathbb{R}^n)} \leq C. \tag{1.2}$$

The crucial point here is that the estimate on the Lipschitz constant of the optimal transport map is independent of dimension, as it is in Caffarelli’s results for log-concave measures.

In the case of spherically symmetric measures, we are able to weaken the assumptions on both the log-concave measure and its perturbation and still obtain a global Lipschitz change of variables. In particular, the Lipschitz constant is controlled only by the L^∞ -norm of the positive and negative parts of the perturbation q , denoted by q^+ and q^- . In the following theorem, we first analyze the 1-dimensional problem:

Theorem 1.2. *Let $V : \mathbb{R} \rightarrow \mathbb{R} \cup \{\infty\}$ be a convex function and $q : \mathbb{R} \rightarrow \mathbb{R}$ be a bounded function such that $e^{-V(x)} dx, e^{-V(x)-q(x)} dx \in \mathcal{P}(\mathbb{R})$. Then, the optimal transport T that takes $e^{-V(x)} dx$ to $e^{-V(x)-q(x)} dx$ is Lipschitz and satisfies*

$$\|\log T'\|_{L^\infty(\mathbb{R})} \leq \|q^+\|_{L^\infty(\mathbb{R})} + \|q^-\|_{L^\infty(\mathbb{R})}. \tag{1.3}$$

We remark that while the map T in Theorem 1.2 is only unique up to sets of $e^{-V(x)} dx$ -measure zero, arguing by approximation, we can find a particular transport T for which the estimate on $\log T'$ in (1.3) is satisfied almost everywhere in \mathbb{R} . Applying this 1-dimensional result to radially symmetric densities, we obtain the following:

Theorem 1.3. *Let $V : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{\infty\}$ be a convex, radially symmetric function and $q : \mathbb{R}^n \rightarrow \mathbb{R}$ be a bounded, radially symmetric function such that $e^{-V(x)} dx, e^{-V(x)-q(x)} dx \in \mathcal{P}(\mathbb{R}^n)$. Then, the optimal transport T that takes $e^{-V(x)} dx$ to $e^{-V(x)-q(x)} dx$ is Lipschitz and satisfies*

$$e^{-\|q^+\|_{L^\infty(\mathbb{R}^n)} - \|q^-\|_{L^\infty(\mathbb{R}^n)}} \text{Id} \leq \nabla T(x) \leq e^{\|q^+\|_{L^\infty(\mathbb{R}^n)} + \|q^-\|_{L^\infty(\mathbb{R}^n)}} \text{Id} \tag{1.4}$$

for a.e. $x \in \mathbb{R}^n$.

Note that the assumption $e^{-V(x)-q(x)} dx \in \mathcal{P}(\mathbb{R}^n)$ in Theorems 1.2 and 1.3, unlike in Theorem 1.1, is nonrestrictive. Since q is not required to be compactly supported, the normalization constant making $e^{-V(x)-q(x)} dx$ a probability measure if it were not already can simply be absorbed into q .

We further remark that the 1-dimensional estimate in Theorem 1.2 is false in higher dimensions when one does not assume that the densities are radially symmetric. More precisely, taking the reference measure $e^{-V(x)} dx$ to be the standard Gaussian measure, the estimate

$$\|D^2\phi - \text{Id}\|_{L^\infty(\mathbb{R}^n)} \leq C\|q\|_{L^\infty(\mathbb{R}^n)} \tag{1.5}$$

cannot be true for $n > 1$ (see Remark 5.2 to understand the relationship between (1.3) and (1.5) for $n = 1$). This is manifest if we recall that the Monge-Ampère equation linearizes to the Poisson equation, which does not enjoy $C_{\text{loc}}^{1,1}$ estimates for bounded right-hand side. In other words, given V and q to be chosen, letting ϕ_ε be the potential such that $\nabla\phi_\varepsilon$ takes $e^{-V(x)} dx$ to $e^{-V(x)-\varepsilon q(x)} dx$ (for simplicity, we omit the normalization constant that makes $e^{-V(x)-\varepsilon q(x)} dx$ a probability measure), and setting $\psi_\varepsilon(x) = (\phi_\varepsilon(x) - |x|^2/2)/\varepsilon$, we have that

$$\begin{aligned} \Delta\psi_\varepsilon + O(\varepsilon) &= \frac{\log \det \nabla^2\phi_\varepsilon}{\varepsilon} = \frac{-V + V(\nabla\phi_\varepsilon) + \varepsilon q(\nabla\phi_\varepsilon)}{\varepsilon} \\ &= \langle x, \nabla\psi_\varepsilon \rangle + q(\nabla\phi_\varepsilon) + O(\varepsilon) \end{aligned}$$

for every $\varepsilon > 0$. The estimate (1.5) implies that $\sup_{\varepsilon>0} \|D^2\psi_\varepsilon\|_{L^\infty(\mathbb{R}^n)} < \infty$ and, therefore, the existence of a $C_{\text{loc}}^{1,1}$ solution to the Poisson equation with bounded right-hand side, an impossibility in higher dimensions.

Although this heuristic argument is convincing, the details of the proof are rather delicate, and we give them in the Appendix for completeness.

ACKNOWLEDGEMENTS. Y. Jhaveri would like to thank Pablo Stinga for helpful conversations. Part of this work was done while the authors were guests of the FIM at ETH Zürich in the Fall of 2014; the hospitality of the Institute is gratefully acknowledged.

2. Preliminaries

We begin with some preliminaries on optimal transportation and the Monge-Ampère equation, and we fix some notation.

Let $\mu, \nu \in \mathcal{P}(\mathbb{R}^n)$. The Monge optimal transport problem for quadratic cost consists of finding the most efficient way to take μ to ν given that the transportation cost to move from a point x to a point y is $|x - y|^2$. Hence, one is led to minimize

$$\text{cost}(T) := \int_{\mathbb{R}^n} |x - T(x)|^2 d\mu(x)$$

among all maps T such that $T_{\#}\mu = \nu$. A relaxed formulation of Monge’s problem, due to Kantorovich, is to minimize

$$\int_{\mathbb{R}^n \times \mathbb{R}^n} |x - y|^2 d\pi(x, y)$$

among all transport plans π , namely the measures $\pi \in \mathcal{P}(\mathbb{R}^n \times \mathbb{R}^n)$ whose marginals are μ and ν . By a classical theorem of Brenier [2], the existence and uniqueness of an optimal transport plan are guaranteed when μ is absolutely continuous and μ and ν have finite second moments. Additionally, the optimality of a transport plan π is equivalent to $\pi = (\text{Id} \times \nabla\phi)_{\#}\mu$ where ϕ is a convex function, often called the potential associated to the optimal transport. As a consequence, it follows that in the Monge problem, unique optimal maps exist as gradients of convex functions.

Theorem 2.1. *Let $\mu, \nu \in \mathcal{P}(\mathbb{R}^n)$ such that $\mu = f(x) dx$ and*

$$\int_{\mathbb{R}^n} |x|^2 d\mu(x) + \int_{\mathbb{R}^n} |y|^2 d\nu(y) < \infty.$$

Then, there exists a unique (up to sets of μ -measure zero) optimal transport T taking μ to ν . Moreover, there is a convex function $\phi : \mathbb{R}^n \rightarrow \mathbb{R}$ such that $T = \nabla\phi$.

A direct consequence of Brenier’s characterization of optimal transports as gradients of convex functions is that

$$\langle x - y, T(x) - T(y) \rangle \geq 0 \quad \text{for a.e. } x, y \in \mathbb{R}^n, \tag{2.1}$$

which follows immediately from the monotonicity of gradients of convex functions.

Suppose now that $\mu = f(x) dx$ and $\nu = g(y) dy$, and let ϕ be a convex function such that $T = \nabla\phi$ for T the optimal transport that takes μ to ν . Assuming that $T = \nabla\phi$ is a smooth diffeomorphism, the standard change of variables formula implies that

$$f(x) = g(T(x)) \det \nabla T(x).$$

Hence, assuming that $g > 0$, we see that ϕ is a solution to the Monge-Ampère equation

$$\det D^2\phi = \frac{f}{g \circ \nabla\phi}.$$

This formal link between optimal transportation and Monge-Ampère (since, to deduce the above equation, we assumed that T was already smooth) is at the heart of the regularity of optimal transport maps (see, for instance, [8] for more details). In particular, Caffarelli showed the following in [3] (see also [9, Theorem 4.5.2]):

Theorem 2.2. *Let $X, Y \subset \mathbb{R}^n$ be bounded open sets, and $f : X \rightarrow \mathbb{R}^+$ and $g : Y \rightarrow \mathbb{R}^+$ be probability densities locally bounded away from zero and infinity. If Y is convex, then for any set $X' \subset\subset X$, the optimal transport $T = \nabla\phi : X \rightarrow Y$ between $f(x) dx$ and $g(y) dy$ is of class $C^{0,\alpha}(X')$ for some $\alpha > 0$. In addition, if $f \in C_{\text{loc}}^{k,\beta}(X)$ and $g \in C_{\text{loc}}^{k,\beta}(Y)$ for some $k \in \mathbb{N} \cup \{0\}$ and $\beta \in (0, 1)$, then $\phi \in C_{\text{loc}}^{k+2,\beta}(X)$.*

As mentioned in [1], Caffarelli’s regularity result on optimal transports can be extended to the case where f and g are defined on all of \mathbb{R}^n and assumed to be locally bounded away from zero and infinity. Lastly, we note that optimal transport maps are stable under approximation (see [18]). In particular, let f_j and g_j be locally uniformly bounded probability densities such that $f_j \rightarrow f$ and $g_j \rightarrow g$ in L^1_{loc} . Then, the associated potentials $\phi_j \rightarrow \phi$ locally uniformly and $\nabla\phi_j \rightarrow \nabla\phi$ in measure.

We fix the following additional notation:

- B_R ball of radius R centered at the origin
- \mathcal{L}^n n -dimensional Lebesgue measure
- \mathcal{H}^d d -dimensional Hausdorff measure
- \mathbb{S}^{n-1} unit sphere in \mathbb{R}^n
- ω_n n -dimensional Lebesgue measure of $B_1 \subset \mathbb{R}^n$.

3. Lipschitz changes of variables between log-concave measures

We begin with two useful results of Caffarelli (see [4]). They provide some motivation, and we briefly recall their proofs both for completeness and because we shall need them later.

Lemma 3.1. *Let $\mu = f(x) dx$, $\nu = g(x) dx \in \mathcal{P}(\mathbb{R}^n)$ with finite second moments and $\nabla\phi = T$ be the optimal transport taking μ to ν . Assume that $\log f \in L^\infty_{\text{loc}}(\mathbb{R}^n)$ and that g is bounded away from zero in the ball B_j for some $j > 0$ and vanishes outside B_j . Then,*

$$T(x) \rightarrow j \frac{x}{|x|} \quad \text{uniformly as } |x| \rightarrow \infty.$$

In particular, for any fixed $\varepsilon > 0$ and for all $\alpha \in \mathbb{S}^{n-1}$, the function $\phi(x + \varepsilon\alpha) + \phi(x - \varepsilon\alpha) - 2\phi(x) \rightarrow 0$ as $|x| \rightarrow \infty$.

Proof. We begin by noticing that, as a consequence of Theorem 2.2, T is continuous on \mathbb{R}^n and, in particular, the map T is well defined at every point.

Let $x_0 \in \mathbb{R}^n$ and $\theta \in (0, \pi/4)$ be fixed, and consider the cone with vertex at $T(x_0)$ and pointing in the x_0 -direction

$$\Gamma := \left\{ y \in \mathbb{R}^n : \angle(x_0, y - T(x_0)) \leq \frac{\pi}{2} - \theta \right\}.$$

By (2.1) we see that

$$\angle(x - x_0, T(x) - T(x_0)) \leq \frac{\pi}{2};$$

hence,

$$\begin{aligned} \angle(x - x_0, x_0) \leq \angle(x - x_0, T(x) - T(x_0)) + \angle(x_0, T(x) - T(x_0)) \leq \pi - \theta \\ \forall x \text{ so that } T(x) \in \Gamma, \end{aligned}$$

and so, up to a set of measure zero, the preimage of Γ under T is contained in the (concave) cone

$$\Omega := \{x \in \mathbb{R}^n : \angle(x_0, x - x_0) \leq \pi - \theta\}.$$

Moreover, since $T_{\#}\mu = \nu$,

$$\inf_{x \in B_j} g(x) \mathcal{L}^n(\Gamma \cap B_j) \leq \nu(\Gamma \cap B_j) = \nu(\Gamma) \leq \mu(\Omega).$$

Let $B = B_{(|x_0| \tan \theta)/2}$, and notice that $\Omega \subseteq \mathbb{R}^n \setminus B$. This proves that $\mu(\Omega) \leq \mu(\mathbb{R}^n \setminus B)$.

Now, $\mu(\mathbb{R}^n \setminus B) \rightarrow 0$ as $|x_0| \rightarrow \infty$ since B covers \mathbb{R}^n as $|x_0| \rightarrow \infty$. Recalling that g is bounded away from zero in B_j , we have that

$$\lim_{|x_0| \rightarrow \infty} \mathcal{L}^n(\Gamma \cap B_j) = 0.$$

Letting $\theta \rightarrow 0$, we see that $T(x_0) \rightarrow j \frac{x_0}{|x_0|}$. As the point x_0 was fixed arbitrarily, $\nabla\phi(x) = T(x) \rightarrow j \frac{x}{|x|}$ uniformly as $|x| \rightarrow \infty$. Thus, ϕ behaves like the cone $j|x|$ at infinity. In particular, for any fixed $\varepsilon > 0$ and for all $\alpha \in \mathbb{S}^{n-1}$, the function $\phi(x + \varepsilon\alpha) + \phi(x - \varepsilon\alpha) - 2\phi(x) \rightarrow 0$ as $|x| \rightarrow \infty$. \square

Thanks to Lemma 3.1, [4,5], Caffarelli proved the following result.

Theorem 3.2. *Let $V, W \in C_{loc}^{1,1}(\mathbb{R}^n)$ be such that $e^{-V(x)} dx, e^{-W(x)} dx \in \mathcal{P}(\mathbb{R}^n)$. Suppose there exist constants $0 < \lambda_W, \Lambda_V < \infty$ such that $D^2V(x) \leq \Lambda_V \text{Id}$ and $\lambda_W \text{Id} \leq D^2W(x)$ for a.e. $x \in \mathbb{R}^n$. Then, the optimal transport T that takes $e^{-V(x)} dx$ to $e^{-W(x)} dx$ is globally Lipschitz and satisfies*

$$\|\nabla T\|_{L^\infty(\mathbb{R}^n)} \leq \sqrt{\Lambda_V/\lambda_W}. \tag{3.1}$$

Proof. By the stability of optimal transports, we may assume that W is equal to infinity outside the ball B_j for some fixed $j > 0$. Indeed, define

$$W^j := \begin{cases} W & \text{in } B_j \\ \infty & \text{in } \mathbb{R}^n \setminus B_j \end{cases}$$

and $c_j \in (0, \infty)$ such that

$$\int_{\mathbb{R}^n} e^{c_j - W^j(x)} dx = 1.$$

Clearly, $e^{c_j - W^j} \rightarrow e^{-W}$ in $L^1(\mathbb{R}^n)$ as $j \rightarrow \infty$. Hence, if we prove (3.1) for the optimal transport T^j that takes $e^{-V(x)} dx$ to $e^{c_j - W^j(x)} dx$, letting $j \rightarrow \infty$ we obtain the same estimate for T .

Also, by Theorem 2.2, the convex potential $\phi : \mathbb{R}^n \rightarrow \mathbb{R}$ associated to the optimal transport T is of class C^3 ; therefore, ϕ satisfies the Monge-Ampère equation

$$\det D^2\phi(x) = \frac{e^{-V(x)}}{e^{-W(\nabla\phi(x))}},$$

or equivalently,

$$\log \det D^2\phi(x) = -V(x) + W(\nabla\phi(x)). \tag{3.2}$$

For fixed $\varepsilon > 0$, we define the incremental quotient of a function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ at $(x, \alpha) \in \mathbb{R}^n \times \mathbb{S}^{n-1}$ by

$$f^\varepsilon(x, \alpha) := f(x + \varepsilon\alpha) + f(x - \varepsilon\alpha) - 2f(x).$$

By the convexity of ϕ we see that $\phi^\varepsilon \geq 0$. Also, it follows by Lemma 3.1 that $\phi^\varepsilon \rightarrow 0$ as $|x| \rightarrow \infty$. Thus ϕ^ε attains a global maximum at some $(x_0, \alpha_0) \in \mathbb{R}^n \times \mathbb{S}^{n-1}$. Up to a rotation, we assume that $\alpha_0 = e_1$. Thus,

$$0 = \nabla\phi^\varepsilon(x_0, e_1) = \nabla\phi(x_0 + \varepsilon e_1) + \nabla\phi(x_0 - \varepsilon e_1) - 2\nabla\phi(x_0). \tag{3.3}$$

Moreover, because e_1 is the maximal direction,

$$0 = \partial_\beta\phi^\varepsilon(x_0, e_1) = \varepsilon\langle \nabla\phi(x_0 + \varepsilon e_1) - \nabla\phi(x_0 - \varepsilon e_1), \beta \rangle \quad \forall \beta \perp e_1.$$

Taking $\beta = e_i$ for $i \neq 1$ and utilizing (3.3), we see that all the components but the first of $\nabla\phi(x_0 + \varepsilon e_1)$, $\nabla\phi(x_0 - \varepsilon e_1)$, and $\nabla\phi(x_0)$ are equal. Let $\delta := \langle \nabla\phi(x_0 + \varepsilon e_1) - \nabla\phi(x_0 - \varepsilon e_1), e_1 \rangle / 2$, and observe that, by (3.3),

$$\begin{aligned} \langle \nabla\phi(x_0), e_1 \rangle \pm \delta &= \frac{1}{2} \langle \nabla\phi(x_0 + \varepsilon e_1) + \nabla\phi(x_0 - \varepsilon e_1), e_1 \rangle \\ &\pm \frac{1}{2} \langle \nabla\phi(x_0 + \varepsilon e_1) - \nabla\phi(x_0 - \varepsilon e_1), e_1 \rangle = \langle \nabla\phi(x_0 \pm \varepsilon e_1), e_1 \rangle. \end{aligned}$$

Hence, we conclude that

$$\nabla\phi(x_0 \pm \varepsilon e_1) = \nabla\phi(x_0) \pm \delta e_1. \tag{3.4}$$

Another consequence of ϕ^ε achieving a maximum at x_0 is

$$D^2\phi(x_0 + \varepsilon e_1) + D^2\phi(x_0 - \varepsilon e_1) - 2D^2\phi(x_0) \leq 0. \tag{3.5}$$

We recall that

$$\lim_{\varepsilon \rightarrow 0^+} \frac{\det(A + \varepsilon B) - \det(A)}{\varepsilon} = \det(A) \operatorname{tr}(A^{-1}B) \tag{3.6}$$

for all square matrices A and B with A invertible. Also, if we set $F(A) := \log \det A$, since F is concave on the space of positive semidefinite $n \times n$ matrices and recalling (3.6), we have

$$\nabla F(D^2\phi(x_0)) = (D^2\phi(x_0))^{-1}$$

and

$$F\left(D^2\phi(x_0 \pm \varepsilon e_1)\right) \leq F\left(D^2\phi(x_0)\right) + \left\langle \left(D^2\phi(x_0)\right)^{-1}, D^2\phi(x_0 \pm \varepsilon e_1) - D^2\phi(x_0) \right\rangle.$$

In particular, from (3.5) and the convexity of ϕ , we deduce that

$$F\left(D^2\phi(x_0 + \varepsilon e_1)\right) + F\left(D^2\phi(x_0 - \varepsilon e_1)\right) - 2F\left(D^2\phi(x_0)\right) \leq 0.$$

Now, let us, for fixed $\varepsilon > 0$, consider the incremental quotient of (3.2) at (x_0, e_1) . Using (3.4), we realize that

$$V^\varepsilon(x_0, e_1) \geq W^\delta(\nabla\phi(x_0), e_1). \tag{3.7}$$

Observe that

$$V^\varepsilon(x_0, e_1) = \int_0^\varepsilon \left(\int_{-t}^t \left\langle D^2V(x_0 + se_1)e_1, e_1 \right\rangle ds \right) dt;$$

hence,

$$V^\varepsilon(x_0, e_1) \leq \Lambda_V \varepsilon^2. \tag{3.8}$$

Furthermore, from (3.4), we similarly see that

$$\lambda_W \delta^2 \leq W^\delta(\nabla\phi(x_0), e_1).$$

Combining this estimate with (3.8) and (3.7), we get

$$\varepsilon \sqrt{\Lambda_V / \lambda_W} \geq \delta. \tag{3.9}$$

Set $C := \sqrt{\Lambda_V / \lambda_W}$. Since

$$\phi^\varepsilon(x_0, e_1) = \int_0^\varepsilon \left\langle \nabla\phi(x_0 + te_1) - \nabla\phi(x_0 - te_1), e_1 \right\rangle dt,$$

the convexity of ϕ , (3.4), and (3.9) give us that

$$\phi^\varepsilon(x_0, e_1) = 2\delta\varepsilon \leq 2C\varepsilon^2,$$

and so

$$\|\nabla T\|_{L^\infty(\mathbb{R}^n)} = \|D^2\phi\|_{L^\infty(\mathbb{R}^n)} \leq 2C.$$

Notice that this is the desired estimate up to a factor 2. We use a bootstrapping argument to remove this factor. Suppose that $0 \leq \|D^2\phi\|_{L^\infty(\mathbb{R}^n)} \leq a_0$ for some $a_0 > C$. For any $0 \leq t \leq \varepsilon$, by (3.4) and (3.9),

$$|\langle \nabla\phi(x_0 + te_1) - \nabla\phi(x_0 - te_1), e_1 \rangle| \leq \min\{2\varepsilon C, 2a_0 t\}.$$

Thus,

$$\phi^\varepsilon(x_0, e_1) \leq \int_0^{\frac{\varepsilon C}{a_0}} 2a_0 t \, dt + \int_{\frac{\varepsilon C}{a_0}}^\varepsilon 2\varepsilon C \, dt = \varepsilon^2 \frac{(2Ca_0 - C^2)}{a_0}.$$

In other words, if $\|D^2\phi\|_{L^\infty(\mathbb{R}^n)} \leq a_0$ with $a_0 > C$, then

$$\|D^2\phi\|_{L^\infty(\mathbb{R}^n)} \leq \frac{(2Ca_0 - C^2)}{a_0}.$$

Starting with $a_0 = 2C$ and repeating the above procedure an infinite number of times, we prove (3.1) since C uniquely solves $(2Ca - C^2)/a = a$. \square

Remark 3.3. Notice that the above proof relies only on the local behavior of our densities e^{-V} and e^{-W} . In particular, the bounds on the Hessians of V and W are only used near the maximum point x_0 and its image $\nabla\phi(x_0)$, respectively. This simple observation will play an important role in the proof of Theorem 1.1.

Remark 3.4. The above result is not ideal. Indeed, if $V = W$, then $T = \text{Id}$ and one would like to have the bound $\|\nabla T\|_{L^\infty(\mathbb{R}^n)} \leq 1$ instead of $\|\nabla T\|_{L^\infty(\mathbb{R}^n)} \leq \sqrt{\Lambda_V/\lambda_V}$.

4. Compactly supported perturbations: proof of Theorem 1.1

In the following lemma, we prove an upper bound on how far points travel under the transport map when the source measure is perturbed in a certain fixed ball B_P . We capture and quantify that our perturbations are compactly supported. Lemma 4.1 will be applied in the proof of Theorem 1.1 to the inverse transport.

Furthermore, given our convex function V , we consider, for $j \in \mathbb{N}$,

$$V^j := \begin{cases} V & \text{in } B_j \\ \infty & \text{in } \mathbb{R}^n \setminus B_j, \end{cases} \tag{4.1}$$

and we approximate $e^{-V(x)} dx$ with compactly supported measures $e^{c_j - V^j(x)} dx$. This approximation is in the spirit of Caffarelli’s approximation in the proof of Theorem 3.2. It allows us to find maximum points of a suitable function and guarantees that they do not escape to infinity in the proof of Theorem 1.1. This approximation procedure is purely technical. Hence, on a first reading of Lemma 4.1, the reader may just take $j = \infty$.

Lemma 4.1. *Let $V \in C^\infty(\mathbb{R}^n)$ be such that $\mu := e^{-V(x)} dx \in \mathcal{P}(\mathbb{R}^n)$. Suppose that $V(0) = \inf_{\mathbb{R}^n} V$ and there exist constants $0 < \lambda, \Lambda < \infty$ such that $\lambda \text{Id} \leq D^2V(x) \leq \Lambda \text{Id}$ for all $x \in \mathbb{R}^n$. Moreover, let $P > 0$, $p \in C_c^\infty(B_P)$, and $c_p \in \mathbb{R}$ be such that $e^{-V(x) + c_p - p(x)} dx \in \mathcal{P}(\mathbb{R}^n)$. Given $j > P$, set V^j as in (4.1) and choose $c_{p,j} \in (0, \infty)$ such that $\mu_{p,j} := e^{c_{p,j} - V^j(x) + c_p - p(x)} dx \in \mathcal{P}(\mathbb{R}^n)$. If T is*

the optimal transport map that takes $\mu_{p,j}$ to μ , then there exist constants $P' = P'(P, \lambda, \Lambda, \|p\|_{L^\infty(\mathbb{R}^n)}) > 0$ and $j' = j'(n, V(0), P, \lambda, \Lambda, \|p\|_{L^\infty(\mathbb{R}^n)}) > P$ such that for all $j \in [j', \infty]$,

$$T(B_P) \subseteq B_{P'}. \tag{4.2}$$

Even though this lemma is not independent of dimension as written (specifically, j' depends on n), the dimensional dependence does not affect the constant P' and disappears in the limit as $j \rightarrow \infty$. Thus, we can indeed prove a global estimate on the optimal transport taking $e^{-V(x)} dx$ to $e^{-V(x)+c_q-q(x)} dx$ that is independent of dimension.

Lemma 4.1 is written under slightly different assumptions than Theorem 1.1. In particular, besides the obvious additional regularity assumptions on V and its perturbation, made only for simplicity, we have not required that the perturbation be semiconvex. That said, if we assume the the distributional Hessian of p is indeed bounded below by $-\lambda_p \text{Id}$, then we can replace the dependence on $\|p\|_{L^\infty(\mathbb{R}^n)}$ with a dependence on λ_p , as explained in the following remark.

Remark 4.2. Let p be a function compactly supported in B_P that satisfies the semi-convexity condition $D^2 p \geq -\lambda_p \text{Id}$ in the sense of distributions. Then, its L^∞ -norm is controlled by a constant depending only on P and λ_p (in particular, it is independent of dimension):

$$\|p\|_{L^\infty(\mathbb{R}^n)} \leq 4\lambda_p P^2. \tag{4.3}$$

First, up to convolving p with a standard convolution kernel, we can assume that p is smooth. Then, we observe that every 1-dimensional restriction $f_\alpha(t) = p(t\alpha)$, for $t \in \mathbb{R}$ and $\alpha \in \mathbb{S}^{n-1}$, is compactly supported in $[-P, P]$ and has second derivative bounded below by $-\lambda_p$. This implies that

$$\|f'_\alpha\|_{L^\infty(\mathbb{R})} \leq 2\lambda_p P. \tag{4.4}$$

Indeed, suppose to the contrary that $f'_\alpha(t_0) > 2\lambda_p P$ for some $t_0 \in [-P, P]$. By integration, we would get

$$0 = f'_\alpha(P) \geq f'_\alpha(t_0) + \int_{t_0}^P f''_\alpha(\tau) d\tau > 2\lambda_p P + \lambda_p(P - t_0) > 0,$$

which is impossible. This proves (4.4), and (4.3) holds by integrating.

Before proceeding with the proof of Lemma 4.1, we recall a Talagrand-type transport inequality. Given $\mu_1, \mu_2 \in \mathcal{P}(\mathbb{R}^n)$, we denote the squared Wasserstein distance between μ_1 and μ_2 by $W_2^2(\mu_1, \mu_2)$ (see [18, Chapter 6] for the general definition), and we consider their relative entropy

$$\text{Ent}(\mu_2|\mu_1) := \begin{cases} \int_{\mathbb{R}^n} \log\left(\frac{d\mu_2}{d\mu_1}\right) d\mu_2 & \text{if } \mu_2 \ll \mu_1 \\ \infty & \text{otherwise.} \end{cases}$$

Here, $d\mu_2/d\mu_1$ is the relative density of μ_2 with respect to μ_1 . If $\mu_1 = e^{-V(x)} dx$ for some $V \in C^2(\mathbb{R}^n)$ such that $D^2V(x) \geq \lambda_V \text{Id}$ for all $x \in \mathbb{R}^n$, we have that (see [6], applied in the particular case when μ_1 and μ_2 are probability measures)

$$W_2^2(\mu_1, \mu_2) \leq \frac{2}{\lambda_V} \text{Ent}(\mu_2|\mu_1). \tag{4.5}$$

In our applications, $W_2^2(\mu_1, \mu_2)$ coincides with the cost of the optimal transport taking μ_2 to μ_1 .

Proof of Lemma 4.1. Notice first that, as a consequence of Theorem 2.2, T is continuous.

Assume there exists a point $x_0 \in B_P$ with $T(x_0) \notin B_{10P}$ (otherwise, the statement is true with $P' = 10P$). We show that $T(x_0) \in B_{P'}$ for some $P' = P'(P, \lambda, \Lambda, \|p\|_{L^\infty(\mathbb{R}^n)}) > 0$ that will be chosen later. Let

$$\bar{x} := x_0 + 3P \frac{T(x_0) - x_0}{|T(x_0) - x_0|},$$

and define the constant C_0 and ball B by $C_0P = |T(x_0) - x_0|$ and $B := B_P(\bar{x})$. Also, let $F : B \rightarrow \mathbb{R}^n$ be the projection of a point $y \in B$ onto the hyperplane through $T(x_0)$ and perpendicular to $y - x_0$. The map F is well-defined because $x_0 \notin B$ (see Figure 4.1). Let us assume that $j' > 6P$, so that $B \subseteq B_{j'}$.

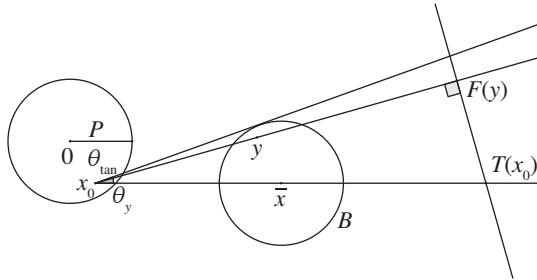


Figure 4.1. The optimal transport sends B far away.

By (2.1), we have that

$$\langle y - x_0, T(y) - T(x_0) \rangle \geq 0 \quad \forall y \in B,$$

and as $F(y)$ is the closest point to y in the set $\{z \in \mathbb{R}^n : \langle y - x_0, z - T(x_0) \rangle \geq 0\}$,

$$|T(y) - x_0| \geq |F(y) - x_0| \quad \forall y \in B$$

(see Figure 4.1). Given any $y \in B$, either x_0, y , and \bar{x} determine a plane, call it Γ_y , within which $x_0, F(y)$, and $T(x_0)$ determine a right triangle, or x_0, y , and \bar{x} are collinear. Thus,

$$|F(y) - x_0| = C_0P \cos \theta_y$$

where θ_y is the angle between $F(y) - x_0$ and $T(x_0) - x_0$. Now, $\Gamma_y \cap \partial B$ is a circle of radius P centered at \bar{x} . Letting θ_{tan} be the angle between the line through x_0 and tangent to $\Gamma_y \cap \partial B$ and the line through $T(x_0)$ and x_0 , we see that $\theta_y \leq \theta_{tan}$. (While there are two such tangent lines, the angles they determine with the line through $T(x_0)$ and x_0 are the same. Again, see Figure 4.1.) Moreover, $|x_0 - \bar{x}| = 3P$ and $\cos \theta_{tan} = 2\sqrt{2}/3$. Consequently,

$$|F(y) - x_0| \geq C_0 P \cos \theta_{tan} \geq \frac{C_0 2\sqrt{2}P}{3}$$

and

$$|T(y) - y| \geq |T(y) - x_0| - |y - x_0| > \frac{C_0 2\sqrt{2}P}{3} - 4P \quad \forall y \in B.$$

Since $V(0) = \inf_{\mathbb{R}^n} V(x)$ and $\lambda \text{Id} \leq D^2 V(x) \leq \Lambda \text{Id}$, by restricting V to 1-dimensional lines through the origin we have that

$$V(0) + \frac{\lambda}{2}|x|^2 \leq V(x) \leq V(0) + \frac{\Lambda}{2}|x|^2 \quad \forall x \in \mathbb{R}^n; \tag{4.6}$$

hence, as $B \subseteq B_{6P}$,

$$V(x) \leq V(0) + 18\Lambda P^2 \quad \forall x \in B.$$

We now estimate $\text{cost}(T)$. Since $B_P \cap B = \emptyset$ and $B \subseteq B_j$, we have

$$\begin{aligned} \text{cost}(T) &\geq \int_B |T(x) - x|^2 e^{c_{p,j} - V^j(x) + c_p} dx \\ &\geq \left[\frac{C_0 2\sqrt{2}P}{3} - 4P \right]^2 e^{c_{p,j} - V(0) - 18\Lambda P^2 + c_p} \mathcal{L}^n(B_P). \end{aligned} \tag{4.7}$$

Furthermore, we claim that the following upper bound on $\text{cost}(T)$ holds:

$$\text{cost}(T) \leq \frac{6}{\lambda} \|p\|_{L^\infty(\mathbb{R}^n)} e^{c_{p,j} + c_p + \|p\|_{L^\infty(\mathbb{R}^n)}} \mu(B_P). \tag{4.8}$$

To see this, first, apply the Talagrand-type transport inequality (4.5) with $\mu_1 = \mu$ and $\mu_2 = \mu_{p,j}$ to find that

$$\text{cost}(T) \leq \frac{2}{\lambda} \int_{\mathbb{R}^n} (c_{p,j} + c_p - p(x)) e^{c_{p,j} - V^j(x) + c_p - p(x)} dx. \tag{4.9}$$

Second, choose $j' > 6P$, so that

$$\int_{\mathbb{R}^n \setminus B_{j'}} e^{-V(0) - \frac{\lambda}{2}|x|^2 + \|p\|_{L^\infty(\mathbb{R}^n)}} dx \leq 1 - \exp\left(-\|p\|_{L^\infty(\mathbb{R}^n)} \int_{B_P} e^{-V(0) - \frac{\lambda}{2}|x|^2} dx\right).$$

Notice that $|c_p| \leq \|p\|_{L^\infty(\mathbb{R}^n)}$ since

$$e^{-c_p} = \int_{\mathbb{R}^n} e^{-p(x)} d\mu(x). \tag{4.10}$$

So, for every $j \geq j'$, observe that

$$\begin{aligned} e^{-c_{p,j}} &= \int_{B_j} e^{-V(x)+c_p-p(x)} dx = 1 - \int_{\mathbb{R}^n \setminus B_j} e^{-V(x)+c_p} dx \\ &\geq 1 - \int_{\mathbb{R}^n \setminus B_{j'}} e^{-V(0)-\frac{\lambda}{2}|x|^2+\|p\|_{L^\infty(\mathbb{R}^n)} dx} \\ &\geq \exp\left(-\|p\|_{L^\infty(\mathbb{R}^n)} \int_{B_P} e^{-V(0)-\frac{\lambda}{2}|x|^2} dx\right), \end{aligned}$$

and then, recalling that $c_{p,j} > 0$, note

$$\begin{aligned} c_{p,j} &\leq \|p\|_{L^\infty(\mathbb{R}^n)} \int_{B_P} e^{-V(0)-\frac{\lambda}{2}|x|^2} dx \\ &\leq \|p\|_{L^\infty(\mathbb{R}^n)} e^{c_{p,j}+c_p+\|p\|_{L^\infty(\mathbb{R}^n)}} \mu(B_P). \end{aligned} \tag{4.11}$$

Now, use Jensen's inequality on (4.10) and that p is supported in B_P to deduce that

$$\begin{aligned} c_p - \int_{B_P} p(x) e^{c_{p,j}-V_j(x)+c_p-p(x)} dx &\leq \int_{B_P} p(x) e^{c_{p,j}+c_p} \left[e^{-c_{p,j}-c_p} - e^{-p(x)} \right] d\mu(x) \\ &\leq 2\|p\|_{L^\infty(\mathbb{R}^n)} e^{c_{p,j}+c_p+\|p\|_{L^\infty(\mathbb{R}^n)}} \mu(B_P). \end{aligned} \tag{4.12}$$

Finally, combine (4.9), (4.11), and (4.12) to see that (4.8) holds as claimed.

In particular, since $\mu(B_P) \leq e^{-V(0)} \mathcal{L}^n(B_P)$, we have that

$$\text{cost}(T) \leq \frac{6}{\lambda} \|p\|_{L^\infty(\mathbb{R}^n)} e^{c_{p,j}-V(0)+c_p+\|p\|_{L^\infty(\mathbb{R}^n)}} \mathcal{L}^n(B_P), \tag{4.13}$$

provided that $j \geq j'$. Thus, (4.7) and (4.13) imply that

$$C_0 \leq C' := 3\sqrt{2} + \frac{9e^{9\Lambda P^2 + \frac{\|p\|_{L^\infty(\mathbb{R}^n)}}{2}}}{2P} \left[\frac{\|p\|_{L^\infty(\mathbb{R}^n)}}{\lambda} \right]^{1/2}.$$

This proves the existence of an upper bound on C_0 depending only on P , λ , Λ and $\|p\|_{L^\infty(\mathbb{R}^n)}$.

Taking $P' := (C' + 1)P$, we deduce that

$$|T(x_0)| \leq |T(x_0) - x_0| + |x_0| \leq C_0P + P \leq P',$$

which proves (4.2). □

The following result is a Pogorelov-type a priori estimate on pure second derivatives of the potential associated to our optimal transport. This technique is inspired by Pogorelov’s original argument for the classical Monge-Ampère equation [16]. In our case, we face the additional difficulty of constructing an auxiliary function h that compensates for the concavity of our perturbation and the growth of our convex function at infinity. Assuming that our auxiliary function attains a finite maximum, we provide a quantitative estimate on the value of h at its finite maximum. This result contains and overcomes the primary obstacles to demonstrating that our optimal transport is globally Lipschitz.

Before stating the result, we introduce some constants and an auxiliary function ψ , all depending only on the constants R , λ , Λ , and λ_q that appear in Theorem 1.1. Define the constants $P > 0$ and $Q > 0$ by

$$P := \frac{2\lambda_q + 4\lambda_q R}{\lambda} + 1 + R \quad \text{and} \quad Q := \frac{\lambda}{2\lambda_q} + 1 + R; \tag{4.14}$$

let $\bar{\psi} \in C^2([0, \infty))$ be given by

$$\begin{aligned} \bar{\psi}(t) &:= \int_0^t \int_0^s \vartheta(r) \, dr \, ds, \\ \vartheta(r) &:= \begin{cases} \lambda_q & r \in [0, R] \\ -\lambda_q r + \lambda_q + \lambda_q R & r \in [R, Q] \\ \frac{\lambda_q \lambda^2 r}{4\lambda_q^2 + 8\lambda_q^2 R - \lambda^2} - \frac{2\lambda_q^2 \lambda + 4\lambda_q^2 \lambda R + \lambda_q \lambda^2 + \lambda_q \lambda^2 R}{4\lambda_q^2 + 8\lambda_q^2 R - \lambda^2} & r \in [Q, P] \\ 0 & r \in [P, \infty); \end{cases} \end{aligned} \tag{4.15}$$

and let $\psi \in C^2(\mathbb{R}^n)$ be defined by

$$\psi(y) := \bar{\psi}(|y|). \tag{4.16}$$

Observe that the function $\bar{\psi}$ is defined in such a way that $\bar{\psi}'' \geq -\lambda/2$ in $[0, \infty)$, $\bar{\psi} = \lambda_q | \cdot |^2/2$ on $[0, R]$, and $\bar{\psi}'$ is supported in B_P (see Figure 4.2).

Proposition 4.3. *Let V , λ , Λ , R , q , λ_q , and c_q be defined as in Theorem 1.1. Assume, additionally, that V and q are smooth. Let P , $\bar{\psi}$, and ψ be defined as in (4.14), (4.15), and (4.16). Given $j > P$, set V^j as in (4.1) and choose $c_{q,j} \in (0, \infty)$ such that $e^{c_{q,j} - V^j(x) + c_q - q(x)} \, dx \in \mathcal{P}(\mathbb{R}^n)$. Also, let $\phi \in C^\infty(\mathbb{R}^n)$ solve*

$$\det D^2 \phi = \frac{e^{-V}}{e^{c_{q,j} - V^j(\nabla \phi) + c_q - q(\nabla \phi)}},$$

and assume that there exist constants j' , $P' > 0$ such that for all $j \in [j', \infty)$,

$$\nabla \phi(\mathbb{R}^n \setminus B_{P'}) \subseteq \mathbb{R}^n \setminus B_P, \tag{4.17}$$

or equivalently, that $[\nabla\phi]^{-1}(B_P) \subseteq B_{P'}$. If

$$h(x, \alpha) := \phi_{\alpha\alpha}(x)e^{\psi(\nabla\phi(x))} \tag{4.18}$$

attains a maximum at some point (x_0, α_0) among all possible $(x, \alpha) \in \mathbb{R}^n \times \mathbb{S}^{n-1}$, then there exists a constant $C = C(R, P', \lambda, \Lambda, \lambda_q) > 0$, yet independent of n , such that

$$h(x_0, \alpha_0) \leq C.$$

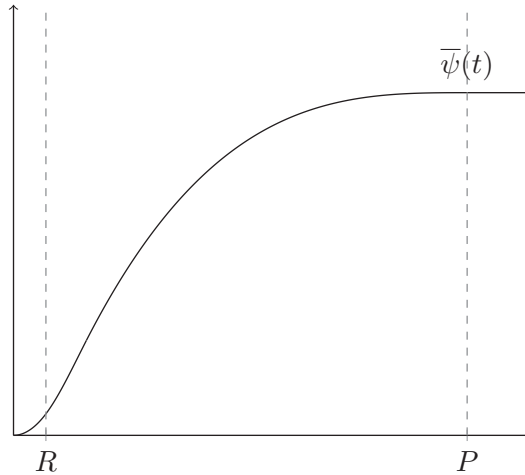


Figure 4.2. The graph of $\bar{\psi}$.

Proof. Since, by assumption, (x_0, α_0) is a maximum point of h , we have $\sup_{|\alpha|=1} \phi_{\alpha\alpha}(x_0) = \phi_{\alpha_0\alpha_0}(x_0)$. This implies that α_0 is an eigenvector of $D^2\phi(x_0)$. Therefore, up to a rotation, we assume that $\alpha_0 = e_1$ and that $D^2\phi$ is diagonal at x_0 . Throughout this proof, the function h is seen as a function of the variable x with α_0 fixed. Then, at x_0 we compute that

$$0 = (\log h)_i = \frac{\phi_{11i}}{\phi_{11}} + \psi_k(\nabla\phi)\phi_{ki}, \tag{4.19}$$

for all $1 \leq i \leq n$, and

$$0 \geq \phi^{ij}(\log h)_{ij} = \phi^{ij} \left[\frac{\phi_{11ij}}{\phi_{11}} - \frac{\phi_{11i}\phi_{11j}}{\phi_{11}^2} + \psi_k(\nabla\phi)\phi_{kij} + \psi_{kl}(\nabla\phi)\phi_{ik}\phi_{jl} \right] \tag{4.20}$$

where we denote the inverse matrix of (ϕ_{ij}) by (ϕ^{ij}) .

Let $\tilde{V}^j := V^j - c_{q,j} + q - c_q$. Using (3.6), we differentiate the equation

$$\log \det D^2\phi = -V + \tilde{V}^j(\nabla\phi) \tag{4.21}$$

in the e_1 -direction twice to obtain

$$\phi^{ij}\phi_{1ij} = -V_1 + \tilde{V}_i^j(\nabla\phi)\phi_{1i}$$

and

$$\phi^{ij}\phi_{11ij} - \phi^{il}\phi^{kj}\phi_{1ij}\phi_{1kl} = -V_{11} + \tilde{V}_i^j(\nabla\phi)\phi_{11i} + \tilde{V}_{ij}^j(\nabla\phi)\phi_{1i}\phi_{1j}. \quad (4.22)$$

By (4.20) and (4.22), we deduce that at x_0

$$\begin{aligned} 0 &\geq \phi^{il}\phi^{kj}\phi_{1ij}\phi_{1kl} - V_{11} + \tilde{V}_i^j(\nabla\phi)\phi_{11i} + \tilde{V}_{ij}^j(\nabla\phi)\phi_{1i}\phi_{1j} \\ &\quad - \frac{\phi^{ij}\phi_{11i}\phi_{11j}}{\phi_{11}} + \phi_{11}\phi^{ij}\psi_k(\nabla\phi)\phi_{kij} + \phi_{11}\phi^{ij}\psi_{kl}(\nabla\phi)\phi_{ik}\phi_{jl}. \end{aligned} \quad (4.23)$$

We estimate each term in (4.23) from below. Recall that (ϕ_{ij}) and (ϕ^{ij}) are diagonal at x_0 . Therefore, $\phi^{ii} = 1/\phi_{ii}$, and we see that

$$\phi^{il}\phi^{kj}\phi_{1ij}\phi_{1kl} - \frac{\phi^{ij}\phi_{11i}\phi_{11j}}{\phi_{11}} = \sum_{i=1}^n \sum_{k=2}^n \phi^{ii}\phi^{kk}\phi_{1ik}^2 \geq 0$$

and

$$\tilde{V}_{ij}^j(\nabla\phi)\phi_{1i}\phi_{1j} = \tilde{V}_{11}^j(\nabla\phi)\phi_{11}^2.$$

Because h has a maximum at e_1 among all directions,

$$\phi_{11}(x_0) \geq \phi_{ii}(x_0), \quad (4.24)$$

and so

$$\phi_{11}\phi^{ij}\psi_{kl}(\nabla\phi)\phi_{ik}\phi_{jl} = \phi_{11}\psi_{ii}(\nabla\phi)\phi_{ii} \geq \psi_{ii}(\nabla\phi)\phi_{ii}^2.$$

Additionally, differentiating (4.21) in the e_k -direction, we have that

$$\phi^{ij}\phi_{kij} = -V_k + \tilde{V}_i^j(\nabla\phi)\phi_{ki}.$$

By (4.19), it then follows that

$$\begin{aligned} \tilde{V}_i^j(\nabla\phi)\phi_{11i} + \phi_{11}\phi^{ij}\psi_k(\nabla\phi)\phi_{kij} &= \tilde{V}_i^j(\nabla\phi)\phi_{11i} + \psi_k(\nabla\phi)(-V_k + \tilde{V}_i^j(\nabla\phi)\phi_{ki})\phi_{11} \\ &= -\psi_k(\nabla\phi)V_k\phi_{11}, \end{aligned}$$

and, consequently, (4.23) becomes

$$0 \geq \tilde{V}_{11}^j(\nabla\phi)\phi_{11}^2 + \sum_{i=1}^n \psi_{ii}(\nabla\phi)\phi_{ii}^2 - \psi_k(\nabla\phi)V_k\phi_{11} - \Lambda. \quad (4.25)$$

If $x_0 \in \mathbb{R}^n \setminus B_{P'}$, (4.17) implies that $\nabla\phi(x_0) \in \mathbb{R}^n \setminus B_P$. Then, $\psi_k(\nabla\phi)V_k\phi_{11} = 0$ since the gradient of ψ is zero outside B_P by construction. If, on the other hand, $x_0 \in B_{P'}$, then

$$\psi_k(\nabla\phi)V_k\phi_{11} \leq \Lambda P' \|\nabla\psi\|_{L^\infty(\mathbb{R}^n)}\phi_{11}.$$

(Here, we have used that $V(0) = \inf_{\mathbb{R}^n} V$ and that $D^2V \leq \Lambda \text{Id}$ to show V_k is bounded above by $\Lambda P'$.) In both cases, we deduce that

$$\psi_k(\nabla\phi)V_k\phi_{11} \leq \tilde{C}\phi_{11}$$

for a constant \tilde{C} depending only on R, P', λ, Λ , and λ_q . Thus, by (4.25), we have that

$$0 \geq \tilde{V}_{11}^j(\nabla\phi)\phi_{11}^2 + \sum_{i=1}^n \psi_{ii}(\nabla\phi)\phi_{ii}^2 - \tilde{C}\phi_{11} - \Lambda. \tag{4.26}$$

We claim that

$$\tilde{V}_{11}^j(\nabla\phi)\phi_{11}^2 + \sum_{i=1}^n \psi_{ii}(\nabla\phi)\phi_{ii}^2 \geq \frac{\lambda}{2}\phi_{11}^2. \tag{4.27}$$

Indeed, let us consider two cases, according to whether or not $\nabla\phi(x_0)$ belongs to B_R . If $\nabla\phi(x_0) \in B_R$, then

$$\tilde{V}_{11}^j(\nabla\phi)\phi_{11}^2 + \sum_{i=1}^n \psi_{ii}(\nabla\phi)\phi_{ii}^2 \geq \lambda\phi_{11}^2 - \lambda_q\phi_{11}^2 + \lambda_q\phi_{11}^2 = \lambda\phi_{11}^2,$$

and (4.27) follows. In the case that $\nabla\phi(x_0) \notin B_R$, we compute the derivatives of ψ in terms of the derivatives of $\bar{\psi}$. Observe that

$$\psi_i(y) = \frac{\bar{\psi}'(|y|)y_i}{|y|} \quad \text{and} \quad \psi_{ii}(y) = \bar{\psi}''(|y|)\frac{y_i^2}{|y|^2} + \frac{\bar{\psi}'(|y|)}{|y|} \left(1 - \frac{y_i^2}{|y|^2}\right).$$

Thus,

$$\psi_{ii}(\nabla\phi) \geq \bar{\psi}''(|\nabla\phi|)\frac{\phi_i^2}{|\nabla\phi|^2} \geq -\frac{\lambda}{2}\frac{\phi_i^2}{|\nabla\phi|^2}$$

since $\bar{\psi}'' \geq -\lambda/2$ in $[0, \infty)$ and $\bar{\psi}' \geq 0$. Then, (4.24) implies that

$$\sum_{i=1}^n \psi_{ii}(\nabla\phi)\phi_{ii}^2 \geq -\frac{\lambda}{2}\sum_{i=1}^n \frac{\phi_i^2}{|\nabla\phi|^2}\phi_{ii}^2 \geq -\frac{\lambda}{2}\phi_{11}^2\sum_{i=1}^n \frac{\phi_i^2}{|\nabla\phi|^2} = -\frac{\lambda}{2}\phi_{11}^2. \tag{4.28}$$

As $\nabla\phi(x_0) \notin B_R$, we know $\tilde{V}_{11}^j(\nabla\phi(x_0)) = V_{11}^j(\nabla\phi(x_0))$. It follows that

$$\tilde{V}_{11}^j(\nabla\phi)\phi_{11}^2 \geq \lambda\phi_{11}^2. \tag{4.29}$$

By (4.28) and (4.29), we deduce that (4.27) holds in this case as well.

Combining (4.26) and (4.27), we observe that

$$0 \geq \frac{\lambda}{2} \phi_{11}^2 - \tilde{C} \phi_{11} - \Lambda. \tag{4.30}$$

Solving the quadratic equation in (4.30), we find that

$$\phi_{11}(x_0) \leq \frac{\tilde{C} + \sqrt{\tilde{C}^2 + 2\lambda\Lambda}}{\lambda} \leq 2\tilde{C}/\lambda + \sqrt{2\Lambda/\lambda}.$$

As ψ is bounded in \mathbb{R}^n by definition, it follows that

$$h(x) \leq h(x_0) \leq \phi_{11}(x_0)e^{\|\psi\|_{L^\infty(\mathbb{R}^n)}} \leq C$$

for a constant C depending on $R, P', \lambda, \Lambda,$ and $\lambda_q,$ yet independent of $n,$ as desired. □

Notice that if $\lambda_q = 0,$ then $\bar{\psi} = 0.$ In this case, the constant \tilde{C} found in the proof above is zero, and we recover the global Lipschitz constant obtained by Caffarelli in Theorem 3.2 up to a factor of $\sqrt{2}$ (this is a better bound than the one provided by the proof of Theorem 3.2 before the final bootstrapping argument).

Proof of Theorem 1.1. We first prove the statement assuming that V and q are smooth. For every $j > R$ set V^j as in (4.1), and choose $c_{q,j} \in (0, \infty)$ such that $e^{c_{q,j}-V^j(x)+c_q-q(x)} dx \in \mathcal{P}(\mathbb{R}^n).$ Let T^j be the optimal transport map that takes $e^{-V(x)} dx$ to $e^{c_{q,j}-V^j(x)+c_q-q(x)} dx.$ Since the density $e^{c_{q,j}-V^j+c_q-q}$ is supported in a convex set, smooth on its support, and is bounded from above and below by positive constants, by Theorem 2.2, we deduce that $T^j \in C^\infty(\mathbb{R}^n).$ By the stability of optimal transport maps, it suffices to show that for all $j \geq j'$ (j' to be chosen possibly depending on n) we have that

$$\|\nabla T^j\|_{L^\infty(\mathbb{R}^n)} \leq C \tag{4.31}$$

for some constant $C > 0$ depending only on $R, \lambda, \Lambda,$ and $\lambda_q.$

Let $P, \psi,$ and h be defined as in (4.14), (4.16), and (4.18). Applying Lemma 4.1 to the optimal transport $[T^j]^{-1},$ we see that there exist constants j' and $P' = P'(R, \lambda, \Lambda, \lambda_q) > 0$ (see Remark 4.2) such that $[T^j]^{-1}(B_P) \subseteq B_{P'}$ for all $j \in [j', \infty];$ that is, letting $\nabla\phi = T^j$ (for simplicity we omit in ϕ the dependence on $j,$ which can be any number greater than j' in the following),

$$\nabla\phi(\mathbb{R}^n \setminus B_{P'}) \subseteq \mathbb{R}^n \setminus B_P. \tag{4.32}$$

We split the proof in two cases, according whether or not h achieves a maximum in $\Omega = \mathbb{R}^n \times \mathbb{S}^{n-1}.$ If there exists $(x_0, \alpha_0) \in \Omega$ such that

$$h(x_0, \alpha_0) = \sup_{\Omega} h(x, \alpha),$$

then we apply Proposition 4.3 and see that

$$\sup_{\mathbb{S}^{n-1}} \|\phi_{\alpha\alpha}\|_{L^\infty(\mathbb{R}^n)} \leq \|h\|_{L^\infty(\Omega)} \leq C,$$

which proves (4.31).

Otherwise, we consider the maxima of h in $\Omega_m := \bar{B}_m \times \mathbb{S}^{n-1}$ with $m \in \mathbb{N}$.

Let

$$h(x_m, \alpha_m) = \sup_{\Omega_m} h(x, \alpha).$$

Notice that $h(x_m, \alpha_m)$ is nondecreasing (and not definitively constant) and $|x_m| \uparrow \infty$ as $m \rightarrow \infty$. Now, consider the functions h^ε approximating h defined by

$$h^\varepsilon(x, \alpha) := [\phi(x + \varepsilon\alpha) + \phi(x - \varepsilon\alpha) - 2\phi(x)]e^{\psi(\nabla\phi(x))} \quad \forall (x, \alpha) \in \Omega.$$

Since ϕ is smooth, we know that $h^\varepsilon \rightarrow h$ locally uniformly in Ω as $\varepsilon \rightarrow 0$. Furthermore, by Lemma 3.1,

$$\lim_{|x| \rightarrow \infty} h^\varepsilon(x, \alpha) = 0 \tag{4.33}$$

uniformly with respect to x and α . Since $h^\varepsilon \geq 0$ (by the convexity of ϕ), the function $h^\varepsilon(x, \alpha)$ has a finite maximum point $(x^\varepsilon, \alpha^\varepsilon)$.

We claim that for sufficiently small ε (possibly depending on n and on the sequence $\{(x_m, \alpha_m)\}_{m \in \mathbb{N}}$)

$$x^\varepsilon \notin B_{P'}. \tag{4.34}$$

Indeed, let m_0 and m_1 be such that $x_{m_0} \notin B_{P'}$ and $h(x_{m_1}, \alpha_{m_1}) > h(x_{m_0}, \alpha_{m_0})$. Since h^ε converges to h locally uniformly, there exists $\varepsilon_0 > 0$ such that

$$|h^\varepsilon(x, \alpha) - h(x, \alpha)| \leq \frac{h(x_{m_1}, \alpha_{m_1}) - h(x_{m_0}, \alpha_{m_0})}{4}$$

for every $x \in \bar{B}_{|x_{m_1}|+1}, \alpha \in \mathbb{S}^{n-1}$, and $\varepsilon \leq \varepsilon_0$. So, for every $\varepsilon \leq \varepsilon_0$, we have that

$$\begin{aligned} h^\varepsilon(x_{m_1}, \alpha_{m_1}) &\geq h(x_{m_1}, \alpha_{m_1}) - |h^\varepsilon(x_{m_1}, \alpha_{m_1}) - h(x_{m_1}, \alpha_{m_1})| \\ &\geq \frac{3h(x_{m_1}, \alpha_{m_1}) + h(x_{m_0}, \alpha_{m_0})}{4}. \end{aligned} \tag{4.35}$$

Thus,

$$\begin{aligned} h^\varepsilon(x, \alpha) &\leq h(x, \alpha) + |h^\varepsilon(x, \alpha) - h(x, \alpha)| \\ &\leq h(x_{m_0}, \alpha_{m_0}) + \frac{h(x_{m_1}, \alpha_{m_1}) - h(x_{m_0}, \alpha_{m_0})}{4} \\ &= \frac{h(x_{m_1}, \alpha_{m_1}) + 3h(x_{m_0}, \alpha_{m_0})}{4} < \frac{3h(x_{m_1}, \alpha_{m_1}) + h(x_{m_0}, \alpha_{m_0})}{4} \end{aligned} \tag{4.36}$$

for every $x \in \overline{B}_{|x_{m_0}|}$, $\alpha \in \mathbb{S}^{n-1}$, and $\varepsilon \leq \varepsilon_0$. Since $B_{P'} \subseteq B_{|x_{m_0}|}$, (4.35) and (4.36) imply that $h^\varepsilon(x, \alpha) \leq h^\varepsilon(x_{m_1}, \alpha_{m_1})$ in $B_{P'}$. Therefore, h^ε satisfies (4.34) for every $\varepsilon \leq \varepsilon_0$.

Recall that ψ is constant outside B_P . Then, by (4.32) and (4.34), we know that for every $\varepsilon \leq \varepsilon_0$, the function $e^{\psi(\nabla\phi(x))}$ is locally constant around x^ε . Therefore, $(x^\varepsilon, \alpha^\varepsilon)$ is also a local maximum point for the incremental quotient $\phi(x + \varepsilon\alpha) + \phi(x - \varepsilon\alpha) - 2\phi(x)$. Moreover, outside B_R the function $V^j - c_{q,j} + q - c_q$ is convex as it coincides with $V^j - c_{q,j} - c_q$. So, proceeding as in the proof of Theorem 3.2 (cf. Remark 3.3), we conclude that (4.31) is also proved in the case that h is not guaranteed to achieve a maximum in Ω .

In order to remove the smoothness assumptions on V and q , we approximate V and q by convolution (adding a small constant to ensure these approximations define probability measures). Then, from what we have shown above, the approximate transports are all globally and uniformly Lipschitz. Thanks to the stability of optimal transports, passing to the limit, we prove (1.2). □

5. Bounded perturbations in 1-dimension and in the radially symmetric case: proofs of Theorems 1.2 and 1.3

Our goal now is to produce optimal global Lipschitz estimates under strong symmetry but weak regularity assumptions on our log-concave measures. Notice that when our perturbation is zero, we recover that our optimal transport is the identity map (cf. Remark 3.4). We begin in 1-dimension and with a technical lemma relating the behavior of our convex base and the cumulative distribution function of the log-concave probability measure it defines.

Lemma 5.1. *Let $V : \mathbb{R} \rightarrow \mathbb{R}$ be a convex function such that $e^{-V(x)} dx \in \mathcal{P}(\mathbb{R})$ and $x_0 \in \mathbb{R}$ be such that $V(x_0) = \inf_{\mathbb{R}} V$. Define $\Phi, \Psi : \mathbb{R} \rightarrow (0, 1)$ by*

$$\Phi(x) := \int_{-\infty}^x e^{-V(t)} dt \quad \text{and} \quad \Psi(x) := \int_x^\infty e^{-V(t)} dt = 1 - \Phi(x). \tag{5.1}$$

Then,

$$V(x) - V(y) \leq \log \Phi(y) - \log \Phi(x) \quad \forall x \leq y \leq x_0 \tag{5.2}$$

and

$$V(x) - V(y) \geq \log \Psi(y) - \log \Psi(x) \quad \forall x_0 \leq x \leq y. \tag{5.3}$$

Proof. Since an analogous argument proves (5.3), we only show (5.2); in other words, we prove that the function $\log \Phi + V$ is nondecreasing in $(-\infty, x_0]$. Let $\hat{x} = \inf\{x : V(x) = V(x_0)\}$. The function $\log \Phi + V$ is clearly nondecreasing in $[\hat{x}, x_0]$, whenever this interval is not a single point. Moreover, it is locally Lipschitz and its derivative is $e^{-V}/\Phi + V'$. Hence, it suffices to show that the derivative is

nonnegative in $(-\infty, \hat{x})$. Since V' is nonincreasing in $(-\infty, \hat{x})$ and by the change of variables formula, we have that for a.e. $x \in (-\infty, \hat{x})$

$$V'(x)\Phi(x) \geq \int_{-\infty}^x V'(t)e^{-V(t)} dt = -e^{-V(x)},$$

which proves our claim. □

Proof of Theorem 1.2. By approximating V with a sequence of convex functions $V_j \rightarrow V$ such that $e^{-V_j(x)} dx \in \mathcal{P}(\mathbb{R})$ and that are finite on \mathbb{R} , we can assume that $V < \infty$ on \mathbb{R} . This reduction follows from the stability of optimal transport maps. Recall that, as a consequence of the push-forward condition $T_{\#}(e^{-V(x)} dx) = e^{-V(x)-q(x)} dx$, T satisfies the mass balance equation

$$\int_{-\infty}^x e^{-V(t)} dt = \int_{-\infty}^{T(x)} e^{-V(t)-q(t)} dt, \tag{5.4}$$

which can be also written as

$$\int_x^{\infty} e^{-V(t)} dt = \int_{T(x)}^{\infty} e^{-V(t)-q(t)} dt \tag{5.5}$$

since the measures $e^{-V(x)} dx$ and $e^{-V(x)-q(x)} dx$ have total mass 1. From (5.4), we deduce that T is differentiable. Indeed, both the functions

$$F(x) := \int_{-\infty}^x e^{-V(t)} dt \quad \text{and} \quad G(x) := \int_{-\infty}^x e^{-V(t)-q(t)} dt$$

are differentiable and their derivatives do not vanish. So, $T(x) = G^{-1} \circ F(x)$ is differentiable as well. Thus, differentiating with respect to x and then taking the logarithm shows that

$$\log(T'(x)) = -V(x) + V(T(x)) + q(T(x)) \quad \forall x \in \mathbb{R}.$$

Consequently,

$$\begin{aligned} V(T(x)) - V(x) - \|q^-\|_{L^\infty(\mathbb{R})} \\ \leq \log(T'(x)) \leq V(T(x)) - V(x) + \|q^+\|_{L^\infty(\mathbb{R})}. \end{aligned} \tag{5.6}$$

On the other hand, (5.4) implies that

$$e^{-\|q^+\|_{L^\infty(\mathbb{R})}} \int_{-\infty}^{T(x)} e^{-V(t)} dt \leq \int_{-\infty}^x e^{-V(t)} dt \leq e^{\|q^-\|_{L^\infty(\mathbb{R})}} \int_{-\infty}^{T(x)} e^{-V(t)} dt$$

since $q \in L^\infty(\mathbb{R})$. Taking the logarithm and defining Φ as in (5.1), we see that

$$-\|q^+\|_{L^\infty(\mathbb{R})} \leq \log \Phi(x) - \log \Phi(T(x)) \leq \|q^-\|_{L^\infty(\mathbb{R})}. \tag{5.7}$$

Analogously, from (5.5), we deduce that

$$-\|q^+\|_{L^\infty(\mathbb{R})} \leq \log \Psi(x) - \log \Psi(T(x)) \leq \|q^-\|_{L^\infty(\mathbb{R})}. \quad (5.8)$$

We claim that

$$-\|q^+\|_{L^\infty(\mathbb{R})} \leq V(T(x)) - V(x) \leq \|q^-\|_{L^\infty(\mathbb{R})} \quad \forall x \in \mathbb{R}. \quad (5.9)$$

To prove this claim, let $x_0 \in \mathbb{R}$ be such that $V(x_0) = \inf_{\mathbb{R}} V$ and consider the sets

$$E_1 := \{x : x \leq x_0 \text{ and } T(x) \leq x_0\} \quad \text{and} \quad E_2 := \{x : x \geq x_0 \text{ and } T(x) \geq x_0\}.$$

Applying (5.2) in E_1 yields that

$$0 \leq V(T(x)) - V(x) \leq \log \Phi(x) - \log \Phi(T(x))$$

if $T(x) \leq x \leq x_0$ and

$$\log \Phi(x) - \log \Phi(T(x)) \leq V(T(x)) - V(x) \leq 0$$

whenever $x \leq T(x) \leq x_0$. Therefore, (5.9) holds in E_1 by (5.7). Similarly, applying (5.3) gives us that (5.9) holds in E_2 by (5.8). Now, we consider three cases:

1. If $T(x_0) = x_0$, the monotonicity of T implies that $E_1 \cup E_2 = \mathbb{R}$, and (5.9) holds in all of \mathbb{R} ;
2. If $T(x_0) > x_0$, then $E_1 \cup E_2 \cup E_+ = \mathbb{R}$ where, thanks to the monotonicity of T , we have

$$E_+ = \{x : x \leq x_0 \text{ and } T(x) \geq x_0\} = \left[T^{-1}(x_0), x_0 \right].$$

Since V attains its minimum at x_0 , V is decreasing on $(-\infty, x_0]$ and increasing on $[x_0, \infty)$. Consequently,

$$V(T^{-1}(x_0)) - V(x_0) \leq V(x) - V(T(x)) \leq V(T(x_0)) - V(x_0) \quad \forall x \in E_+.$$

As $T^{-1}(x_0) \in E_1$ and $x_0 \in E_2$, our above analysis shows that (5.9) holds in E_+ ;

3. If $T(x_0) < x_0$, an analogous argument to one used to prove case 2 demonstrates that $E_1 \cup E_2 \cup E_- = \mathbb{R}$ where $E_- = [x_0, T^{-1}(x_0)]$ and proves (5.9) also in E_- .

Therefore, by (5.6) and (5.9), we deduce (1.3). \square

Remark 5.2. From the numerical inequality $|\log(x)| \geq x - 1$, which holds for $x \in [0, e^2]$, we see that if ϕ is the potential associated to T in Theorem 1.2, then provided that $\|q\|_{L^\infty(\mathbb{R})} \leq 1$, there exists a constant $C > 0$ such that

$$\|\phi'' - 1\|_{L^\infty(\mathbb{R})} \leq C \|q\|_{L^\infty(\mathbb{R})}.$$

We now move to the radially symmetric case in n dimensions.

Proof of Theorem 1.3. Let $\bar{V}, \bar{q} : \mathbb{R} \rightarrow \mathbb{R} \cup \{\infty\}$ be two functions such that $\bar{V} = \bar{q} = \infty$ on $(-\infty, 0)$, and $V(x) = \bar{V}(|x|)$ and $q(x) = \bar{q}(|x|)$ for every $x \in \mathbb{R}^n$. Now, consider the function

$$T(x) := \tilde{T}(|x|) \frac{x}{|x|}$$

where $\tilde{T} : \mathbb{R} \rightarrow \mathbb{R}$ is the optimal transport that takes $e^{-\bar{V}(r)}r^{n-1} dr$ to $e^{-\bar{V}(r)-\bar{q}(r)}r^{n-1} dr$.

Set $\mathbb{R}^+ := [0, \infty)$. We first claim that the optimal transport \tilde{T} is Lipschitz and satisfies

$$\left\| \log \tilde{T}' \right\|_{L^\infty(\mathbb{R}^+)} \leq \|q^+\|_{L^\infty(\mathbb{R}^+)} + \|q^-\|_{L^\infty(\mathbb{R}^+)}. \tag{5.10}$$

Indeed, let $\tilde{V} : \mathbb{R} \rightarrow \mathbb{R} \cup \{\infty\}$ be defined by $\tilde{V}(r) = \bar{V}(r) - (n - 1) \log r$ on \mathbb{R}^+ and infinity otherwise, and let $\tilde{q} = \bar{q}$ on \mathbb{R}^+ and zero elsewhere. Observe that \tilde{V} is convex and \tilde{q} is bounded. Hence, applying Proposition 1.2 with $V = \tilde{V}$ and $q = \tilde{q}$ proves (5.10).

We now conclude the proof. Notice that T is continuous. Furthermore, T is an admissible change of variables from $e^{-V(x)} dx$ to $e^{-V(x)-q(x)} dx$. To see this, we show that for every bounded, Borel function $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}$,

$$\int_{\mathbb{R}^n} \varphi(T(x))e^{-V(x)} dx = \int_{\mathbb{R}^n} \varphi(x)e^{-V(x)-q(x)} dx. \tag{5.11}$$

The formula (5.11) can be rewritten, using polar coordinates and the definition of T , as

$$\begin{aligned} & \int_0^\infty \int_{\mathbb{S}^{n-1}} \varphi(\tilde{T}(r)\alpha) d\mathcal{H}^{n-1}(\alpha) e^{-\bar{V}(r)}r^{n-1} dr \\ &= \int_0^\infty \int_{\mathbb{S}^{n-1}} \varphi(r\alpha) d\mathcal{H}^{n-1}(\alpha) e^{-\bar{V}(r)-\bar{q}(r)}r^{n-1} dr, \end{aligned}$$

which is, in turn, satisfied if we use the test function $\bar{\varphi}(r) = \int_{\mathbb{S}^{n-1}} \varphi(r\alpha) d\mathcal{H}^{n-1}(\alpha)$ and recall the definition of T .

Now, let $\xi \in \mathbb{S}^{n-1}$ and $x \in \mathbb{R}^n \setminus \{0\}$. Since $\tilde{T}(0) = 0$, we observe that

$$\begin{aligned} \nabla T(x)[\xi] &= [\xi|x|^{-1} - x|x|^{-3}\langle x, \xi \rangle] \tilde{T}'(|x|) + x|x|^{-2} \tilde{T}'(|x|)\langle x, \xi \rangle \\ &= [\xi - x|x|^{-2}\langle x, \xi \rangle] \tilde{T}'(t) + x|x|^{-2} \tilde{T}'(|x|)\langle x, \xi \rangle, \end{aligned}$$

where $t \in (0, |x|)$. By (5.10), we deduce that

$$e^{-\|q^+\|_{L^\infty(\mathbb{R}^n)} - \|q^-\|_{L^\infty(\mathbb{R}^n)}} \leq \langle \xi, \nabla T(x)[\xi] \rangle \leq e^{\|q^+\|_{L^\infty(\mathbb{R}^n)} + \|q^-\|_{L^\infty(\mathbb{R}^n)}},$$

which proves (1.4). To conclude, we show that T is the optimal transport taking $e^{-V(x)} dx$ to $e^{-V(x)-q(x)} dx$. Let $\tilde{\phi} : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ be the convex potential associated to \tilde{T} . By construction, $T(x) = \nabla(\tilde{\phi}(|x|))$ and $\tilde{\phi}(|x|)$ is a convex function. Since optimal transports are characterized by being gradients of convex functions, T is the optimal transport taking $e^{-V(x)} dx$ to $e^{-V(x)-q(x)} dx$. \square

6. Appendix

We now show that the linear bound in Remark 5.2 is specific to the 1-dimensional case.

Proposition 6.1. *Let $n \in \mathbb{N}$ and $V(x) = |x|^2/2 + (n/2) \log(2\pi)$, so that e^{-V} is the standard Gaussian density in \mathbb{R}^n . Then, for every $C > 0$, there exists a bounded, continuous perturbation p such that $\|p\|_{L^\infty(\mathbb{R}^n)} \leq 1$ and $e^{-V(x)-p(x)} dx \in \mathcal{P}(\mathbb{R}^n)$ and the optimal transport $T = \nabla\phi$ that takes $e^{-V(x)} dx$ to $e^{-V(x)-p(x)} dx$ satisfies*

$$\|D^2\phi - \text{Id}\|_{L^\infty(\mathbb{R}^n)} > C\|p\|_{L^\infty(\mathbb{R}^n)}.$$

Proof. Suppose, to the contrary, that for every bounded, continuous function $p : \mathbb{R}^n \rightarrow \mathbb{R}$ with $\|p\|_{L^\infty(\mathbb{R}^n)} \leq 1$, the optimal transport $T = \nabla\phi$ that takes $e^{-V(x)} dx$ to $e^{-V(x)-p(x)} dx$ satisfies

$$\|D^2\phi - \text{Id}\|_{L^\infty(\mathbb{R}^n)} \leq C_0\|p\|_{L^\infty(\mathbb{R}^n)} \tag{6.1}$$

for some $C_0 > 0$. In particular, let $q \in L^\infty(\mathbb{R}^n) \cap C^0(\mathbb{R}^n)$, and for all $\varepsilon \geq 0$, define c_ε by

$$e^{c_\varepsilon} = \int_{\mathbb{R}^n} e^{-V(x)-\varepsilon q(x)} dx.$$

By construction, $e^{-V(x)-\varepsilon q(x)-c_\varepsilon} dx \in \mathcal{P}(\mathbb{R}^n)$. Thus, let ϕ_ε be the potential associated to the optimal transport that takes $e^{-V(x)} dx$ to $e^{-V(x)-\varepsilon q(x)-c_\varepsilon} dx$, and remember that ϕ_ε solves the Monge-Ampère equation

$$\det D^2\phi_\varepsilon = e^{-V+V(\nabla\phi_\varepsilon)+\varepsilon q(\nabla\phi_\varepsilon)+c_\varepsilon}. \tag{6.2}$$

Note that $c_\varepsilon \rightarrow 0$ as $\varepsilon \rightarrow 0$. Also, since

$$|c'_\varepsilon| = \left| \frac{(e^{c_\varepsilon})'}{e^{c_\varepsilon}} \right| = \left| \int_{\mathbb{R}^n} -q(x)e^{-V(x)-\varepsilon q(x)-c_\varepsilon} dx \right| \leq \|q\|_{L^\infty(\mathbb{R}^n)},$$

c_ε is Lipschitz as a function of ε and

$$\frac{|c_\varepsilon|}{\varepsilon} \leq \|q\|_{L^\infty(\mathbb{R}^n)}. \tag{6.3}$$

In addition, by the dominated convergence theorem,

$$c'_\varepsilon \rightarrow \iota_q := \int_{\mathbb{R}^n} -q(x)e^{-V(x)} dx \quad \text{as } \varepsilon \rightarrow 0. \tag{6.4}$$

Without loss of generality, we assume that $\phi_\varepsilon(0) = 0$. Now, define

$$\psi_\varepsilon(x) := \frac{\phi_\varepsilon(x) - |x|^2/2}{\varepsilon}.$$

By (6.1) applied to $p = \varepsilon q + c_\varepsilon$ and (6.3), we see that if $\varepsilon \leq \frac{1}{2\|q\|_{L^\infty(\mathbb{R}^n)}}$, then

$$\|D^2\psi_\varepsilon\|_{L^\infty(\mathbb{R}^n)} \leq (C_0 + 1)\|q\|_{L^\infty(\mathbb{R}^n)}. \tag{6.5}$$

Recall that, for any $n \times n$ matrix A , there exists a $K > 0$, depending only on $\|A\|$, such that for all ε sufficiently small $|\log \det(\text{Id} + \varepsilon A) - \varepsilon \text{tr } A| \leq \varepsilon^2 K$. Therefore, there exist an $\varepsilon_0 > 0$ and a collection of functions g_ε with

$$\sup_{\varepsilon \leq \varepsilon_0} \|g_\varepsilon\|_{L^\infty(\mathbb{R}^n)} < \infty \tag{6.6}$$

such that for all $\varepsilon \leq \varepsilon_0$,

$$\varepsilon \Delta \psi_\varepsilon(x) + \varepsilon^2 g_\varepsilon(x) = \log \det (\text{Id} + \varepsilon D^2 \psi_\varepsilon) = \log \det D^2 \phi_\varepsilon.$$

Thus, by (6.2) and our choice of V ,

$$\begin{aligned} \Delta \psi_\varepsilon(x) + \varepsilon g_\varepsilon(x) &= \frac{V(\nabla \phi_\varepsilon(x)) - V(x) + \varepsilon q(\nabla \phi_\varepsilon(x)) + c_\varepsilon}{\varepsilon} \\ &= \int_0^1 \langle (1-t)\nabla \phi_\varepsilon(x) + tx, \nabla \psi_\varepsilon(x) \rangle dt + q(\nabla \phi_\varepsilon(x)) + \frac{c_\varepsilon}{\varepsilon} \\ &= \langle x, \nabla \psi_\varepsilon(x) \rangle + \frac{\varepsilon}{2} |\nabla \psi_\varepsilon(x)|^2 + q(\nabla \phi_\varepsilon(x)) + \frac{c_\varepsilon}{\varepsilon}. \end{aligned} \tag{6.7}$$

We claim that, up to a subsequence, there exists a function $\psi_0 \in C^{1,1}_{\text{loc}}(\mathbb{R}^n)$ such that $\psi_\varepsilon \rightarrow \psi_0$ in $C^1_{\text{loc}}(\mathbb{R}^n)$ and $D^2\psi_\varepsilon \rightharpoonup D^2\psi_0$ weakly- $*$ in $L^\infty(\mathbb{R}^n)$ as $\varepsilon \rightarrow 0$. To this end, by Arzelà-Ascoli, it suffices to show that ψ_ε are locally bounded in $C^{1,1}$. Since $\psi_\varepsilon(0) = 0$, by (6.5), it is enough to prove that

$$\liminf_{\varepsilon \rightarrow 0} |\nabla \psi_\varepsilon(0)| < \infty. \tag{6.8}$$

Assume, to the contrary, that $\lim_{\varepsilon \rightarrow 0} |\nabla \psi_\varepsilon(0)| = \infty$. Notice that (6.7) implies that for all $\varepsilon \leq \varepsilon_0$ and $x \in \mathbb{R}^n$,

$$\begin{aligned} & \left| \int_0^1 \langle (1-t)\nabla\phi_\varepsilon(x) + tx, \nabla\psi_\varepsilon(0) \rangle dt \right| \\ & \leq \left| \int_0^1 \langle (1-t)\nabla\phi_\varepsilon(x) + tx, \nabla\psi_\varepsilon(x) \rangle dt \right| + (|\nabla\phi_\varepsilon(x)| + |x|)|\nabla\psi_\varepsilon(0) - \nabla\psi_\varepsilon(x)| \\ & \leq |\Delta\psi_\varepsilon(x)| + \varepsilon|g_\varepsilon(x)| + |q(\nabla\phi_\varepsilon(x))| + \frac{|c_\varepsilon|}{\varepsilon} \\ & \quad + (|\nabla\phi_\varepsilon(x)| + |x|)|x| \sup_{\varepsilon \leq \varepsilon_0} \|D^2\psi_\varepsilon\|_{L^\infty(\mathbb{R}^n)}. \end{aligned}$$

Let $\alpha_\varepsilon = \nabla\psi_\varepsilon(0)/|\nabla\psi_\varepsilon(0)| \in \mathbb{S}^{n-1}$, and note that up to subsequences $\alpha_\varepsilon \rightarrow \alpha_0 \in \mathbb{S}^{n-1}$ as $\varepsilon \rightarrow 0$. Furthermore, let $\eta \in C_c^\infty(B_{1/2}(\alpha_0))$ be a nonnegative function that integrates to one. Then, by (6.3), we deduce that

$$\begin{aligned} & |\nabla\psi_\varepsilon(0)| \int_{\mathbb{R}^n} \left| \int_0^1 \langle (1-t)\nabla\phi_\varepsilon(x) + tx, \alpha_\varepsilon \rangle dt \right| \eta(x) dx \\ & = \int_{\mathbb{R}^n} \left| \int_0^1 \langle (1-t)\nabla\phi_\varepsilon(x) + tx, \nabla\psi_\varepsilon(0) \rangle dt \right| \eta(x) dx \tag{6.9} \\ & \leq \sup_{\varepsilon \leq \varepsilon_0} \varepsilon \|g_\varepsilon\|_{L^\infty(\mathbb{R}^n)} + 2\|q\|_{L^\infty(\mathbb{R}^n)} \\ & \quad + \sup_{\varepsilon \leq \varepsilon_0} \|D^2\psi_\varepsilon\|_{L^\infty(\mathbb{R}^n)} \int_{\mathbb{R}^n} (|\nabla\phi_\varepsilon(x)||x| + |x|^2 + 1)\eta(x) dx. \end{aligned}$$

Recall that $D^2\phi_\varepsilon$ converges uniformly to the identity matrix by (6.1) applied to ϕ_ε and εq . By the stability and uniqueness of optimal transports, $\nabla\phi_\varepsilon$ converges locally uniformly to the identity map as $\varepsilon \rightarrow 0$. In particular, $|\nabla\phi_\varepsilon(x)| \leq 2$ for every $x \in B_{1/2}(\alpha_0)$ and ε sufficiently small, and we obtain that

$$\lim_{\varepsilon \rightarrow 0} \int_{\mathbb{R}^n} \left| \int_0^1 \langle (1-t)\nabla\phi_\varepsilon(x) + tx, \alpha_\varepsilon \rangle dt \right| \eta(x) dx = \int_{\mathbb{R}^n} \langle x, \alpha_0 \rangle \eta(x) dx \geq \frac{1}{2}$$

by dominated convergence. Thus, taking the limit in (6.9) and noticing that the right-hand side is bounded as $\varepsilon \rightarrow 0$ thanks to (6.5) and (6.6), we see that

$$\infty = \lim_{\varepsilon \rightarrow 0} |\nabla\psi_\varepsilon(0)| \left| \int_0^1 \langle (1-t)\nabla\phi_\varepsilon(x) + tx, \alpha_\varepsilon \rangle dt \right| < \infty,$$

which, being impossible, proves (6.8) and shows that $\psi_\varepsilon \rightarrow \psi_0$ in $C_{loc}^1(\mathbb{R}^n)$ and $D^2\psi_\varepsilon \rightharpoonup D^2\psi_0$ weakly-* in $L^\infty(\mathbb{R}^n)$ as $\varepsilon \rightarrow 0$ for some function $\psi_0 \in C_{loc}^{1,1}(\mathbb{R}^n)$.

Now, reformulating (6.7), we see that for any $\eta \in C_c^\infty(\mathbb{R}^n)$,

$$\begin{aligned} & \int_{\mathbb{R}^n} \left(\Delta \psi_\varepsilon(x) + \varepsilon g_\varepsilon(x) - q(\nabla \phi_\varepsilon(x)) - \frac{c_\varepsilon}{\varepsilon} \right) \eta(x) dx \\ &= \int_{\mathbb{R}^n} \left(\langle x, \nabla \psi_\varepsilon(x) \rangle + \frac{\varepsilon}{2} |\nabla \psi_\varepsilon(x)|^2 \right) \eta(x) dx. \end{aligned} \tag{6.10}$$

Thus, recalling (6.4) and that q is continuous, we can pass to the limit and obtain that

$$\int_{\mathbb{R}^n} (\Delta \psi_0(x) - \langle x, \nabla \psi_0(x) \rangle) \eta(x) dx = \int_{\mathbb{R}^n} (q(x) + \iota_q) \eta(x) dx$$

for all $\eta \in C_c^\infty(\mathbb{R}^n)$. Since q was arbitrary, we have shown that for every $q \in L^\infty(\mathbb{R}^n) \cap C^0(\mathbb{R}^n)$, there exists a function $\psi_0 \in C_{loc}^{1,1}(\mathbb{R}^n)$ solution to

$$\Delta \psi_0(x) - \langle x, \nabla \psi_0(x) \rangle = q(x) + \iota_q. \tag{6.11}$$

We now show that this is impossible. Recall that there exists a bounded, continuous g and $\psi \in C_{loc}^{1,\alpha}(B_2) \cap C^\infty(B_2 \setminus \{0\})$, for any $\alpha \in (0, 1)$, such that $\Delta \psi(x) = g(x)$ in B_2 , yet $\psi \notin C^{1,1}(B_2)$. In particular, $\lim_{x \rightarrow 0} |D^2 \psi(x)| = \infty$. (See [10, Chapter 3].) Define

$$h(x) := \begin{cases} g(x) - \langle x, \nabla \psi(x) \rangle & x \in B_1 \\ g(x/|x|) - \langle x/|x|, \nabla \psi(x/|x|) \rangle & x \in \mathbb{R}^n \setminus B_1, \end{cases}$$

and observe that, since $\psi \in C_{loc}^{1,\alpha}(B_2)$ and g is bounded and continuous, $h \in L^\infty(\mathbb{R}^n) \cap C^0(\mathbb{R}^n)$. By construction, there exists a $\psi_0 \in C_{loc}^{1,1}(\mathbb{R}^n)$ that solves (6.11) with $q = h$. Then, for $\psi_1 := \psi_0 - \psi$ we have that $\Delta \psi_1(x) - \langle x, \nabla \psi_1(x) \rangle = \iota_h$ in B_1 . Thus, $\psi_1 \in C^\infty(B_1)$ by elliptic regularity, a contradiction since $\psi \notin C_{loc}^{1,1}(B_1)$ and $\psi_0 \in C^{1,1}(B_1)$. □

References

- [1] S. ALESKER, S. DAR and V. MILMAN, *A remarkable measure preserving diffeomorphism between two convex bodies in \mathbb{R}^n* , *Geom. Dedicata* **74** (1999), 201–212.
- [2] Y. BRENIER, *Polar factorization and monotone rearrangement of vector-valued functions*, *Comm. Pure Appl. Math.* **44** (1991), 365–417.
- [3] L. A. CAFFARELLI, *The regularity of mappings with a convex potential*, *J. Amer. Math. Soc.* **5** (1992), 99–104.
- [4] L. A. CAFFARELLI, *Monotonicity properties of optimal transportation and the FKG and related inequalities*, *Comm. Math. Phys.* **214** (2000), 547–563.
- [5] L. A. CAFFARELLI, *Erratum: Monotonicity properties of optimal transportation and the FKG and related inequalities*, *Comm. Math. Phys.* **225** (2002), 449–450.
- [6] D. CORDERO-ERAUSQUIN, *Some applications of mass transport to Gaussian-type inequalities*, *Arch. Ration. Mech. Anal.* **161** (2002), 257–269.

- [7] D. CORDERO-ERAUSQUIN, M. FRADELIZI and B. MAUREY, *The (B) conjecture for the Gaussian measure of dilates of symmetric convex sets and related problems*, J. Funct. Anal. **214** (2004), 410–427.
- [8] G. DE PHILIPPIS and A. FIGALLI, *The Monge-Ampère equation and its link to optimal transportation*, Bull. Amer. Math. Soc. (N.S.) **51** (2014), 527–580.
- [9] A. FIGALLI, “The Monge-Ampère Equation and its Applications”, Zürich Lectures in Advanced Mathematics, European Mathematical Society (EMS), Zürich, 2017.
- [10] Q. HAN and F. LIN, “Elliptic Partial Differential Equations”, 2nd Edition, Courant Lecture Notes in Mathematics, Vol. 1, New York University, Courant Institute of Mathematical Sciences, New York, American Mathematical Society, Providence, 1997.
- [11] G. HARGÉ, *A convex/log-concave correlation inequality for Gaussian measure and an application to abstract Wiener spaces*, Probab. Theory Related Fields **130** (2004), 415–440.
- [12] Y.-H. KIM and E. MILMAN, *A generalization of Caffarelli’s contraction theorem via (reverse) heat flow*, Math. Ann. **354** (2012), 827–862.
- [13] B. KLARTAG, *Marginals of geometric inequalities*, In: “Geometric Aspects of Functional Analysis”, V. D. Milman and G. Schechtman (eds.), Lecture Notes in Mathematics, Vol. 1910, Springer Berlin, 2007, 133–166.
- [14] A. V. KOLESNIKOV, *Global Hölder estimates for optimal transportation*, (Russian) Mat. Zametki **88** (2010), 708–728; (English) Math. Notes **88** (2010), 678–695.
- [15] A. V. KOLESNIKOV, *On Sobolev regularity of mass transport and transportation inequalities*, Theory Probab. Appl. **57** (2013), 243–264.
- [16] A. V. POGORELOV, *The regularity of generalized solutions of the equation $\det(\partial^2 u / \partial x_i \partial x_j) = \varphi(x_1, x_2, \dots, x_n) > 0$* (Russian), Dokl. Akad. Nauk SSSR **200** (1971), 534–537.
- [17] S. I. VALDIMARSSON, *On the Hessian of the optimal transport potential*, Ann. Sc. Norm. Super. Pisa Cl. Sci. (5) **6** (2007), 441–456.
- [18] C. VILLANI, “Optimal Transport, Old and New”, Grundlehren der mathematischen Wissenschaften, Vol. 338, Springer-Verlag Berlin Heidelberg, 2009.

Institute for Theoretical Studies
ETH Zürich, Clausiusstrasse 47
CH-8092 Zürich, Switzerland
maria.colombo@eth-its.ethz.ch

The University of Texas at Austin
Mathematics Dept. RLM 8.100
2515 Speedway Stop C1200
Austin, TX 78712, USA
figalli@math.utexas.edu
yjhaveri@math.utexas.edu