RIGIDITY AND STABILITY OF CAFFARELLI'S LOG-CONCAVE PERTURBATION THEOREM

GUIDO DE PHILIPPIS AND ALESSIO FIGALLI

To Nicola Fusco, for his 60th birthday, con affetto e ammirazione.

Abstract. In this note we establish some rigidity and stability results for Caffarelli's log-concave perturbation theorem. As an application we show that if a 1-log-concave measure has almost the same Poincaré constant as the Gaussian measure, then it almost splits off a Gaussian factor.

1. Introduction

Let γ_n denote the centered Gaussian measure in $\mathbb{R}^n$, i.e., $\gamma_n = (2\pi)^{-n/2}e^{-|x|^2/2}dx$, and let μ be a probability measure on $\mathbb{R}^n$. By a classical theorem of Brenier [2], there exists a convex function $\varphi : \mathbb{R}^n \to \mathbb{R}$ such that $T = \nabla\varphi : \mathbb{R}^n \to \mathbb{R}^n$ transports γ_n onto μ, i.e., $T_\sharp\gamma_n = \mu$, or equivalently

$$\int h \circ T \, d\gamma_n = \int h \, d\mu \qquad \text{for all continuous and bounded functions } h \in C_b(\mathbb{R}^n).$$

In the sequel we will refer to T as the *Brenier map* from γ_n to μ.

In [4, 5] Caffarelli proved that if μ is "more log-concave" than γ_n, then T is 1-Lipschitz, that is, all the eigenvalues of $D^2\varphi$ are bounded from above by 1. Here is the exact statement:

Theorem 1.1 (Caffarelli)**.** *Let γ_n be the Gaussian measure in $\mathbb{R}^n$, and let $\mu = e^{-V}dx$ be a probability measure satisfying $D^2V \geq \mathrm{Id}_n$. Consider the Brenier map $T = \nabla\varphi$ from γ_n to μ. Then T is 1-Lipschitz. Equivalently, $0 \leq D^2\varphi(x) \leq \mathrm{Id}_n$ for a.e. x.*

This theorem allows one to show that optimal constants in several functional inequalities are extremized by the Gaussian measure. More precisely, let F, G, H, L, J be continuous functions on $\mathbb{R}$ and assume that F, G, H, J are nonnegative, and that H and J are increasing. For $\ell \in \mathbb{R}_+$ let

$$\lambda(\mu,\ell) := \inf\left\{ \frac{H\Big(\int J(|\nabla u|)\,d\mu\Big)}{F\Big(\int G(u)\,d\mu\Big)} \; : \qquad u \in \mathrm{Lip}(\mathbb{R}^n)\,, \int L(u)\,d\mu = \ell \right\}. \tag{1.1}$$

Then

$$\lambda(\gamma_n,\ell) \leq \lambda(\mu,\ell). \tag{1.2}$$

Indeed, given a function u admissible in the variational formulation for μ, we set $v := u \circ T$ and note that, since $T_\sharp\gamma_n = \mu$,

$$\int K(v)\,d\gamma_n = \int K(u\circ T)\,d\gamma_n = \int K(u)\,d\mu \qquad \text{for } K = G, L.$$

In particular, this implies that v is admissible in the variational formulation for γ_n. Also, thanks to Caffarelli's Theorem,

$$|\nabla v| \leq |\nabla u|\circ T\,|\nabla T| \leq |\nabla u|\circ T,$$

therefore

$$H\Big(\int J(|\nabla v|)\,d\gamma_n\Big) \le H\Big(\int J(|\nabla u|)\circ T\,d\gamma_n\Big) = H\Big(\int J(|\nabla u|)\,d\mu\Big).$$

Thanks to these formulas, (1.2) follows easily.

Note that the classical Poincaré and Log-Sobolev inequalities fall in the above general framework. For instance, choosing $H(t) = F(t) = L(t) = t$, $\ell = 0$, and $J(t) = F(t) = |t|^p$ with $p \ge 1$, we deduce that

$$\inf\left\{\frac{\int|\nabla u|^p\,d\mu}{\int|u|^p\,d\mu}\ :\ u\in \mathrm{Lip}(\mathbb{R}^n)\,,\int u\,d\mu = 0\right\} \ge \inf\left\{\frac{\int|\nabla u|^p\,d\gamma_n}{\int|u|^p\,d\gamma_n}\ :\ u\in \mathrm{Lip}(\mathbb{R}^n)\,,\int u\,d\gamma_n = 0\right\}. \tag{1.3}$$

Two questions that naturally arise from the above considerations are:

- *Rigidity*: What can be said about μ when $\lambda(\mu,\ell) = \lambda(\gamma_n,\ell)$?
- *Stability*: What can be said about μ when $\lambda(\mu,\ell) \approx \lambda(\gamma_n,\ell)$?

Looking at the above proof, these two questions can usually be reduced to the study of the corresponding ones concerning the optimal map T in Theorem 1.1 (here $|A|$ denotes the operator norm of a matrix A):

- *Rigidity*: What can be said about μ when $|\nabla T(x)| = 1$ for a.e. x ?
- *Stability*: What can be said about μ when $|\nabla T(x)| \approx 1$ (in suitable sense)?

Our first main result state that if $|\nabla T(x)| = 1$ for a.e. x then μ "splits off" a Gaussian factor. More precisely, it splits off as many Gaussian factors as the number of eigenvalues of $\nabla T = D^2\varphi$ that are equal to 1. In the following statement and in the sequel, given $p\in\mathbb{R}^k$ we denote by $\gamma_{p,k}$ the Gaussian measure in $\mathbb{R}^k$ with barycenter p, that is, $\gamma_{p,k} = (2\pi)^{-k/2}e^{-|x-p|^2/2}dx$.

Theorem 1.2 (Rigidity)**.** *Let γ_n be the Gaussian measure in $\mathbb{R}^n$, and let $\mu = e^{-V}dx$ be a probability measure with $D^2V \ge \mathrm{Id}_n$. Consider the Brenier map $T = \nabla\varphi$ from γ_n to μ, and let*

$$0 \le \lambda_1(D^2\varphi(x)) \le \dots \le \lambda_n(D^2\varphi(x)) \le 1$$

be the eigenvalues of the matrix $D^2\varphi(x)$. If $\lambda_{n-k+1}(D^2\varphi(x)) = 1$ for a.e. x then $\mu = \gamma_{p,k}\otimes e^{-W(x')}dx'$, where $W : \mathbb{R}^{n-k}\to\mathbb{R}$ satisfies $D^2W \ge \mathrm{Id}_{n-k}$.

Our second main result is a quantitative version of the above theorem. Before stating it let us recall that, given two probability measures $\mu,\nu\in\mathcal{P}(\mathbb{R}^n)$, the 1-Wasserstein distance between them is defined as

$$W_1(\mu,\nu) := \inf\Big\{\int|x-y|\,d\sigma(x,y)\ :\quad \sigma\in\mathcal{P}(\mathbb{R}^n\times\mathbb{R}^n)\text{ such that }(\mathrm{pr}_1)_\sharp\sigma = \mu,\ (\mathrm{pr}_2)_\sharp\sigma = \nu\Big\},$$

where pr_1 (resp. pr_2) is the projection of $\mathbb{R}^n\times\mathbb{R}^n$ onto the first (resp. second) factor. Our stability result is formulated in terms of the W_1-distance between probability measure as this distance natural comes out from our strategy of proof. Our result could also be proved with other notions of distances meterizing the weak topology (for instance, any Wasserstein distance W_p), as well as stronger notion of distances (such as the total variation), but we shall not investigate this here.

Theorem 1.3 (Stability). *Let γ_n be the Gaussian measure in $\mathbb{R}^n$ and let $\mu = e^{-V}dx$ be a probability measure with $D^2V \geq \mathrm{Id}_n$. Consider the Brenier map $T = \nabla\varphi$ from γ_n to μ, and let*

$$0 \leq \lambda_1(D^2\varphi(x)) \leq \cdots \leq \lambda_n(D^2\varphi(x)) \leq 1$$

be the eigenvalues of $D^2\varphi(x)$. Let $\varepsilon \in (0,1)$ and assume that

$$1-\varepsilon \leq \int \lambda_{n-k+1}(D^2\varphi(x))\, d\gamma_n(x) \leq 1\,. \tag{1.4}$$

Then there exists a probability measure $\nu = \gamma_{p,k} \otimes e^{-W(x')}dx'$, with $W : \mathbb{R}^{n-k} \to \mathbb{R}$ satisying $D^2W \geq \mathrm{Id}_{n-k}$, such that

$$W_1(\mu,\nu) \lesssim \frac{1}{|\log\varepsilon|^{1/4-}}. \tag{1.5}$$

In the above statement, and in the rest of the note, we are employing the following notation:

$$X \lesssim Y^{\beta-} \qquad \text{if } X \leq C(n,\alpha)Y^\alpha \text{ for all } \alpha < \beta.$$

Analogously,

$$X \gtrsim Y^{\beta-} \qquad \text{if } C(n,\alpha)X \geq Y^\alpha \text{ for all } \alpha < \beta.$$

Remark 1.4. We do not expect the stability estimate in the previous theorem to be sharp. In particular, in dimension 1 an elementary argument (but completely specific to the one dimensional case) gives a linear control in ε. Indeed, assuming (up to translating μ) that

$$\int x\, d\mu = 0, \tag{1.6}$$

set $\psi(x) := x^2/2 - \varphi(x)$. Then, since $\psi'' = (x-T)' > 0$, our assumption can be rewritten as

$$\int |(x-T)'|\, d\gamma_1 = \int \psi''\, d\gamma_1 \leq \varepsilon.$$

Also, since $T_\#\gamma_1 = \mu$, (1.6) yields

$$\int T(x)\, d\gamma_1 = 0 = \int x\, d\gamma_1.$$

Hence, by the L^1-Poincaré inequality for the Gaussian measure we obtain

$$W_1(\mu,\gamma_1) \leq \int |x-y|\, d\sigma_T(x,y) = \int |x-T(x)|\, d\gamma_1 \leq C \int |(x-T)'|\, d\gamma_1 \leq C\,\varepsilon,$$

where $\sigma_T := (\mathrm{Id} \times T)_\#\gamma_1$.

As explained above, Theorems 1.2 and 1.3 can be applied to study the structure of 1-log-concave measures (i.e., measures of the form $e^{-V}dx$ with $D^2V \geq \mathrm{Id}_n$) that almost achieve equality in (1.2). To simplify the presentation and emphasize the main ideas, we limit ourselves to a particular instance of (1.1), namely the optimal constant in the L^2-Poincaré inequality for μ:

$$\lambda_\mu := \inf\left\{ \frac{\int |\nabla u|^2\, d\mu}{\int u^2\, d\mu} : \qquad u \in \mathrm{Lip}(\mathbb{R}^n)\,, \int u\, d\mu = 0 \right\}.$$

It is well-known that $\lambda_{\gamma_n} = 1$ and that $\{u_i(x) = x_i\}_{1\le i\le n}$ are the corresponding minimizers. In particular it follows by (1.3) that, for every 1-log-concave measure μ,

$$\int u^2\,d\mu \le \int |\nabla u|^2\,d\mu \qquad \text{for all } u \in \mathrm{Lip}(\mathbb{R}^n) \text{ with } \int u\,d\mu = 0. \tag{1.7}$$

As a consequence of Theorems 1.2 and 1.3 we have:

Theorem 1.5. *Let $\mu = e^{-V}dx$ be a probability measure with $D^2V \ge \mathrm{Id}_n$, and assume there exist k functions $\{u_i\}_{1\le i\le k} \subset W^{1,2}(\mathbb{R}^n,\mu)$, $k \le n$, such that*

$$\int u_i\,d\mu = 0, \qquad \int u_i^2\,d\mu = 1, \qquad \int \nabla u_i \cdot \nabla u_j\,d\mu = 0 \qquad \forall\, i \ne j,$$

and

$$\int |\nabla u_i|^2\,d\mu \le 1+\varepsilon$$

for some $\varepsilon > 0$. Then there exists a probability measure $\nu = \gamma_{p,k} \otimes e^{-W(x')}dx'$, with $W : \mathbb{R}^{n-k} \to \mathbb{R}$ satisfying $D^2W \ge \mathrm{Id}_{n-k}$, such that

$$W_1(\mu,\nu) \lesssim \frac{1}{|\log\varepsilon|^{1/4-}}.$$

In particular, if there exist n orthogonal functions $\{u_i\}_{1\le i\le n}$ that attain the equality in (1.7) then $\mu = \gamma_{n,p}$.

We conclude this introduction recalling that the rigidity version of the above theorem (i.e., the case $\varepsilon = 0$) has already been proved by Cheng and Zho in [6, Theorem 2] with completely different techniques.

2. Proof of Theorem 1.2

To prove Theorem 1.2, we first recall the following classical estimate due to Alexandrov (see for instance [8, Theorem 2.2.4 and Example 2.1.2(1)] for a proof):

Lemma 2.1. *Let Ω be an open bounded convex set, and let $u : \Omega \to \mathbb{R}$ be a $C^{1,1}$ convex function such that $u = 0$ on $\partial\Omega$. Then there exists a dimensional constant $C_n > 0$ such that*

$$|u(x)|^n \le C_n\,\mathrm{diam}(\Omega)^{n-1}\,\mathrm{dist}(x,\partial\Omega)\int_\Omega \det D^2u \qquad \forall\, x \in \Omega.$$

Proof of Theorem 1.2. Set $\psi(x) := |x|^2/2 - \varphi(x)$ and note that, as a consequence of Theorem 1.1, $\psi : \mathbb{R}^n \to \mathbb{R}$ is a $C^{1,1}$ convex function with $0 \le D^2\psi \le \mathrm{Id}$. Also, our assumption implies that

$$\lambda_1(D^2\psi(x)) = \ldots = \lambda_k(D^2\psi(x)) = 0 \qquad \text{for a.e. } x \in \mathbb{R}^d. \tag{2.1}$$

We are going to show that ψ depends only on $n-k$ variables. As we shall show later, this will immediately imply the desired conclusion. In order to prove the above claim, we note it is enough to prove it for $k = 1$, since then one can argue recursively on $\mathbb{R}^{n-1}$ and so on.

Note that (2.1) implies that

$$\det D^2\psi \equiv 0. \tag{2.2}$$

Up to translate μ we can subtract a linear function to ψ and assume without loss of generality that $\psi(x) \ge \psi(0) = 0$.

Consider the convex set $\Sigma := \{\psi = 0\}$. We claim that Σ contains a line. Indeed, if not, this set would contain an exposed point $\bar{x}$. Up to a rotation, we can assume that $\bar{x} = a\, e_1$ with $a \geq 0$. Also, since $\bar{x}$ is an exposed point,

$$\Sigma \subset \{x_1 \leq a\} \quad \text{and} \quad \Sigma \cap \{x_1 = a\} = \{\bar{x}\}.$$

Hence, by convexity of Σ, the set $\Sigma \cap \{x_1 \geq -1\}$ is compact.

Consider the affine function

$$\ell_\eta(x) := \eta(x_1 + 1), \qquad \eta > 0 \text{ small},$$

and define $\Sigma_\eta := \{\psi \leq \ell_\eta\}$. Note that, as $\eta \to 0$, the sets Σ_η converge in the Hausdorff distance to the compact set $\Sigma \cap \{x_1 \geq -1\}$. In particular, this implies that Σ_η is bounded for η sufficiently small.

We now apply Lemma 2.1 to the convex function $\psi - \ell_\eta$ inside Σ_η, and it follows by (2.2) that (note that $D^2\ell_\eta \equiv 0$)

$$|\psi(x) - \ell_\eta(x)|^n \leq C_n\big(\operatorname{diam}(\Sigma_\eta)\big)^n \int_{\Sigma_\eta} \det D^2\psi = 0 \qquad \forall\, x \in \Sigma_\eta.$$

In particular this implies that $\psi(0) = \ell_\eta(0) = \eta$, a contradiction to the fact that $\psi(0) = 0$.

Hence, we proved that $\{\psi = 0\}$ contains a line, say $\mathbb{R}e_1$. Consider now a point $x \in \mathbb{R}^n$. Then, by convexity of ψ,

$$\psi(x) + \nabla\psi(x) \cdot (se_1 - x) \leq \psi(se_1) = 0 \qquad \forall\, s \in \mathbb{R},$$

and by letting $s \to \pm\infty$ we deduce that $\partial_1\psi(x) = \nabla\psi(x) \cdot e_1 = 0$. Since x was arbitrary, this means that $\partial_1\psi \equiv 0$, hence $\psi(x) = \psi(0, x')$, $x' \in \mathbb{R}^{n-1}$.

Going back to φ, this proves that

$$T(x) = (x_1, x' - \nabla\psi(x')),$$

and because $\mu = T_\# \gamma_n$ we immediately deduce that $\mu = \gamma_1 \otimes \mu_1$ where $\mu_1 := (\mathrm{Id}_{n-1} - \nabla\psi)_\# \gamma_{n-1}$.

Finally, to deduce that $\mu_1 = e^{-W} dx'$ with $D^2 W \geq \mathrm{Id}_{n-1}$ we observe that $\mu_1 = (\pi')_\# \mu$ where $\pi' : \mathbb{R}^n \to \mathbb{R}^{n-1}$ is the projection given by $\pi'(x_1, x') := x'$. Hence, the result is a consequence of the fact that 1-log-concavity is preserved when taking marginals, see [1, Theorem 4.3] or [9, Theorem 3.8]. $\square$

3. Proof of Theorem 1.3

To prove Theorem 1.3, we first recall a basic properties of convex sets (see for instance [3, Lemma 2] for a proof).

Lemma 3.1. *Given S an open bounded convex set in $\mathbb{R}^n$ with barycenter at 0, let $\mathcal{E}$ denote an ellipsoid of minimal volume with center 0 and containing S. Then there exists a dimensional constant $\kappa_n > 0$ such that $\kappa_n \mathcal{E} \subset S$.*

Thanks to this result, we can prove the following simple geometric lemma:

Lemma 3.2. *Let κ_n be as in Lemma 3.1, set $c_n := \kappa_n/2$, and consider $S \subset \mathbb{R}^n$ an open convex set with barycenter at 0. Assume that $S \subset B_R$ and $\partial S \cap \partial B_R \neq \emptyset$. Then there exists a unit vector $v \in \mathbb{S}^{n-1}$ such that $\pm c_n R v \in S$.*

Proof. By scaling we can assume that $R = 1$.

Let $v \in \partial S \cap \partial B_1$, and consider the ellipsoid $\mathcal{E}$ provided by Lemma 3.1. Since $v \in \overline{\mathcal{E}}$ and $\mathcal{E}$ is symmetric with respect to the origin, also $-v \in \overline{\mathcal{E}}$. Hence

$$\pm c_n v \in c_n \overline{\mathcal{E}} \subset \kappa_n \mathcal{E} \subset S,$$

as desired. □

In order to complete the proof of Theorem 1.3 we recall the following geometric result, see [3, Lemma 1].

Lemma 3.3. *Let $\psi : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be a nonnegative convex function with $\psi(0) = 0$. Assume that ψ is finite in a neighbourhood of 0 and that the graph of ψ does not contains lines. Then there exists $p \in \mathbb{R}^n$ such that the open convex set*

$$S_1 := \{x : \psi(x) \le p \cdot x + 1\}$$

is nonempty, bounded, and with barycenter at 0.

Proof of Theorem 1.3. As in the proof of Theorem 1.2 we set $\psi := |x|^2/2 - \varphi$. Then, inequality (1.4) gives

$$\int \lambda_k(D^2\psi)\, d\gamma_n \le \varepsilon. \tag{3.1}$$

Up to subtract a linear function (i.e., substituting μ with one of its translation, which does not affect the conlclusion of the theorem) we can assume that $\psi(x) \ge \psi(0) = 0$, therefore $\nabla\psi(0) = \nabla\varphi(0) = 0$. Since $(\nabla\varphi)_\# \gamma_n = \mu$ and $\|D^2\varphi\|_\infty \le 1$, these conditions imply that

$$\int |x|\, d\mu(x) = \int |\nabla\varphi(x)|\, d\gamma_n(x) = \int |\nabla\varphi(x) - \nabla\varphi(0)|\, d\gamma_n(x) \le \int |x|\, d\gamma_n(x) \le C_n.$$

In particular

$$W_1(\mu, \gamma) \le W_1(\mu, \delta_0) + W_1(\delta_0, \gamma) \le C_n.$$

This proves that (1.5) holds true with $\nu = \gamma_n$ and with a constant $C \approx |\log \varepsilon_0|^{1/4}$ whenever $\varepsilon \ge \varepsilon_0$. Hence, when showing the validity of (1.5), we can safely assume that $\varepsilon \le \varepsilon_0(n) \ll 1$. Furthermore, we can assume that the graph of ψ does not contain lines (otherwise, by the proof of Theorem 1.2, we would deduce that μ splits a Gaussian factor, and we could simply repeat the argument in $\mathbb{R}^{n-1}$).

Thus we can apply Lemma 3.3 to deduce the existence of a slope $p \in \mathbb{R}^n$ such that

$$S_1 = \{x \in \mathbb{R}^n : \psi(x) < p \cdot x + 1\}$$

is nonempty, bounded, and with barycenter at 0. Applying Lemma 2.1 to the convex function $\tilde\psi(x) := \psi(x) - p \cdot x - 1$ inside the set S_1, we get (note that $D^2\tilde\psi = D^2\psi$)

$$1 \le \Big(-\min_{S_1} \tilde\psi\Big)^n \le C_n \big(\mathrm{diam}(S_1)\big)^n \int_{S_1} \det D^2\psi. \tag{3.2}$$

Consider now the smallest radius $R > 0$ such that $S_1 \subset B_R$ (note that $R < +\infty$ since S_1 is bounded). Since $\gamma_n \ge c_n e^{-R^2/2}$ in B_R and $\lambda_i(D^2\psi) \le 1$ for all $i = 1, \dots, n$, (3.1) implies that

$$\int_{B_R} \det D^2\psi \le C_n e^{R^2/2} \varepsilon.$$

Hence, using (3.2), since $\operatorname{diam}(S_1) \le 2R$ we get

$$1 \le C_n R^n e^{R^2/2} \varepsilon$$

which yields

$$R \gtrsim |\log \varepsilon|^{1/2+}. \tag{3.3}$$

Now, up to a rotation and by Lemma 3.2, we can assume that

$$\pm c_n R e_1 \in S_1.$$

Consider $1 \ll \rho \ll R^{1/2}$ to be chosen. Since $S_1 \subset B_R$ and $\psi \ge 0$ we get that $|p| \le 1/R$, therefore $\psi \le 2$ on $S_1 \subset B_R$. Hence

$$2 \ge \psi(z) \ge \psi(x) + \langle \nabla\psi(x), z - x\rangle \ge \langle \nabla\psi(x), z - x\rangle \qquad \forall\, z \in S_1,\, x \in B_\rho.$$

Thus, since $|\nabla\psi| \le \rho$ in B_ρ (by $\|D^2\psi\|_{L^\infty(\mathbb{R}^n)} \le 1$ and $|\nabla\psi(0)| = 0$), choosing $z = \pm c_n R e_1$ we get

$$|\partial_1 \psi| \le \frac{C_n \rho^2}{R} \qquad \text{inside } B_\rho. \tag{3.4}$$

Consider now $\bar{x}_1 \in [-1, 1]$ (to be fixed later) and define $\psi_1(x') := \psi(\bar{x}_1, x')$ with $x' \in \mathbb{R}^{n-1}$. Integrating (3.4) with respect to x_1 inside $B_{\rho/2}$, we get

$$|\psi - \psi_1| \le C_n \frac{\rho^3}{R} \qquad \text{inside } B_{\rho/2}.$$

Thus, using the interpolation inequality

$$\|\nabla\psi - \nabla\psi_1\|^2_{L^\infty(B_{\rho/4})} \le C_n \|\psi - \psi_1\|_{L^\infty(B_{\rho/2})} \|D^2\psi - D^2\psi_1\|_{L^\infty(B_{\rho/2})}$$

and recalling that $\|D^2\psi\|_{L^\infty(\mathbb{R}^n)} \le 1$ (hence $\|D^2\psi_1\|_{L^\infty(\mathbb{R}^{n-1})} \le 1$), we get

$$|\nabla\psi - \nabla\psi_1| \le C_n \frac{\rho^{3/2}}{R^{1/2}} \qquad \text{inside } B_{\rho/4}.$$

If $k = 1$ we stop here, otherwise we notice that (3.1) implies that

$$\int_{\mathbb{R}} d\gamma_1(x_1) \int_{\mathbb{R}^{n-1}} \det D^2_{x'x'}\psi(x_1, x')\, d\gamma_{n-1}(x') \le \int_{\mathbb{R}} d\gamma_1(x_1) \int_{\mathbb{R}^{n-1}} \lambda_2(D^2\psi)(x_1, x')\, d\gamma_{n-1}(x') \le \varepsilon,$$

where we used that[1]

$$\lambda_1\big(D^2\psi|_{\{0\}\times\mathbb{R}^{n-1}}\big) \le \lambda_2(D^2\psi)$$

and that (since $0 \le D^2\psi \le \mathrm{Id}_n$)

$$\det D^2_{x'x'}\psi(x_1, x') \le \lambda_1\big(D^2\psi|_{\{0\}\times\mathbb{R}^{n-1}}\big).$$

Hence, by Fubini's Theorem, there exists $\bar{x}_1 \in [-1, 1]$ such that $\psi_1(x') = \psi(\bar{x}_1, x')$ satisfies

$$\int_{\mathbb{R}^{n-1}} \det D^2\psi_1\, d\gamma_{n-1}(x) \le C_n \varepsilon.$$

[1] This inequality follows from the general fact that, given $A \in \mathbb{R}^{n\times n}$ symmetric matrix and $W \subset \mathbb{R}^n$ a k-dimensional vector space,

$$\lambda_1\big(A|_W\big) = \min_{v\in W} \frac{Av\cdot v}{|v|^2} \le \max_{\substack{v\in W'\subset\mathbb{R}^n \\ W'\ k\text{-dim}}} \min_{W'} \frac{Av\cdot v}{|v|^2} = \lambda_{n-k+1}(A).$$

This allows us to repeat the argument above in $\mathbb{R}^{n-1}$ with

$$\widetilde{\psi}_1(x') := \psi_1(x') - \nabla_{x'}\psi_1(0)\cdot x' - \psi_1(0)$$

in place of ψ, and up to a rotation we deduce that

$$|\nabla\widetilde{\psi}_1 - \nabla\psi_2| \le C_n\frac{\rho^{3/2}}{R^{1/2}} \qquad \text{inside } B_{\rho/4}.$$

where $\psi_2(x'') := \psi_1(\bar{x}_2, x'')$, where $\bar{x}_2 \in [-1,1]$ is arbitrary. By triangle inequality, this yields

$$|\nabla\psi + p' - \nabla\psi_2| \le C_n\frac{\rho^{3/2}}{R^{1/2}} \qquad \text{inside } B_{\rho/4},$$

where $p' = -(0, \nabla_{x'}\psi(\bar{x}_1, 0))$. Note that, since $|\bar{x}_1| \le 1$, $\nabla\psi(0) = 0$, and $\|D^2\psi\|_\infty \le 1$, we have $|p| \le 1$. Iterating this argument k times, we conclude that

$$|\nabla\psi + \bar{p} - \nabla\psi_k| \le C_n\frac{\rho^{3/2}}{R^{1/2}} \qquad \text{inside } B_{\rho/4}$$

where $\bar{p} = (p, p'') \in \mathbb{R}^k \times \mathbb{R}^{n-k} = \mathbb{R}^n$ with $|\bar{p}| \le C_n$,

$$\psi_k(y) := \psi(\bar{x}_1, \dots, \bar{x}_k, y), \qquad y \in \mathbb{R}^{n-k},$$

and $\bar{x}_i \in [-1,1]$. Recalling that $\nabla\varphi = x - \nabla\psi$, we have proved that

$$T(x) = \nabla\varphi(x) = (x_1 + p_1, \dots, x_k + p_k, S(y) + p'') + Q(x),$$

where $Q := -(\nabla\psi - \nabla\psi_k + \bar{p})$ satisfies

$$\|Q\|_{L^\infty(B_\rho)} \le C_n\frac{\rho^{3/2}}{R^{1/2}} \qquad \text{and} \qquad |Q(x)| \le C_n(1 + |x|)$$

(in the second bound we used that $T(0) = \nabla\varphi(0) = 0$, $|p| \le C_n$, and T is 1-Lipschitz). Hence, if we set $\nu := (S + p'')_\#\gamma_{n-k}$, we have

$$W_1(\mu, \gamma_{p,k}\otimes\nu) \le \int |Q|\,d\gamma_n \le C_n\frac{\rho^{3/2}}{R^{1/2}} + C_n\int_{\mathbb{R}^n\setminus B_\rho}|x|\,d\gamma_n = C_n\frac{\rho^{3/2}}{R^{1/2}} + C_n\rho^n e^{-\rho^2/2},$$

so, by choosing $\rho := (\log R)^{1/2}$, we get

$$W_1(\mu, \gamma_{p,k}\otimes\nu) \lesssim \frac{1}{R^{1/2-}}.$$

Consider now $\pi_k : \mathbb{R}^n \to \mathbb{R}^n$ and $\bar{\pi}_{n-k} : \mathbb{R}^n \to \mathbb{R}^{n-k}$ the orthogonal projection onto the first k and the last $n-k$ coordinates, respectively. Define $\mu_1 := (\pi_k)_\#(e^{-V}dx)$, $\mu_2 := (\bar{\pi}_{n-k})_\#(e^{-V}dx)$, and note that these are 1-log-concave measures in $\mathbb{R}^k$ and $\mathbb{R}^{n-k}$ respectively (see [1, Theorem 4.3] or [9, Theorem 3.8]). In particular $\mu_2 = e^{-W}$ with $D^2W \ge \mathrm{Id}_{n-k}$. Moreover, since W_1 decreases under orthogonal projection,

$$W_1(\mu_2, \nu) = W_1\big((\bar{\pi}_{n-k})_\#\mu, (\bar{\pi}_{n-k})_\#(\gamma_{p,k}\otimes\nu)\big) \le W_1(\mu, \gamma_{p,k}\otimes\nu) \lesssim \frac{1}{R^{1/2-}},$$

thus

$$\begin{aligned} W_1(\mu, \gamma_{p,k}\otimes\mu_2) &\le W_1(\mu, \gamma_{p,k}\otimes\nu) + W_1(\gamma_{p,k}\otimes\nu, \gamma_{p,k}\otimes\mu_2) \\ &\le W_1(\mu, \gamma_{p,k}\otimes\nu) + W_1(\nu, \mu_2) \lesssim \frac{1}{R^{1/2-}} \end{aligned}$$

where we used the elementary fact that $W_1(\gamma_{p,k}\otimes\nu,\gamma_{p,k}\otimes\mu_2)\le W_1(\nu,\mu_2)$. Recalling (3.3), this proves that

$$W_1(\mu,\gamma_{p,k}\otimes\mu_2)\lesssim\frac{1}{|\log\varepsilon|^{1/4-}},$$

concluding the proof. □

4. Proof of Theorem 1.5

Proof of Theorem 1.5. As in the proof of Theorem 1.3, it is enough to prove the result when $\varepsilon\le\varepsilon_0\ll 1$.

Let $\{u_i\}_{1\le i\le k}$ be as in the statement, and set $v_i:=u_i\circ T$, where $T=\nabla\varphi:\mathbb{R}^n\to\mathbb{R}^n$ is the Brenier map from γ_n to μ. Note that since $T_\#\gamma_n=\mu$,

$$\int v_i\,d\gamma_n=\int u_i\circ T\,d\gamma_n=\int u_i\,d\mu=0.$$

Also, since $|\nabla T|\le 1$ and by our assumption on u_i,

$$\begin{aligned}\int|\nabla v_i|^2\,d\gamma_n&\le\int|\nabla u_i|^2\circ T\,d\gamma_n=\int|\nabla u_i|^2\,d\mu\\&\le(1+\varepsilon)\int u_i^2\,d\mu=(1+\varepsilon)\int v_i^2\,d\gamma_n\le(1+\varepsilon)\int|\nabla v_i|^2\,d\gamma_n,\end{aligned}$$

where the last inequality follows from the Poincaré inequality for γ_n applied to v_i. Since

$$\int|\nabla u_i|^2\,d\mu\le 1+\varepsilon,$$

this proves that

$$0\le\int\Big(|\nabla u_i|^2\circ T-|\nabla v_i|^2\Big)\,d\gamma_n\le\varepsilon\int|\nabla v_i|^2\,d\mu\le\varepsilon(1+\varepsilon).\tag{4.1}$$

Moreover, by Theorem 1.1, $\nabla T=D^2\varphi$ is a symmetric matrix satisfying $0\le\nabla T\le\mathrm{Id}_n$, therefore $(\mathrm{Id}-\nabla T)^2\le\mathrm{Id}-(\nabla T)^2$. Hence, since $\nabla v_i=\nabla T\cdot\nabla u_i\circ T$, it follows by (4.1) that

$$\begin{aligned}\int|\nabla u_i\circ T-\nabla v_i|^2\,d\gamma_n&=\int|(\mathrm{Id}_n-\nabla T)\cdot\nabla u_i\circ T|^2\,d\gamma_n\\&=\int(\mathrm{Id}_n-(\nabla T))^2[\nabla u_i\circ T,\nabla u_i\circ T]\,d\gamma_n\\&\le\int(\mathrm{Id}_n-(\nabla T)^2)[\nabla u_i\circ T,\nabla u_i\circ T]\,d\gamma_n\\&=\int\Big(|\nabla u_i|^2\circ T-|\nabla v_i|^2\Big)\,d\gamma_n\le 2\varepsilon,\end{aligned}\tag{4.2}$$

where, given a matrix A and a vector v, we have used the notation $A[v,v]$ for $Av\cdot v$. In particular, recalling the orthogonality constraint $\int\nabla u_i\cdot\nabla u_j\,d\mu=0$, we deduce that

$$\int\nabla v_i\cdot\nabla v_j\,d\gamma_n=O(\sqrt{\varepsilon}).\tag{4.3}$$

In addition, if we set

$$f_i(x):=\frac{\nabla u_i\circ T(x)}{|\nabla u_i\circ T(x)|}$$

then, using again that $|\nabla T| \leq 1$,

$$\int |\nabla(u_i \circ T)|^2 \Big(1 - |\nabla T \cdot f_i|^2\Big)\, d\gamma \leq \int |\nabla u_i|^2 \circ T \Big(1 - |\nabla T \cdot f_i|^2\Big)\, d\gamma_n \leq 2\varepsilon. \tag{4.4}$$

Now, for $j \in \mathbb{N}$, let $H_j : \mathbb{R} \to \mathbb{R}$ be the one dimensional Hermite polynomial of degree j:

$$H_j(t) = \frac{(-1)^j}{\sqrt{j!}} e^{t^2/2} \left(\frac{d}{dt}\right)^j e^{-t^2/2}$$

see [7, Section 9.2]. It is well known (see for instance [7, Theorem 9.7]) that for $J = (j_1, \ldots, j_n) \in \mathbb{N}^n$ the functions

$$H_J(x_1, \ldots, x_n) := H_{j_1}(x_1) H_{j_2}(x_2) \cdot \cdots \cdot H_{j_n}(x_n)$$

form a Hilbert basis of $L^2(\mathbb{R}^n, \gamma_n)$. Hence, since $\alpha_0^i = \int v_i\, d\gamma_n = 0$, we can write

$$v_i = \sum_{J \in \mathbb{N}^n \setminus \{0\}} \alpha_J^i H_J.$$

By elementary computations (see for instance [7, Proposition 9.3]), we get

$$1 = \int v_i^2 d\gamma_n = \sum_{J \in \mathbb{N}^n \setminus \{0\}} \big(\alpha_J^i\big)^2, \qquad \int |\nabla v_i|^2 d\gamma_n = \sum_{J \in \mathbb{N}^n \setminus \{0\}} |J| \big(\alpha_J^i\big)^2,$$

where $|J| = \sum_{m=1}^n j_m$. Hence, combining the above equations with the bound $\int |\nabla v_i|^2 d\gamma_n \leq (1+\varepsilon)$, we obtain

$$\varepsilon \geq \int |\nabla v_i|^2 d\gamma_n - \int v_i^2 d\gamma_n = \sum_{J \in \mathbb{N}^n, |J| \geq 2} (|J| - 1) \big(\alpha_J^i\big)^2 \geq \frac{1}{2} \sum_{J \in \mathbb{N}^n, |J| \geq 2} |J| \big(\alpha_J^i\big)^2.$$

Recalling that the first Hermite polynomials are just linear functions (since $H_1(t) = t$), using the notation

$$\alpha_j^i := \alpha_J^i \qquad \text{with } J = e_j \in \mathbb{N}^n$$

we deduce that

$$v_i(x) = \sum_{j=1}^n \alpha_j^i x_j + z(x), \qquad \text{with} \qquad \|z\|^2_{W^{1,2}(\mathbb{R}^n, \gamma_n)} = O(\varepsilon).$$

In particular, if we define the vector

$$V_i := \sum_{j=1}^n \alpha_j^i e_j \in \mathbb{R}^n,$$

and we recall that $\int |\nabla v_i|^2\, d\gamma_n = 1 + O(\varepsilon)$ and the almost orthogonality relation (4.3), we infer that $|V_i| = 1 + O(\varepsilon)$ and $|V_i \cdot V_l| = O(\sqrt{\varepsilon})$ for all $i \neq l \in \{1, \ldots, k\}$.

Hence, up to a rotation, we can assume that $|V_i - e_i| = O(\sqrt{\varepsilon})$ for all $i = 1, \ldots, k$, and (4.2) yields

$$\int |\nabla(u_i \circ T) - e_i|^2\, d\gamma_n \leq C\, \varepsilon. \tag{4.5}$$

Since $0 \leq 1 - |\nabla T \cdot f_i|^2 \leq 1$, it follows by (4.4) and (4.5) that

$$\int \Big(1 - |\nabla T \cdot f_i|^2\Big)\, d\gamma_n \leq 2 \int \Big(|\nabla(u_i \circ T)|^2 + |\nabla(u_i \circ T) - e_i|^2\Big)\Big(1 - |\nabla T \cdot f_i|^2\Big)\, d\gamma_n \leq C\varepsilon. \tag{4.6}$$

Set $w_i := \nabla u_i \circ T$ so that $f_i = \frac{w_i}{|w_i|}$. We note that, since all the eigenvalues of $\nabla T = D^2\varphi$ are bounded by 1, given $\delta \ll 1$ the following holds: whenever

$$|\nabla T \cdot w_i - e_i| \le \delta \qquad \text{and} \qquad |\nabla T \cdot f_i| \ge 1-\delta$$

then $|w_i| = 1 + O(\delta)$. In particular,

$$|\nabla T \cdot f_i - e_i| \le C\delta.$$

Hence, if $\delta \le \delta_0$ where δ_0 is a small geometric constant, this implies that the vectors f_i are a basis of $\mathbb{R}^k$, and

$$\nabla T|_{\mathrm{span}(f_1,\ldots,f_k)} \ge (1 - C_0\delta)\,\mathrm{Id}_k$$

for some dimensional constant C_0. Defining $\psi(x) := |x|^2/2 - \varphi(x)$, this proves that

$$\left\{x\,:\,\sum_{i=1}^k\Big[|\nabla T(x)\cdot w_i(x) - e_i| + \Big(1-|\nabla T(x)\cdot f_i(x)|\Big)\Big] \le \delta\right\} \subset \left\{x\,:\,\lambda_{n-k+1}(D^2\psi(x)) \le C_0\delta\right\} \tag{4.7}$$

for all $0 < \delta \le \delta_0$. Hence, by the layer-cake formula, (4.5), and (4.6),

$$\begin{aligned}
&\int_{\{\lambda_{n-k+1}(D^2\psi)\le C_0\delta_0\}} \lambda_{n-k+1}(D^2\psi)\,d\gamma_n = C_0\int_0^{\delta_0}\gamma_n\big(\{\lambda_{n-k+1}(D^2\psi) > C_0 s\}\big)\,ds\\
&\le C_0\int_0^{\delta_0}\gamma_n\left(\left\{\sum_{i=1}^k\Big[|\nabla T(x)\cdot w_i(x) - e_i| + \Big(1-|\nabla T(x)\cdot f_i(x)|\Big)\Big] > s\right\}\right)ds\\
&\le C_0\sum_{i=1}^k\int\Big(|\nabla T\cdot w_i - e_i| + \big(1-|\nabla T\cdot f_i|\big)\Big)\,d\gamma_n \le C\sqrt{\varepsilon}.
\end{aligned} \tag{4.8}$$

On the other hand, it follows by (4.7) that

$$\begin{aligned}
&\left\{x\,:\,\lambda_{n-k+1}(D^2\psi(x)) > C_0\delta\right\}\\
&\qquad\subset\bigcup_{i=1}^k\left[\left\{x\,:\,|\nabla T(x)\cdot w_i(x) - e_i| > \frac{\delta}{2k}\right\}\cup\left\{x\,:\,\Big(1-|\nabla T(x)\cdot f_i(x)|\Big) > \frac{\delta}{2k}\right\}\right].
\end{aligned}$$

Thus, (4.5), (4.6), and Chebishev's inequality yield

$$\begin{aligned}
\gamma_n\big(\{\lambda_{n-k+1}(D^2\psi) > C_0\delta_0\}\big) &\le \sum_{i=1}^k\gamma_n\left(\left\{|\nabla T\cdot w_i - e_i| > \frac{\delta_0}{2k}\right\}\right)\\
&\qquad + \sum_{i=1}^k\gamma_n\left(\left\{1-|\nabla T\cdot f_i| > \frac{\delta_0}{2k}\right\}\right) \le C\,\frac{\varepsilon}{\delta_0^2}.
\end{aligned} \tag{4.9}$$

Hence, since δ_0 is a small but fixed geometric constant, combining (4.8) and (4.9), and recalling that $\lambda_{n-k+1}(D^2\psi) \le 1$, we obtain

$$\int\lambda_{n-k+1}(D^2\psi)\,d\gamma_n \le C\sqrt{\varepsilon}.$$

This implies that (1.4) holds with $C\sqrt{\varepsilon}$ in place of ε, and the result follows by Theorem 1.3. $\square$

Acknowledgements. G.D.P. is supported by the MIUR SIR-grant "Geometric Variational Problems" (RBSI14RVEZ). G.D.P is a member of the "Gruppo Nazionale per l'Analisi Matematica, la Probabilità e le loro Applicazioni" (GNAMPA) of the Istituto Nazionale di Alta Matematica (INdAM). A.F. was partially supported by NSF Grants DMS-1262411 and DMS-1361122.

References

[1] Brascamp H., Lieb E: *On extensions of the Brunn-Minkowski and Prékopa-Leindler theorems, including inequalities for log con- cave functions, and with an application to the diffusion equation* . J. Functional Analysis **22** (1976) 366–389.
[2] Brenier Y.: *Polar factorization and monotone rearrangement of vector-valued functions.* Comm. Pure Appl. Math. **44** (1991), no. 4, 375–417.
[3] Caffarelli L.: *Boundary regularity of maps with convex potentials.* Comm. Pure Appl. Math. **45** (1992), no. 9, 1141–1151.
[4] Caffarelli L: *Monotonicity properties of optimal transportation and the FKG and related inequalities.* Comm. Math. Phys. **214** (2000), 547–563.
[5] Caffarelli L: *Erratum: Monotonicity properties of optimal transportation and the FKG and related inequalities.* Comm. Math. Phys **225** (2002), 449–450.
[6] Cheng X., Zho D.: *Eigenvalues of the drifted Laplacian on complete metric measure spaces.* Commun. Contemp. Math. `http://dx.doi.org/10.1142/S0219199716500012`.
[7] Da Prato, Giuseppe: *An introduction to infinite-dimensional analysis.*. Universitext. Springer-Verlag, Berlin, 2006. x+209 pp.
[8] Figalli, A: *The Monge-Ampère Equation and its Applications.* Zürich Lectures in Advanced Mathematics, to appear.
[9] Saumard A., Wellner J.: *Log-concavity and strong log-concavity: a review.* Stat. Surv. **8** (2014), 45–114

SISSA, Via Bonomea 265, 34136 Trieste, Italy.
E-mail address: `guido.dephilippis@sissa.it`

ETH Zürich, Department of Mathematics, Rämistrasse 101, 8092 Zürich, Switzerland.
E-mail address: `alessio.figalli@math.ethz.ch`