Eigenvalue distribution of optimal transportation

Bo'az B. Klartag,[*]   Alexander V. Kolesnikov[†]

Abstract

We investigate the Brenier map $\nabla\Phi$ between the uniform measures on two convex domains in $\mathbb{R}^n$, or more generally, between two log-concave probability measures on $\mathbb{R}^n$. It is shown that the eigenvalues of the Hessian matrix $D^2\Phi$ exhibit remarkable concentration properties on a multiplicative scale, regardless of the choice of the two measures or the dimension n.

1   Introduction

Let μ and ν be two absolutely-continuous probability measures on $\mathbb{R}^n$. It was discovered by Brenier [4] and McCann [19] that there exists a convex function Φ on $\mathbb{R}^n$ with $(\nabla\Phi)_*\mu = \nu$, i.e.,

$$\int_{\mathbb{R}^n} b(\nabla\Phi(x))d\mu(x) = \int_{\mathbb{R}^n} b(x)d\nu(x) \tag{1}$$

for any ν-integrable function $b : \mathbb{R}^n \to \mathbb{R}$. Moreover, the Brenier map $x \mapsto \nabla\Phi(x)$ is uniquely determined μ-almost everywhere. In this paper we consider the case where μ and ν are log-concave probability measures. An absolutely-continuous probability measure on $\mathbb{R}^n$ is called log-concave if it has a density ρ which satisfies

$$\rho(\lambda x + (1-\lambda)y) \geq \rho(x)^\lambda \rho(y)^{1-\lambda} \qquad (x, y \in \mathbb{R}^n, 0 < \lambda < 1).$$

The uniform measure on any convex domain is log-concave, as well as the Gaussian measure. Write $Supp(\mu)$ for the interior of the support of μ, which is an open, convex set in $\mathbb{R}^n$. We make the assumption that

($\star$) The function Φ is C^2-smooth in $Supp(\mu)$.

It follows from the works of Caffarelli [5, 6, 1] that ($\star$) holds true when each of the measures μ and ν satisfies the following extra condition: Either the support of the measure is the entire $\mathbb{R}^n$, or else the support is a bounded, convex domain and density of the measure is bounded away from zero and from infinity in this convex domain. It is fair to say that Caffarelli's regularity theory covers most cases of interest, yet it is very plausible that ($\star$) is in fact always correct, without any extra conditions.

[*]School of Mathematical Sciences, Tel Aviv University, Tel Aviv 69978, Israel. E-mail: klartagb@tau.ac.il

[†]Faculty of Mathematics, National Research University Higher School of Economics, Moscow, Russia. Email: sascha77@mail.ru

As it turns out, the positive-definite Hessian matrix $D^2\Phi(x)$ exhibits remarkable regularity in the behavior of its eigenvalues. We write $Var[X]$ for the variance of the random variable X.

Theorem 1.1. *Let μ, ν be absolutely-continuous, log-concave probability measures on $\mathbb{R}^n$. Let $\nabla\Phi$ be the Brenier map between μ and ν, and assume $(\star)$. Write $0 < \lambda_1(x) \leq \ldots \leq \lambda_n(x)$ for the eigenvalues of the matrix $D^2\Phi(x)$, repeated according to their multiplicity. Let X be a random vector in $\mathbb{R}^n$ that is distributed according to μ. Then, for $i = 1, \ldots, n$,*

$$Var\left[\log\lambda_i(X)\right] \leq 4.$$

Thus, on a multiplicative scale, the eigenvalues of $D^2\Phi$ are quite stable. Note that the multiplicative scale is indeed the natural scale in the generality of Theorem 1.1: By applying appropriate linear transformations to μ and ν, one may effectively multiply all eigenvalues by an arbitrary positive constant. The variance bound in Theorem 1.1 follows from a Poincaré inequality which we now formulate. For $x \in Supp(\mu)$ set

$$\Lambda(x) = \left(\log\lambda_1(x), \ldots, \log\lambda_n(x)\right).$$

We write $|\cdot|$ for the standard Euclidean norm in $\mathbb{R}^n$.

Theorem 1.2. *Under the notation and assumptions of Theorem 1.1, for any locally-Lipschitz function $f : \mathbb{R}^n \to \mathbb{R}$ with $\mathbb{E}\,|f(\Lambda(X))| < \infty$,*

$$Var\left[f(\Lambda(X))\right] \leq 4\mathbb{E}|\nabla f|^2(\Lambda(X)),$$

whenever the right-hand side is finite. At the points in which f is not continuously differentiable, we define $|\nabla f|$ via (36) below.

Denote $\pi = \Lambda_*(\mu)$, the push-forward of the measure μ under the map Λ. Theorem 1.2 is a spectral gap estimate for the metric-measure space $(\mathbb{R}^n, |\cdot|, \pi)$. Gromov and Milman [13] proved that a spectral gap estimate implies exponential concentration of Lipschitz functions. Therefore Theorem 1.2 admits the following immediate corollary:

Corollary 1.3. *We work under the notation and assumptions of Theorem 1.1. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a 1-Lipschitz function (i.e., $|f(x) - f(y)| \leq |x - y|$). Denote $A = \mathbb{E}f(\Lambda(X))$. Then A is finite and*

$$\mathbb{E}\exp(c\,|f(\Lambda(X)) - A|) \leq 2,$$

where $c > 0$ is a universal constant.

Remark 1.4. Corollary 1.3 implies that $\mathbb{E}e^{c|\Lambda(X)|} < \infty$. Consequently, one may replace the condition $\mathbb{E}\,|f(\Lambda(X))| < \infty$ in Theorem 1.2 by the requirement that $e^{-c|x|}|f(x)|$ is bounded in $\mathbb{R}^n$, for a certain universal constant $c > 0$.

Our next result is that the diagonal elements of the matrix $D^2\Phi(x)$ are also concentrated on a logarithmic scale, pretty much like the eigenvalues.

2

Theorem 1.5. *We work under the notation and assumptions of Theorem 1.1. Fix $v \in \mathbb{R}^n$, let $H(x) = \log \left(D^2\Phi(x)v \cdot v \right)$ and denote $Y = H(X)$. Then,*

(i) *$Var[Y] \leq 4$.*

(ii) *For any locally-Lipschitz function $f : \mathbb{R} \to \mathbb{R}$ with $\mathbb{E}|f(Y)| < \infty$,*

$$Var[f(Y)] \leq 4\mathbb{E}|f'|^2(Y).$$

(iii) *For any 1-Lipschitz function $f : \mathbb{R} \to \mathbb{R}$, denoting $A = \mathbb{E}f(Y)$ we have that $A \in \mathbb{R}$ and*

$$\mathbb{E}\exp(c|f(Y) - A|) \leq 2,$$

where $c > 0$ is a universal constant.

All of the assertions made so far follow from Theorem 5.1 below, which is in fact a sound reformulation of [14, Theorem 1.4]. The results in [14] were obtained under a technical assumption dubbed "regularity at infinity", which we have to address in this paper. Our argument is based on analysis of the transportation metric: This means that we use the positive-definite Hessian $D^2\Phi$ in order to define a Riemannian metric in $Supp(\mu)$. The weighted Riemannian manifold

$$M_{\mu,\nu} = \left(Supp(\mu), D^2\Phi, \mu \right)$$

was studied in [17], where it was shown that the associated Ricci-Bakry-Émery tensor is non-negative when μ and ν are log-concave. We will also consider the map

$$x \mapsto D^2\Phi(x)$$

from $Supp(\mu) \subseteq \mathbb{R}^n$ into the space of positive-definite matrices. The space of positive-definite matrices is endowed with a natural Riemannian metric, which fits very nicely with computations related to the weighted Riemannian manifold $M_{\mu,\nu}$. This leads to a certain Poincaré inequality with respect to the standard Riemannian metric on the space of positive-definite matrices, formulated in Theorem 5.1 below .

We have tried to make the exposition self-contained, apart from the regularity theory of mass-transport. The rest of this paper is organized as follows: In Section 2 we recall some well-known constructions related to positive-definite matrices. In Section 3 and Section 4 we prove the main results under regularity assumptions by employing the Bakry-Émery Γ_2-calculus. Section 5 is devoted to the elimination of these regularity assumptions. In Section 6 we complete the proofs of the theorems formulated above. Below we write $x \cdot y$ for the standard scalar product of $x, y \in \mathbb{R}^n$. We denote derivatives by $\partial_k f = f_k = \partial f / \partial x_k$ and $f_{ij} = \partial^2 f / (\partial x_i \partial x_j)$. By a smooth function we mean a C^∞-smooth one. We write $\log$ for the natural logarithm, and $Tr(A)$ stands for the trace of the matrix A.

2 Positive-definite quadratic forms

This section surveys standard material on positive-definite matrices. Denote by $M_n^+(\mathbb{R})$ the collection of all symmetric, positive-definite $n \times n$ matrices. For a function $f : (0, \infty) \to \mathbb{R}$ and $A \in M_n^+(\mathbb{R})$ we may define the symmetric matrix $f(A)$ via the spectral theorem. In other words,

$$f\left(\sum_{i=1}^n \lambda_i v_i \otimes v_i\right) = \sum_{i=1}^n f(\lambda_i) v_i \otimes v_i$$

for any orthonormal basis $v_1, \ldots, v_n \in \mathbb{R}^n$ and $\lambda_1, \ldots, \lambda_n > 0$, where we write $x \otimes x = (x_i x_j)_{i,j=1,\ldots,n}$ for $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$.

Lemma 2.1. *For any $A, B \in M_n^+(\mathbb{R})$,*

$$\left\|\log\left(A^{1/2} B A^{1/2}\right)\right\|_{HS} \leq \|\log(A)\|_{HS} + \|\log(B)\|_{HS} \tag{2}$$

where $\|\cdot\|_{HS}$ stands for the Hilbert-Schmidt norm.

Proof. For an $n \times n$ matrix T and $k = 1, \ldots, n$ we define

$$D_k(T) = \sup_{\substack{E \subseteq \mathbb{R}^n \\ \dim(E)=k}} \frac{Vol_k(T(B^n \cap E))}{Vol_k(B^n \cap E)}, \tag{3}$$

where $B^n = \{x \in \mathbb{R}^n \, ; \, |x| < 1\}$, and the supremum in (3) runs over all k-dimensional subspaces in $\mathbb{R}^n$. Thus, an application of the linear transformation A may increase k-dimensional volumes by a factor of at most $D_k(A)$. It follows that for any $n \times n$ matrices A and B,

$$D_k(AB) \leq D_k(A)D_k(B) \qquad (k = 1, \ldots, n). \tag{4}$$

In the case in which $A \in M_n^+(\mathbb{R})$, we have $D_k(A) = \prod_{i=1}^k \lambda_i$, where $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n > 0$ are the eigenvalues of A. Assume that $A, B \in M_n^+(\mathbb{R})$. Denote the eigenvalues of the symmetric, positive-definite matrix $A^{1/2}BA^{1/2}$ by $e^{\gamma_1} \geq \ldots \geq e^{\gamma_n} > 0$. Then, for $k = 1, \ldots, n$,

$$\prod_{i=1}^k e^{\gamma_i} = D_k\left(A^{1/2}BA^{1/2}\right) \leq D_k(A^{1/2})D_k(B)D_k(A^{1/2}) = D_k(A)D_k(B) = \prod_{i=1}^k (e^{\alpha_i} e^{\beta_i}), \tag{5}$$

where $e^{\alpha_1} \geq \ldots \geq e^{\alpha_n} > 0$ are the eigenvalues of A and $e^{\beta_1} \geq \ldots \geq e^{\beta_n} > 0$ are the eigenvalues of B. We will now apply a lemma of Weyl [25], see also Polya [20]. According to the inequality of Weyl and Polya, the inequalities (5) entail that

$$\sum_{i=1}^n h(\gamma_i) \leq \sum_{i=1}^n h(\alpha_i + \beta_i) \tag{6}$$

for any convex, non-decreasing function $h : \mathbb{R} \to \mathbb{R}$. For $t \in \mathbb{R}$ denote $t_+ = \max\{t, 0\}$. The function $t \mapsto (t_+)^2$ is convex and non-decreasing, hence from (6),

$$\sum_{i=1}^n ((\gamma_i)_+)^2 \leq \sum_{i=1}^n ((\alpha_i + \beta_i)_+)^2. \tag{7}$$

4

By using (4) for the inverse matrices, we conclude that for $k = 1, \ldots, n$,

$$\prod_{i=n-k+1}^{n} e^{-\gamma_i} = D_k \left(A^{-1/2} B^{-1} A^{-1/2} \right) \leq D_k(A^{-1}) D_k(B^{-1}) = \prod_{i=n-k+1}^{n} (e^{-\alpha_i} e^{-\beta_i}).$$

The inequality of Weyl and Polya now implies that $\sum_{i=1}^{n} h(-\gamma_i) \leq \sum_{i=1}^{n} h(-\alpha_i - \beta_i)$ for any convex, non-decreasing function h. By using again $h(t) = (t_+)^2$, we get

$$\sum_{i=1}^{n} ((-\gamma_i)_+)^2 \leq \sum_{i=1}^{n} ((-\alpha_i - \beta_i)_+)^2. \tag{8}$$

Adding (7) and (8), we finally obtain

$$\sum_{i=1}^{n} \gamma_i^2 \leq \sum_{i=1}^{n} (\alpha_i + \beta_i)^2 \leq \left(\sqrt{\sum_{i=1}^{n} \alpha_i^2} + \sqrt{\sum_{i=1}^{n} \beta_i^2} \right)^2, \tag{9}$$

where we used the Cauchy-Schwartz inequality in the last passage. By taking the square root of (9) we deduce (2). $\qquad \square$

For two matrices $A, B \in M_n^+(\mathbb{R})$ set

$$dist(A, B) = \left\| \log \left(A^{-1/2} B A^{-1/2} \right) \right\|_{HS}. \tag{10}$$

Equivalently, $dist(A, B)$ equals $\sqrt{\sum_i \log^2 \lambda_i}$, where $\lambda_1, \ldots, \lambda_n > 0$ are the eigenvalues of the matrix $A^{-1}B$ which is conjugate to $A^{-1/2} B A^{-1/2}$. The latter equivalent definition of $dist$ shows that for any invertible $n \times n$ matrix T,

$$dist(A, B) = dist\left(T^t A T, T^t B T \right) \qquad (A, B \in M_n^+(\mathbb{R}^n)), \tag{11}$$

where A^t is the transpose of the matrix A. Observe also that $dist(A, B) = dist\left(A^{-1}, B^{-1} \right)$ for any $A, B \in M_n^+(\mathbb{R})$. Lemma 2.1 states that for $A, B \in M_n^+(\mathbb{R})$,

$$dist(A, B) \leq dist(A, \mathrm{Id}) + dist(\mathrm{Id}, B), \tag{12}$$

where Id is the identity matrix. From (11) and (12) one realizes that $dist$ satisfies the triangle inequality in $M_n^+(\mathbb{R})$, hence it is a metric. For $A \in M_n^+(\mathbb{R}^n)$ and a symmetric $n \times n$ matrix B we denote

$$\|B\|_A = \left\| A^{-1/2} B A^{-1/2} \right\|_{HS} = \sqrt{Tr\left[(A^{-1}B)^2 \right]}.$$

For a smooth curve $\gamma : [a, b] \to M_n^+(\mathbb{R})$ set

$$Length(\gamma) = \int_a^b \|\dot{\gamma}(s)\|_{\gamma(s)} \, ds, \tag{13}$$

where $\dot{\gamma}(s) = \frac{d\gamma(s)}{ds}$ is a symmetric $n \times n$ matrix. Then $Length$ is invariant under conjugations. That is, the length of the curve $\gamma(s)$ equals that of the curve $T^t \gamma(s) T$ for any invertible $n \times n$ matrix T.

Lemma 2.2. *(i) For any $A \in M_n^+(\mathbb{R}^n)$ and a symmetric $n \times n$ matrix B,*

$$\lim_{\varepsilon \to 0} \frac{dist^2(A + \varepsilon B, A)}{\varepsilon^2} = \|B\|_A^2 = Tr\left[(A^{-1}B)^2\right]. \tag{14}$$

(ii) Let $A, B \in M_n^+(\mathbb{R}^n)$ and consider the curve

$$\gamma_{A,B}(s) = A^{1/2}\left(A^{-1/2}BA^{-1/2}\right)^s A^{1/2} \qquad (0 \le s \le 1).$$

Then $\gamma_{A,B}$ is a curve connecting A and B with $Length(\gamma_{A,B}) = dist(A, B)$.

Proof. The invariance property (11) implies that

$$dist(A + \varepsilon B, A) = dist(\mathrm{Id} + \varepsilon A^{-1/2}BA^{-1/2}, \mathrm{Id}).$$

It therefore suffices to prove (i) under the additional assumption that $A = \mathrm{Id}$. Let $\lambda_1, \ldots, \lambda_n > 0$ be the eigenvalues of B. It follows from (10) that

$$\lim_{\varepsilon \to 0} \frac{dist^2(\mathrm{Id} + \varepsilon B, \mathrm{Id})}{\varepsilon^2} = \lim_{\varepsilon \to 0} \frac{\sum_{i=1}^n \log^2(1 + \varepsilon \lambda_i)}{\varepsilon} = \sum_{i=1}^n \lambda_i^2,$$

and (i) follows from the fact that $\|B\|_A^2 = \sum_i \lambda_i^2$. We now turn to the proof of (ii). Again, we may reduce matters to the case where $A = \mathrm{Id}$ by noting that

$$\gamma_{A,B}(s) = A^{1/2}\gamma_{\mathrm{Id}, A^{-1/2}BA^{-1/2}}(s)A^{1/2} \qquad (0 \le s \le 1).$$

Abbreviate $\gamma(s) = \gamma_{A,B}(s) = \gamma_{\mathrm{Id},B}(s)$. Since $\gamma(s) = B^s$ then $\dot{\gamma}(s) = B^s \log(B)$ and hence, for any $0 \le s \le 1$,

$$\|\dot{\gamma}(s)\|_{\gamma(s)} = \left\| B^{-s/2}\left(B^s \log(B)\right)B^{-s/2}\right\|_{HS} = \|\log(B)\|_{HS} = dist(\mathrm{Id}, B).$$

From the definition (13) it follows that $Length(\gamma) = dist(\mathrm{Id}, B)$, and (ii) is proven. $\qquad \square$

The right-hand side of (14) depends quadratically on B, and therefore Lemma 2.2 tells us that our distance function $dist$ on $M_n^+(\mathbb{R})$ is induced by a Riemannian metric. We refer to this Riemannian metric as the *standard Riemannian metric* on $M_n^+(\mathbb{R})$. The next two lemmas describe certain Lipschitz functions on $M_n^+(\mathbb{R})$.

Lemma 2.3. *Fix $v \in \mathbb{R}^n$ and set $f(A) = \log(Av \cdot v)$ for $A \in M_n^+(\mathbb{R})$. Then f is a 1-Lipschitz function with respect to the standard Riemannian metric on $M_n^+(\mathbb{R})$.*

Proof. The map f is clearly smooth. Fix $A \in M_n^+(\mathbb{R})$ and let us show that the norm of the Riemannian gradient of f at the point A is bounded by one. For any symmetric $n \times n$ matrix B we have

$$\frac{d}{dt}f(A + tB)\Big|_{t=0} = \frac{Bv \cdot v}{Av \cdot v}.$$

Thus, in order to prove the lemma, it suffices to show that

$$\frac{Bv \cdot v}{Av \cdot v} \le \|B\|_A = \|A^{-1/2}BA^{-1/2}\|_{HS}. \tag{15}$$

6

By switching to another orthonormal basis, if necessary, we may assume that A is a diagonal matrix. Denote by $\lambda_1, \ldots, \lambda_n > 0$ the numbers on the diagonal of A. Denote $B = (b_{ij})_{i,j=1,\ldots,n}$ and $v = (v_1, \ldots, v_n) \in \mathbb{R}^n$. From the Cauchy-Schwartz inequality,

$$\sum_{i,j=1}^n b_{ij} v_i v_j \leq \sqrt{\sum_{i,j=1}^n \frac{b_{ij}^2}{\lambda_i \lambda_j}} \sqrt{\sum_{i,j=1}^n \lambda_i \lambda_j v_i^2 v_j^2} = \sqrt{\sum_{i,j=1}^n \frac{b_{ij}^2}{\lambda_i \lambda_j}} \left(\sum_{i=1}^n \lambda_i v_i^2 \right),$$

which is equivalent to the desired inequality (15). $\qquad\square$

Lemma 2.4. *For $A \in M_n^+(\mathbb{R})$ denote its eigenvalues by $\lambda_1(A) \geq \ldots \geq \lambda_n(A) > 0$. Consider the map $\Lambda : M_n^+(\mathbb{R}) \to \mathbb{R}^n$ defined via*

$$\Lambda(A) = (\log(\lambda_1(A)), \ldots, \log(\lambda_n(A))). \tag{16}$$

Then Λ is a 1-Lipschitz map, with respect to the standard Riemannian metric on $M_n^+(\mathbb{R})$, and the standard Euclidean metric on $\mathbb{R}^n$.

Proof. Write $\mathcal{F} \subseteq M_n^+(\mathbb{R})$ the collection of all positive-definite, symmetric matrices with n distinct eigenvalues. Then $\mathcal{F}$ is an open, dense set. The function Λ is continuous, since the eigenvalues vary continuously with the matrix. It therefore suffices to prove that

$$|\Lambda(A_1) - \Lambda(A_2)| \leq dist(A_1, A_2) \qquad \text{for } A_1, A_2 \in \mathcal{F}.$$

Fix $A_1, A_2 \in \mathcal{F}$. Consider the curve $\gamma(s) = \gamma_{A_1, A_2}(s/dist(A_1, A_2))$ where $\gamma_{A_1, A_2}(s)$ is as in Lemma 2.2. Then γ is a length-minimizing curve between A_1 and A_2, parametrized by Riemannian arclength. We claim that $\gamma(s) \in \mathcal{F}$ for all but finitely many values of s. Indeed, the resultant of $\gamma(s)$ is a real-analytic function of s which is not identically zero, hence its zeros are isolated. Since $\Lambda \circ \gamma$ is continuous, in order to prove the lemma it suffices to show that

$$\left| \frac{d\Lambda(\gamma(s))}{ds} \right| \leq 1 \tag{17}$$

for all s with $\gamma(s) \in \mathcal{F}$. Let us fix s_0 with $\gamma(s_0) \in \mathcal{F}$. Denote $A = \gamma(s_0)$ and $B = \dot{\gamma}(s_0)$. Since γ is parameterized by arclength, then

$$\|B\|_A = \|A^{-1/2} B A^{-1/2}\|_{HS} = 1. \tag{18}$$

Let $v_1, \ldots, v_n \in \mathbb{R}^n$ be the orthonormal basis of eigenvectors that corresponds to the eigenvalues $\lambda_1(A), \ldots, \lambda_n(A)$ of the matrix A. Then,

$$\left. \frac{d\lambda_i(\gamma(s))}{ds} \right|_{s=s_0} = B v_i \cdot v_i \qquad (i = 1, \ldots, n). \tag{19}$$

The relation (19) is standard, see, e.g. Reed and Simon [21, Section XII.1]. Consequently,

$$\left. \frac{d\Lambda(\gamma(s))}{ds} \right|_{s=s_0} = \left(\frac{B v_1 \cdot v_1}{\lambda_1(A)}, \ldots, \frac{B v_n \cdot v_n}{\lambda_n(A)} \right). \tag{20}$$

However, by (18),

$$\sum_{i=1}^{n} \left(\frac{Bv_i \cdot v_i}{\lambda_i(A)} \right)^2 = \sum_{i=1}^{n} \left(A^{-1/2} B A^{-1/2} v_i \cdot v_i \right)^2 \leq \|A^{-1/2} B A^{-1/2}\|_{HS}^2 = 1. \quad (21)$$

Now (17) follows from (20) and (21). $\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

Corollary 2.5. *Whenever A and B are positive-definite $n \times n$ matrices,*

$$\sum_{i=1}^{n} \log^2 \frac{\lambda_i}{\mu_i} \leq \left\| \log \left(A^{-1/2} B A^{-1/2} \right) \right\|_{HS}^2$$

where $\lambda_1 \geq \ldots \geq \lambda_n > 0$ are the eigenvalues of A and $\mu_1 \geq \ldots \geq \mu_n > 0$ are the eigenvalues of B.

3   Bakry-Émery Γ_2-calculus

Let μ and ν be two absolutely-continuous, log-concave probability measures on $\mathbb{R}^n$. Assume that $d\mu = e^{-V(x)} dx$ and $d\nu = e^{-W(x)} dx$, for certain smooth, convex functions $V, W : \mathbb{R}^n \to \mathbb{R}$. Let $\nabla \Phi$ be the Brenier map between μ and ν. Caffarelli's regularity theory states that $\Phi : \mathbb{R}^n \to \mathbb{R}$ is a smooth, convex function. Therefore (1) implies that the transport equation

$$-V(x) = \log \det D^2 \Phi(x) - W(\nabla \Phi(x)) \qquad\qquad (22)$$

holds everywhere in $\mathbb{R}^n$. In particular, the matrix $D^2 \Phi(x) = (\Phi_{ij}(x))_{i,j=1,\ldots,n}$ is invertible and hence positive-definite for any $x \in \mathbb{R}^n$. The inverse matrix to $D^2 \Phi(x)$ is denoted by $\left(D^2 \Phi(x) \right)^{-1} = \left(\Phi^{ij}(x) \right)_{i,j=1,\ldots,n}$. We use the Einstein summation convention, thus an index that appears twice in an expression, once as a subscript and once as a superscript, is being summed upon. We also use abbreviations such as $\Phi^i_{jk} = \Phi^{i\ell} \Phi_{jk\ell}$ and $\Phi^{ij}_k = \Phi^{i\ell} \Phi^{jm} \Phi_{km\ell}$. Differentiating (22), we obtain

$$V_j(x) = -\Phi^i_{ji}(x) + \sum_{i=1}^{n} \Phi_{ij}(x) W_i(\nabla \Phi(x)) \qquad\qquad (j = 1, \ldots, n,\ x \in \mathbb{R}^n). \quad (23)$$

Following [17], we use the positive-definite matrices $D^2 \Phi(x)$ in order to induce a Riemannian metric on $\mathbb{R}^n$, and consider the weighted Riemannian manifold

$$M = M_{\mu,\nu} = \left(\mathbb{R}^n, D^2 \Phi, \mu \right).$$

See Grigor'yan [11] and Bakry, Gentil and Ledoux [3] for background on weighted Riemannian manifolds and the Γ_2-calculus. For a smooth function $u : \mathbb{R}^n \to \mathbb{R}$ we have $|\nabla_M u|_M^2 = \Phi^{ij} u_i u_j$ where $|\nabla_M u|_M^2$ stands for the square of the Riemannian norm of the Riemannian gradient of u. The Dirichlet form associated with the weighted Riemannian manifold $M_{\mu,\nu}$ is defined, for smooth functions $u, v : \mathbb{R}^n \to \mathbb{R}$, via

$$\Gamma(u, v) = \int_{\mathbb{R}^n} \langle \nabla_M u, \nabla_M v \rangle_M \, d\mu = \int_{\mathbb{R}^n} \left(\Phi^{ij} u_i v_j \right) d\mu$$

whenever the integral converges. The Laplacian associated with the weighted Riemannian manifold $M_{\mu,\nu}$ is defined, for a smooth function $u : \mathbb{R}^n \to \mathbb{R}$, by

$$Lu = \Phi^{ij} u_{ij} - \sum_{j=1}^{n} W_j(\nabla\Phi(x))u_j = \Phi^{ij} u_{ij} - \left(\Phi_i^{ij} + \Phi^{ij} V_i\right) u_j, \tag{24}$$

where the last equality holds in view of (23). Integrating by parts, we verify that

$$-\int_{\mathbb{R}^n} (Lu)v d\mu = -\int_{\mathbb{R}^n} \left(\Phi^{ij} u_{ij} - \left[\Phi_i^{ij} + \Phi^{ij} V_i\right] u_j\right) v e^{-V} = \int_{\mathbb{R}^n} \left(\Phi^{ij} u_i v_j\right) d\mu = \Gamma(u,v)$$

for any smooth functions $u, v : \mathbb{R}^n \to \mathbb{R}$ with one of them compactly-supported. The next step is to consider the *Carré du Champ* of $M_{\mu,\nu}$: As in Bakry and Émery [2], for a smooth function $u : K \to \mathbb{R}$ we define

$$\Gamma_2(u) = \frac{1}{2} L\left(|\nabla_M u|_M^2\right) - \langle\nabla_M u, \nabla_M(Lu)\rangle_M = \frac{1}{2} L\left(\Phi^{ij} u_i u_j\right) - \Phi^{ij}(Lu)_i u_j. \tag{25}$$

Lemma 3.1. *For any smooth function $u : \mathbb{R}^n \to \mathbb{R}$ we have the pointwise inequality*

$$\Gamma_2(u) \geq \frac{1}{4} \Phi_\ell^{ik} \Phi_k^{j\ell} u_i u_j.$$

Lemma 3.1 is proven in [14] by introducing a Kähler structure and interpreting the left-hand side of (26) below as the Hilbert-Schmidt norm of a certain Hessian operator restricted to a subspace. There are several additional ways to prove Lemma 3.1. The brute-force way, is to use a tedious but straightforward computation which shows that

$$\Gamma_2(u) = \Phi^{kl} \Phi^{ij} u_{ik} u_{j\ell} - \Phi^{ijk} u_{ij} u_k + \frac{1}{2}\left(\Phi_\ell^{ik} \Phi_k^{j\ell} + \Phi^{ik} \Phi^{j\ell} V_{k\ell}\right) u_i u_j + \frac{1}{2} \sum_{i,j=1}^{n} (W_{ij} \circ \nabla\Phi) u_i u_j.$$

This computation is more or less equivalent to reproving Bochner's formula. Then, one proves the pointwise inequality

$$\Phi^{kl} \Phi^{ij} u_{ik} u_{j\ell} - \Phi^{ijk} u_{ij} u_k + \frac{1}{4} \Phi_\ell^{ik} \Phi_k^{j\ell} u_i u_j \geq 0, \tag{26}$$

by representing the left-hand side of (26) as the trace of the square of the matrix $B = (b_i^j)_{i,j=1,\ldots,n}$ where $b_i^j = \Phi^{jk} u_{ki} - \frac{1}{2}\Phi_i^{jk} u_k$. The product $A = (D^2\Phi)B$ is a symmetric matrix, hence

$$Tr\left(B^2\right) = Tr\left[\left((D^2\Phi)^{-1/2} A (D^2\Phi)^{-1/2}\right)^2\right] \geq 0.$$

Lemma 3.1 follows from (26) and from the fact that $D^2 V$ and $D^2 W$ are positive semi-definite matrices.

Another approach to Lemma 3.1 is to use the notation of Riemannian geometry as in [17], and use the Bochner formula. We first observe that identity (23) in the case $j = 1$ has the simple form

$$L\Phi_1 = -V_1. \tag{27}$$

9

Differentiating (27) and using $\partial_k(\Phi^{ij}) = -\Phi_k^{ij}$, we obtain

$$L(\Phi_{11}) - \Phi_1^{jk}\Phi_{1jk} - \sum_{j,k=1}^n \Phi_{j1}\Phi_{1k}\,(W_{jk} \circ \nabla\Phi) = -V_{11}. \tag{28}$$

The Bochner-Lichnerowicz-Weitzenböck formula states that for any smooth $u : \mathbb{R}^n \to \mathbb{R}$,

$$\Gamma_2(u) = \|D_M^2 u\|_M^2 + Ric_M(\nabla_M u, \nabla_M u), \tag{29}$$

where $\|D_M^2 u\|_M^2$ is the Hilbert-Schmidt norm of the Riemannian Hessian of u, and Ric_M is the Bakry-Émery-Ricci tensor of the weighted Riemmannian manifold $M = M_{\mu,\nu}$. Let us analyze the term in (29) involving the Hessian of u. The Christofell symbols of our Riemannian metric are $\Gamma_{ij}^k = \frac{1}{2}\Phi_{ij}^k$, and therefore $(D_M^2 u)_{ij} = u_{ij} - \frac{1}{2}\Phi_{ij}^k u_k$ and

$$\|D_M^2 u\|_M^2 = \Phi^{ik}\Phi^{jm}\left(u_{ij} - \frac{1}{2}\Phi_{ij}^\ell u_\ell\right)\left(u_{mk} - \frac{1}{2}\Phi_{mk}^s u_s\right).$$

In the particular case of $u = \Phi_1$, we obtain $(D_M^2 \Phi_1)_{jk} = \frac{1}{2}\Phi_{1jk}$ and hence $\|D_M^2 \Phi_1\|_M^2 = \frac{1}{4}\Phi_{1j}^k\Phi_{1k}^j$. Furthermore, the vector field $\nabla_M \Phi_1$ satisfies $\nabla_M \Phi_1 = \partial/\partial x_1$ and $|\nabla_M \Phi_1|_M^2 = \Phi_{11}$. Since $L\Phi_1 = -V_1$, the Bochner formula (29) for $u = \Phi_1$ takes the form

$$\frac{1}{2}L\left(\Phi_{11}\right) = -\langle \nabla_M \Phi_1, \nabla_M V_1\rangle_M + \frac{1}{4}\Phi_{1j}^k\Phi_{1k}^j + Ric_M(\nabla_M u, \nabla_M u)$$

$$= -V_{11} + \frac{1}{4}\Phi_{1j}^k\Phi_{1k}^j + (Ric_M)_{11}. \tag{30}$$

From (28) and (30) we obtain a formula for the Bakry-Émery-Ricci tensor:

$$(Ric_M)_{11} = \frac{1}{4}\Phi_{1j}^k\Phi_{1k}^j + \frac{1}{2}V_{11} + \frac{1}{2}\sum_{j,k=1}^n \Phi_{j1}\Phi_{1k}\,(W_{jk} \circ \nabla\Phi).$$

It is clear that there is nothing special about the derivative $u = \Phi_1$, and that we could have repeated the argument with $u = \nabla\Phi \cdot \theta$ for any $\theta \in \mathbb{R}^n$. We thus obtain the formula

$$(Ric_M)_{i\ell} = \frac{1}{4}\Phi_{ij}^k\Phi_{\ell k}^j + \frac{1}{2}V_{i\ell} + \frac{1}{2}\sum_{j,k=1}^n \Phi_{ji}\Phi_{\ell k}\,(W_{jk} \circ \nabla\Phi). \tag{31}$$

Since D^2V and D^2W are positive semi-definite, then for any smooth $u : \mathbb{R}^n \to \mathbb{R}$,

$$\Gamma_2(u) \geq Ric_M(\nabla_M u, \nabla_M u) \geq \frac{1}{4}\Phi_j^{ik}\Phi_k^{j\ell} u_i u_\ell$$

and the third proof of Lemma 3.1 is complete.

Having finished with Lemma 3.1, let us introduce one of the main ideas in this paper, which was absent from [14]. The idea is to consider the map

$$\mathbb{R}^n \ni x \mapsto D^2\Phi(x) \in M_n^+(\mathbb{R}). \tag{32}$$

Denote by $(g_{ij}(x))_{i,j=1,...,n}$ the pull-back of the standard Riemannian metric on $M_n^+(\mathbb{R})$ via the map (32). It follows from Lemma 2.2 that g_{ij} is given by the formula

$$g_{ij} = Tr\left[(D^2\Phi)^{-1} \cdot \partial_i\left(D^2\Phi\right) \cdot (D^2\Phi)^{-1} \cdot \partial_j\left(D^2\Phi\right)\right] = \Phi_{ik}^\ell \Phi_{j\ell}^k. \tag{33}$$

Note that the positive semi-definite matrix $(g_{ij}(x))_{i,j=1,...,n}$ is not necessarily invertible, and it could happen that distinct points of $\mathbb{R}^n$ have zero Riemannian distance with respect to the Riemannian metric (g_{ij}). The metric g_{ij} resembles an expression appearing in Lemma 3.1, a fact that will be exploited in the next section.

4 Dualizing the Bochner inequality

It is by now well-known that in the presence of convexity assumptions, Poincaré-type inequalities may be deduced from Bochner's formula via a dualization procedure. In this section we investigate the Poincaré inequality that is dual to Lemma 3.1. This Poincaré inequality was also obtained in [14], but in a cumbersome formulation and under an undesired assumption called "regularity at infinite", which we eliminate here.

We begin with an easy case. Throughout this section we assume, in addition to the smoothness assumptions made in the beginning of Section 3, that there exists $\varepsilon_0 > 0$ for which

$$D^2\Phi(x) \geq \varepsilon_0 \cdot \mathrm{Id} \qquad\qquad (x \in \mathbb{R}^n) \tag{34}$$

in the sense of symmetric matrices. Write $C_c^\infty(\mathbb{R}^n)$ for the space of all compactly-supported, smooth functions on $\mathbb{R}^n$. The following lemma is a variant of a well-known fact (see, e.g., Strichartz [23]), that compactly-supported functions are dense in Sobolev spaces when the Riemannian manifold is complete. Our assumption (34) implies the completeness of the Riemannian manifold $M = M_{\mu,\nu}$.

Lemma 4.1. *Let $f \in L^2(\mu)$ satisfy $\int f d\mu = 0$. Then there exists a sequence $u_k \in C_c^\infty(\mathbb{R}^n)$ with*

$$\|Lu_k - f\|_{L^2(\mu)} \overset{k\to\infty}{\longrightarrow} 0.$$

Proof. Recall that $\int(Lu)d\mu = 0$ for all $u \in C_c^\infty(\mathbb{R}^n)$. In order to show that the linear space $\{Lu \,;\, u \in C_c^\infty(\mathbb{R}^n)\}$ is dense, we analyze its orthogonal complement. Let $f \in L^2(\mu)$ be in the orthogonal complement, i.e., for any $u \in C_c^\infty(\mathbb{R}^n)$,

$$\int_{\mathbb{R}^n} f(Lu)d\mu = 0. \tag{35}$$

Our goal is to show that $f \equiv Const$. Note that (35) means that f is a weak solution of $Lf \equiv 0$. Since L is elliptic, then f is smooth and $Lf \equiv 0$ in the classical sense. Thus,

$$L(f^2) = 2fLf + 2|\nabla_M f|^2 = 2|\nabla_M f|^2.$$

11

Therefore, for any $\eta \in C_c^\infty(\mathbb{R}^n)$,

$$\int_{\mathbb{R}^n} |\nabla_M(\eta f)|^2 d\mu = \int_{\mathbb{R}^n} \left[\eta^2 |\nabla_M f|^2 + \frac{1}{2} \nabla_M(f^2) \cdot \nabla_M(\eta^2) + f^2 |\nabla_M \eta|^2 \right] d\mu$$

$$= \int_{\mathbb{R}^n} \left[\eta^2 |\nabla_M f|^2 - \frac{1}{2} \eta^2 L(f^2) + f^2 |\nabla_M \eta|^2 \right] d\mu = \int_{\mathbb{R}^n} |\nabla_M \eta|^2 f^2 d\mu.$$

However, according to our assumption (34), we have $|\nabla_M \eta|^2 = \Phi^{ij} \eta_i \eta_j \leq \varepsilon_0^{-1} |\nabla \eta|^2$. Let η_R be a smooth cutoff function in $\mathbb{R}^n$, that equals one on a Euclidean ball of radius R centered at the origin, that equals zero outside a Euclidean ball of radius $2R$, and that satisfies $|\nabla \eta_R| \leq 2/R$ throughout $\mathbb{R}^n$. Then,

$$\int_K |\nabla_M(\eta_R f)|^2 d\mu \leq \int_{\mathbb{R}^n} |\nabla_M \eta|^2 f^2 d\mu \leq \varepsilon_0^{-1} \int_{\mathbb{R}^n} |\nabla \eta_R|^2 f^2 d\mu \leq \frac{2}{R\varepsilon_0} \int_{\mathbb{R}^n} f^2 d\mu \xrightarrow{R \to \infty} 0,$$

since $f \in L^2(\mu)$. Therefore $\nabla f \equiv 0$ and f is constant. $\qquad\square$

Suppose that F is a locally-Lipschitz function on a Riemannian manifold such as $M_n^+(\mathbb{R})$. By the Rademacher theorem, the gradient ∇F is well-defined almost everywhere with respect to the Riemannian volume measure. In order to have a function $|\nabla F|$ that is defined everywhere, in this paper we set

$$|\nabla F|(x) = \limsup_{\substack{y \to x \\ z \to x}} \frac{|f(y) - f(z)|}{dist(y, z)} = \lim_{\varepsilon \to 0^+} \sup_{\substack{y, z \in B(x, \varepsilon) \\ y \neq z}} \frac{|f(y) - f(z)|}{dist(y, z)} \tag{36}$$

where $dist$ is the Riemannian distance, and $B(x, \varepsilon) = \{y \,;\, dist(x, y) < \varepsilon\}$. Since F is locally-Lipschitz, then the function $|\nabla F|$ is locally-bounded and upper semi-continuous. Clearly, at any point x where F is continuously differentiable, $|\nabla F|(x)$ equals the Riemannian length of $\nabla F(x)$.

Proposition 4.2. *Denote by θ the push-forward of the measure μ under the map (32). Then for any locally-Lipschitz function $F : M_n^+(\mathbb{R}) \to \mathbb{R}$ that belongs to $L^2(\theta)$ with $\int_{M_n^+(\mathbb{R})} F d\theta = 0$,*

$$\int_{M_n^+(\mathbb{R})} F^2 d\theta \leq 4 \int_{M_n^+(\mathbb{R})} |\nabla F|^2 d\theta,$$

whenever the right-hand side is finite.

Proof. Since F is locally-Lipschitz in $L^2(\theta)$, then the function f defined via

$$f(x) = F\left(D^2 \Phi(x)\right) \qquad (x \in \mathbb{R}^n),$$

is locally-Lipschitz in $\mathbb{R}^n$ and belongs to $L^2(\mu)$. Abbreviate $H = |\nabla F|^2$ and $h(x) = H\left(D^2 \Phi(x)\right)$. From the definition (36) of $|\nabla F|$, for any $x \in \mathbb{R}^n$ in which f is differentiable,

$$h(x) \geq \sup \left\{ \sum_{i=1}^n V^i f_i \,;\, \sum_{i,j=1}^n g_{ij} V^i V^j \leq 1, V^1, \ldots, V^n \in \mathbb{R} \right\}, \tag{37}$$

where f_i and g_{ij} are evaluated at the point x. In the case where the matrix $(g_{ij}(x))_{i,j=1,\ldots,n}$ is invertible, we may express the supremum in (37) in terms of the inverse matrix, yet it is the formula (37) which is valid in the general case. Setting $U_i = \Phi_{ij}V^j$, we reformulate (37) as

$$h(x) \geq \sup \left\{ \Phi^{ij}U_j f_i \, ; \, g_{ij}\Phi^{ki}\Phi^{\ell j}U_k U_\ell \leq 1, \, U_1, \ldots, U_n \in \mathbb{R} \right\}. \tag{38}$$

The formula (38) is valid for almost any $x \in \mathbb{R}^n$, since f is differentiable almost everywhere in $\mathbb{R}^n$ by the Rademacher theorem. We would like to show that for any $u \in C_c^\infty(\mathbb{R}^n)$,

$$-\int_{\mathbb{R}^n} f(Lu)d\mu \leq 2\sqrt{\int_{\mathbb{R}^n} h^2 d\mu} \cdot \sqrt{\int_{\mathbb{R}^n} (Lu)^2 d\mu}. \tag{39}$$

To that end, we observe that since u is compactly-supported,

$$\int_{\mathbb{R}^n} \Gamma_2(u)d\mu = \frac{1}{2}\int_{\mathbb{R}^n} L\left(\Phi^{ij}u_i u_j\right) d\mu - \int_{\mathbb{R}^n} \Phi^{ij}(Lu)_i u_j d\mu = -\int_{\mathbb{R}^n} \Phi^{ij}(Lu)_i u_j d\mu = \int_{\mathbb{R}^n} (Lu)^2 d\mu.$$

Therefore Lemma 3.1 and (33) imply that for any $u \in C_c^\infty(\mathbb{R}^n)$,

$$\int_{\mathbb{R}^n} (Lu)^2 d\mu \geq \frac{1}{4}\int_{\mathbb{R}^n} \Phi^{ik}\Phi^{j\ell}g_{k\ell}u_i u_j d\mu.$$

Since f is locally-Lipschitz, we may safely integrate by parts, and obtain that for any $u \in C_c^\infty(\mathbb{R}^n)$,

$$-\int_{\mathbb{R}^n} f(Lu)d\mu = \int_{\mathbb{R}^n} \Phi^{ij}f_i u_j d\mu \leq \int_{\mathbb{R}^n} h(x)\sqrt{g_{ij}\Phi^{ki}\Phi^{\ell j}u_k u_\ell}d\mu(x)$$

$$\leq \sqrt{\int_{\mathbb{R}^n} h^2 d\mu}\sqrt{\int_{\mathbb{R}^n} g_{ij}\Phi^{ki}\Phi^{\ell j}u_k u_\ell \, d\mu} \leq 2\sqrt{\int_{\mathbb{R}^n} h^2 d\mu}\sqrt{\int_{\mathbb{R}^n} (Lu)^2 d\mu}$$

and (39) is proven. Since $\int_{M_n^+(\mathbb{R})} F d\theta = 0$ then also $\int_{\mathbb{R}^n} f d\mu = 0$. From Lemma 4.1 there exists a sequence $u_k \in C_c^\infty(\mathbb{R}^n)$ with $Lu_k \to -f$ in $L^2(\mu)$. We plug in $u = u_k$ in (39), and take the limit $k \to \infty$. This yields

$$\int_{\mathbb{R}^n} f^2 d\mu \leq 2\sqrt{\int_{\mathbb{R}^n} h^2 d\mu} \cdot \sqrt{\int_{\mathbb{R}^n} f^2 d\mu}.$$

Hence,

$$\int_{\mathbb{R}^n} f^2 d\mu \leq 4\int_K h^2 d\mu.$$

Since $h(x) = H(D^2\Phi)$ with $H = |\nabla F|^2$, the proposition is proven. $\qquad \square$

5 Regularity issues

This section explains how to eliminate assumption (34) and also the smoothness assumptions of the previous two sections.

Theorem 5.1. *Assume that μ and ν are absolutely-continuous, log-concave probability measures on $\mathbb{R}^n$. Let $\nabla\Phi$ be the Brenier map between μ and ν, and assume condition $(\star)$ from Section 1. Denote by θ the push-forward of the measure μ under the map $x \mapsto D^2\Phi(x)$.*

Then for any θ-integrable, locally-Lipschitz function $F : M_n^+(\mathbb{R}) \to \mathbb{R}$,

$$\int_{M_n^+(\mathbb{R})} F^2 d\theta - \left(\int_{M_n^+(\mathbb{R})} F d\theta\right)^2 \leq 4\int_{M_n^+(\mathbb{R})} |\nabla F|^2 d\theta, \tag{40}$$

whenever the right-hand side is finite, and $|\nabla F|$ is interpreted as in (36).

The strategy for proving Theorem 5.1 is to approximate Φ by a sequence of functions Φ_N that satisfy assumption (34), and to prove the pointwise (even local uniform) convergence $D^2\Phi_N(x) \overset{N\to\infty}{\Longrightarrow} D^2\Phi(x)$. We discuss below two possible justifications of this convergence, because we believe that both of them may be useful. The first proof occupies Subsection 5.1, and it is based on various results from the regularity theory of the Monge-Ampère equation. The log-concavity of the measures is not really required for the first proof, and it suffices to assume that the densities are locally Hölder.

The second proof in Subsection 5.2 is in fact an alternative approach to Caffareli's $C^{1,\alpha}$-regularity results in the log-concave case. The argument in Subsection 5.2 is more self-contained, and it is based on integration-by-parts arguments. The log-concavity of the target measure plays an important role here, and we further assume a certain integrability condition on the logarithmic derivative of the density of μ. This integrability condition is rather mild in our opinion, and it is satisfied in many cases of interest.

5.1 First proof of Theorem 5.1

As before, we write e^{-V} and e^{-W} for the densities of μ and ν, respectively. By log-concavity, the functions V and W are locally-Lipschitz in the open sets $Supp(\mu)$ and $Supp(\nu)$, respectively. From condition $(\star)$ the function Φ is C^2-smooth, and the push-forward equation (1) implies that

$$\det D^2\Phi(x) = e^{-V(x)+W(\nabla\Phi(x))} \tag{41}$$

for any $x \in Supp(\mu)$. In particular, $D^2\Phi(x)$ is invertible, and hence positive-definite for all $x \in Supp(\mu)$. Thus Φ is strictly-convex. The modulus of convexity of Φ at the point x is defined to be

$$\omega_\Phi(x;\delta) = \inf\left\{\Phi(y) - (\Phi(x) + \nabla\Phi(x)\cdot(y-x)) \; ; \; y \in \mathbb{R}^n, \, |y-x| = \delta\right\}.$$

Then $\omega_\Phi(x;\delta)$ is a positive, continuous function of $x \in Supp(\mu)$ and $\delta > 0$, when we restrict attention to x and δ for which $\overline{B(x,\delta)} \subseteq Supp(\mu)$. Here, $B(x,\delta) = \{y \in \mathbb{R}^n \; ; \; |y-x| < \delta\}$. Next, the Legendre transform

$$\Phi^*(x) = \sup_{\substack{y\in\mathbb{R}^n \\ \Phi(y)<\infty}} \left[x\cdot y - \Phi(y)\right]$$

is also C^2-smooth and strictly-convex in $Supp(\nu)$, with $y \mapsto \nabla\Phi^*(y)$ being the inverse map to $x \mapsto \nabla\Phi(x)$. Thus $\nabla\Phi$ is a C^1-diffeomorphism of $Supp(\mu)$ and $Supp(\nu)$. The reader is referred to Rockafellar [22] for the basic properties of the Legendre transform.

We will approximate μ and ν by sequences of probability measures μ_N and ν_N with the following properties:

(i) The probability measure μ_N (respectively ν_N) has a density in $\mathbb{R}^n$ of the form e^{-V_N} (respectively e^{-W_N}).

(ii) The functions $V_N, W_N : \mathbb{R}^n \to \mathbb{R}$ are smooth and for any $x \in \mathbb{R}^n$,

$$D^2 V_N(x) \geq \frac{1}{N} \cdot \mathrm{Id}, \qquad D^2 W_N(x) \leq N \cdot \mathrm{Id}.$$

(iii) $V_N \longrightarrow V$ locally uniformly in $Supp(\mu)$, and similarly, $W_N \longrightarrow W$ locally uniformly in $Supp(\nu)$.

It is quite standard to approximate μ and ν in this manner. For instance, in order to obtain μ_N (or ν_N), we may convolve μ (or ν) with a Gaussian of a tiny variance, then multiply the resulting density by a Gaussian of a huge variance, and then normalize to obtain a probability density. Denote by $\nabla \Phi_N$ the Brenier map between μ_N and ν_N. We use again Caffarelli's regularity theory, to conclude that $\Phi_N : \mathbb{R}^n \to \mathbb{R}$ is a smooth, strictly-convex function, with

$$\det D^2 \Phi_N(x) = e^{-V_N(x) + W_N(\nabla \Phi_N(x))} \qquad\qquad (x \in \mathbb{R}^n). \qquad (42)$$

The following lemma should be known to experts on the Monge-Ampère equation, yet we could not find it in the literature.

Lemma 5.2. *There exists an increasing sequence $\{N_j\}$ such that*

$$D^2 \Phi_{N_j}(x) \overset{j \to \infty}{\longrightarrow} D^2 \Phi(x)$$

locally uniformly in $x \in Supp(\mu)$.

Proof. Fix $x_0 \in Supp(\mu)$. It suffices to find $\{N_j\}$ such that $D^2 \Phi_{N_j} \longrightarrow D^2 \Phi$ uniformly in a neighborhood of x_0. A standard convexity argument (e.g., [15, Section 2]) based on (iii) and the fact that $\int e^{-V} = \int e^{-W} = 1$ shows that there exist $A, B > 0$ with

$$\min \left\{ \inf_N V_N(x), \inf_N W_N(x), V(x), W(x) \right\} \geq A|x| - B, \qquad\qquad (x \in \mathbb{R}^n). \qquad (43)$$

Therefore,

$$\sup_N \int_{\mathbb{R}^n} |\nabla \Phi_N|^2 e^{-V_N(x)} dx = \sup_N \int_{\mathbb{R}^n} |x|^2 e^{-W_N(x)} dx \leq \int_{\mathbb{R}^n} |x|^2 e^{B - A|x|} dx < \infty. \quad (44)$$

Recall that $V_N \to V$ locally uniformly in $Supp(\mu)$, according to (iii). From (44) we learn that $\sup_N \|\Phi_N\|_{\dot{H}^1(K)} < \infty$ for any compact $K \subset Supp(\mu)$. Here,

$$\|u\|_{\dot{H}^1(K)}^2 = \int_K |\nabla u(x)|^2 dx.$$

From the Rellich-Kondrachov compactness theorem (e.g., [8, Section 4.6]), we conclude that there exists a subsequence Φ_{N_j}, numbers $C_j \in \mathbb{R}$ and a certain function $F : Supp(\mu) \to \mathbb{R}$ such that for any compact $K \subset Supp(\mu)$, the sequence $\Phi_{N_j} + C_j$ converges to F in $L^2(K)$.

Passing to another subsequence, which we conveniently denote again by $\{\Phi_N\}$, and using [22, Theorem 10.9], we may assume that F is convex and that the convergence is locally-uniform in $Supp(\mu)$. Thus, from [22, Theorem 24.5],

$$\nabla\Phi_N(x) \overset{N\to\infty}{\longrightarrow} \nabla F(x) \tag{45}$$

for almost any $x \in Supp(\mu)$. However, $(\nabla\Phi_N)_*\mu_N = \nu_N$. From (iii), (43) and (45) we conclude that $(\nabla F)_*\mu = \nu$. From the uniqueness of the Brenier map, we deduce that $\nabla F = \nabla\Phi$ almost everywhere in $Supp(\mu)$. Since Φ is C^2-smooth, then we may apply [22, Theorem 25.7], and upgrade (45) to

$$\nabla\Phi_N(x) \overset{N\to\infty}{\longrightarrow} \nabla\Phi(x) \tag{46}$$

locally uniformly in $Supp(\mu)$. The convexity arguments in [22, Section 25] also show that $\nabla\Phi_N^* \to \nabla\Phi^*$ locally uniformly in $Supp(\nu)$. As for the modulus of convexity, we have

$$\omega_{\Phi_N}(x;\delta) \overset{N\to\infty}{\longrightarrow} \omega_\Phi(x;\delta), \qquad \text{and respectively,} \qquad \omega_{\Phi_N^*}(y;\delta) \overset{N\to\infty}{\longrightarrow} \omega_{\Phi^*}(y;\delta) \tag{47}$$

locally uniformly in the set $\{(x,\delta) \in Supp(\mu) \times (0,\infty)\,;\, \overline{B(x,\delta)} \subset Supp(\mu)\}$, and respectively, in the set $\{(y,\delta) \in Supp(\nu) \times (0,\infty)\,;\, \overline{B(y,\delta)} \subset Supp(\nu)\}$.

We will now invoke the estimates of Gutierrez and Huang [12] and Forzani and Maldonado [9, 10], which are constructive versions of Caffarelli's $C^{1,\alpha}$-regularity theory. We are allowed to apply [12, Theorem 2.1] and [9, Theorem 15] locally near x_0, thanks to (iii), (42), (46) and (47). From [9, Theorem 15] we learn that there exist $\alpha, \delta, C > 0$ such that for any $x, y \in B(x_0, \delta)$ and $N \geq 1$,

$$|\nabla\Phi_N(x) - \nabla\Phi_N(y)| \leq C|x-y|^\alpha. \tag{48}$$

The function V is locally-Lipschitz. From (iii) and [22, Theorem 24.5], the sequence $\{V_N\}$ is uniformly locally-Lipschitz: This means that for any compact subset $K \subset Supp(\mu)$, the Lipschitz constant of V_N is bounded by some finite number C_K, independent of N. Similarly, the sequence $\{W_N\}$ is also uniformly locally-Lipschitz. Together with (46) and (48) we deduce that there exist $\hat{C} > 0$ such that $u_N(x) = -V_N(x) + W_N(\nabla\Phi_N(x))$ satisfies

$$|u_N(x) - u_N(y)| \leq \hat{C}|x-y|^\alpha \qquad\qquad (x, y \in B(x_0, \delta), N \geq 1).$$

Recalling the Monge-Ampère equation (42), we learn that that there exists $\tilde{C} > 0$ such that

$$\left|\det D^2\Phi_N(x) - \det D^2\Phi_N(y)\right| \leq \tilde{C}|x-y|^\alpha. \qquad\qquad (x, y \in B(x_0, \delta), N \geq 1).$$

We are finally in good shape for applying the $C^{2,\alpha}$-estimates from Trudinger and Wang [24, Theorem 3.2]. These estimates yield the existence of $\bar{C} > 0$ such that for any $x, y \in B(x_0, \delta/2)$ and $N \geq 1$,

$$\|D^2\Phi_N(x) - D^2\Phi_N(y)\|_{HS} \leq \bar{C}|x-y|^\alpha. \tag{49}$$

The uniform $C^{2,\alpha}$-estimate in (49) allows us to apply the Arzella-Ascoli theorem. All we need is to denote $K = B(x_0, \delta/2)$ and observe that

$$\int_K (\Delta\Phi_N)\xi = -\int_K \nabla\Phi_N \cdot \nabla\xi \overset{N\to\infty}{\longrightarrow} -\int_K \nabla\Phi \cdot \nabla\xi = \int_K (\Delta\Phi)\xi,$$

16

where ξ is any smooth, compactly-supported function in K. Hence the sequence $\{\int_K \Delta\Phi_N\}_{N\geq 1}$ is bounded, and since $D^2\Phi_N$ is positive-definite, also the sequence $\{\int_K \|D^2\Phi_N\|_{HS}\}_{N\geq 1}$ is bounded. From (49) and the Arzella-Ascoli theorem, there exists a subsequence, denoted still by $\{\Phi_N\}$, such that $D^2\Phi_N \longrightarrow D^2\Phi$ uniformly on $K = B(x_0, \delta/2)$. $\qquad\square$

Remark 5.3. Our proof of Lemma 5.2 does not make any use of the log-concavity of μ and ν. By inspecting the proof above, we see that Lemma 5.2 holds true as long as V and W are locally Hölder, and V_N, W_N are uniformly locally Hölder.

In order to simplify the notation, we denote the sequence $\{\Phi_{N_j}\}$ from Lemma 5.2 by $\{\Phi_N\}$. Properties (i), (ii) and (iii) above are still satisfied.

Corollary 5.4. *Denote by θ_N the push-forward of the measure μ_N under the map $x \mapsto D^2\Phi_N(x)$. Then for any bounded, continuous function $b : M_n^+(\mathbb{R}) \to \mathbb{R}$,*

$$\int_{M_n^+(\mathbb{R})} b\, d\theta_N \overset{N\to\infty}{\longrightarrow} \int_{M_n^+(\mathbb{R})} b\, d\theta. \tag{50}$$

Furthermore, if $b : M_n^+(\mathbb{R}) \to \mathbb{R}$ is bounded and upper semi-continuous, then

$$\limsup_{N\to\infty} \int_{M_n^+(\mathbb{R})} b\, d\theta_N \leq \int_{M_n^+(\mathbb{R})} b\, d\theta. \tag{51}$$

Proof. In order to prove (50), we need to show that

$$\int_{\mathbb{R}^n} b\left(D^2\Phi_N(x)\right) e^{-V_N(x)} dx \overset{N\to\infty}{\longrightarrow} \int_{\mathbb{R}^n} b\left(D^2\Phi(x)\right) e^{-V(x)} dx.$$

This follows from Lemma 5.2 and the dominated convergence theorem, since (43) provides an integrable majorant. Next, assume that b is bounded and upper semi-continuous. Then for any $x \in Supp(\mu)$,

$$\limsup_{N\to\infty} b\left(D^2\Phi_N(x)\right) e^{-V_N(x)} \leq b(D^2\Phi(x)) e^{-V(x)}.$$

Now (51) follows from Fatou's lemma, since we have an integrable majorant by (43). $\qquad\square$

Proof of Theorem 5.1. Assume first that the locally-Lipschitz function F is compactly supported. We observe that for any fixed N, assumption (34) holds true. Indeed, we may apply a refinement of Caffarelli's contraction theorem [7] which appears in [18], and obtain from (ii) that for any $x \in \mathbb{R}^n$,

$$D^2\Phi_N(x) \geq \frac{1}{N^2} \cdot \mathrm{Id}.$$

We may therefore apply Proposition 4.2, and conclude that for any $N \geq 1$,

$$\int_{M_n^+(\mathbb{R})} F^2 d\theta_N - \left(\int_{M_n^+(\mathbb{R})} F d\theta_N\right)^2 \leq 4\int_{M_n^+(\mathbb{R})} |\nabla F|^2 d\theta_N.$$

Recall that $|\nabla F|^2$ is upper semi-continuous and bounded, while F is continuous and bounded. By taking the limit as $N \to \infty$ and using Corollary 5.4, we obtain that

$$\int_{M_n^+(\mathbb{R})} F^2 d\theta - \left(\int_{M_n^+(\mathbb{R})} F d\theta \right)^2 \leq 4 \int_{M_n^+(\mathbb{R})} |\nabla F|^2 d\theta,$$

and (40) is proven in the case where F is a compactly-supported function.

The next step is to prove (40) under the additional assumption that $F \in L^2(\theta)$. To that end we pick a smooth function $\theta_R : M_n^+(\mathbb{R}) \to [0, 1]$, such that θ_R equals one on $B(\mathrm{Id}, R)$ and it vanishes outside $B(\mathrm{Id}, 2R)$, with $|\nabla \theta_R| \leq 2/R$. Set $F_R = \theta_R F$. We have just proven that (40) holds true when F is replaced by F_R. Clearly, $F_R \longrightarrow F$ in $L^2(\theta)$ as $R \longrightarrow \infty$. All that remains is to show that

$$\limsup_{R \to \infty} \int_{M_n^+(\mathbb{R})} |\nabla F_R|^2 d\theta \leq \int_{M_n^+(\mathbb{R})} |\nabla F|^2 d\theta. \tag{52}$$

The functions θ_R and F are continuous, and therefore we may use the Leibnitz rule

$$|\nabla F_R| \leq |F| |\nabla \theta_R| + \theta_R |\nabla F| \leq |\nabla F| + 2|F|/R,$$

where we interpret $|\nabla F|$ and $|\nabla F_R|$ in the sense of definition (36). Since $F, |\nabla F| \in L^2(\theta)$, then (52) follows in the case where $F \in L^2(\theta)$.

Finally, in order to eliminate the assumption that $F \in L^2(\theta)$, we replace F by $F_R = \max\{-R, \min\{F, R\}\}$, apply the inequality for F_R, and let R tend to infinity. For all but countably many values of R, the level set $\{A \in M_n^+(\mathbb{R}) \, ; \, F(A) = R\}$ has zero θ-measure. Consequently, we have the inequality $\int |\nabla F_R|^2 d\theta \leq \int |\nabla F|^2 d\theta$ for all but countably many values of R, and (40) follows. $\qquad \square$

5.2 Second proof: Log-concave target measure

In our second proof we will exploit the fact that ν is log-concave, but we will not require the log-concavity of μ. Throughout this subsection we make the following additional assumption:

Assumption (A): For some $p > n$,

$$\int_{\mathbb{R}^n} |\nabla V|^p e^{-V} \, dx < \infty,$$

where the derivatives V_i are understood in the logarithmic derivative sense, i.e.

$$\int_{\mathbb{R}^n} \xi V_i d\mu = - \int_{\mathbb{R}^n} \xi_i d\mu, \quad \xi \in C_c^\infty(\mathbb{R}^n), i = 1, \ldots, n.$$

By the Morrey embedding theorem (see, e.g., [8, Section 4.5]), the function V is locally Hölder. We will approximate μ and ν by sequences of probability measures μ_N and ν_N having properties (i), (ii) and (iii) from Subsection 5.1. We also require a fourth property:

18

(iv) There exists $p > n$ such that

$$\sup_N \int_{\mathbb{R}^n} |\nabla V_N|^p e^{-V_N}\, dx < \infty.$$

The approach outlined in Subsection 5.1, to convolve with a tiny Gaussian and then multiply by the density of a huge Gaussian, yields also property (iv). Recall that the Brenier map $\nabla \Phi_N$ between μ_N and ν_N is smooth and that it satisfies (42). The central ingredient of this subsection is the following a priori estimate:

Proposition 5.5. *Assume that functions V, W and Φ are smooth on the entire $\mathbb{R}^n$ and that ν is a log-concave measure. Then for every $q \geq 2$, $0 < \tau < 1$, $i = 1, \ldots, n$ there exists $C(q, \tau) > 0$*

$$\int_{\mathbb{R}^n} \Phi_{ii}^q d\mu \leq C(q, \tau)\left(\int_{\mathbb{R}^n} |V_i|^{\frac{2q}{2-\tau}}\, d\mu + \int_{\mathbb{R}^n} |x_i|^{\frac{2q}{\tau}}\, d\nu\right). \tag{53}$$

Proof. Assume in addition that $D^2 W \geq \frac{1}{C} \cdot \mathrm{Id}$, $D^2 V \leq C \cdot \mathrm{Id}$. In this case $D^2 \Phi \leq C^2 \cdot \mathrm{Id}$. Recall formula (28),

$$L(\Phi_{ii}) - \Phi_i^{jk}\Phi_{ijk} - \sum_{j,k=1}^n \Phi_{ji}\Phi_{ik}W_{jk} \circ \nabla\Phi = -V_{ii},$$

which is obtained by differentiating the change of variables formula (22) along x_i. Let us multiply this formula by Φ_{ii}^p, $p \geq 0$ and make a formal integration by parts with respect to μ. Using the convexity of W we get

$$\int V_{ii}\Phi_{ii}^p\, d\mu \geq p\int \Phi_{ii}^{p-1}\langle (D^2\Phi)^{-1}\nabla\Phi_{ii}, \nabla\Phi_{ii}\rangle d\mu + \int \Phi_{ii}^p\Phi_i^{jk}\Phi_{ijk}d\mu. \tag{54}$$

Let us justify this formula. To this end we fix a compactly supported function $\eta \geq 0$ and integrate with respect to $\eta \cdot \mu$.

$$\int V_{ii}\Phi_{ii}^p\eta\, d\mu \geq \int \langle (D^2\Phi)^{-1}\nabla\eta, \nabla\Phi_{ii}\rangle\Phi_{ii}^p d\mu + p\int \Phi_{ii}^{p-1}\langle (D^2\Phi)^{-1}\nabla\Phi_{ii}, \nabla\Phi_{ii}\rangle\eta d\mu + \int \Phi_{ii}^p\Phi_i^{jk}\Phi_{ijk}\eta d\mu.$$

Applying the Cauchy inequality we get

$$-\int \langle (D^2\Phi)^{-1}\nabla\eta, \nabla\Phi_{ii}\rangle\Phi_{ii}^p d\mu \leq \frac{4}{\varepsilon}\int \frac{\langle (D^2\Phi)^{-1}\nabla\eta, \nabla\eta\rangle}{\eta}\Phi_{ii}^{p+1}d\mu + \varepsilon\int \langle (D^2\Phi)^{-1}\nabla\Phi_{ii}, \nabla\Phi_{ii}\rangle\Phi_{ii}^{p-1}\eta d\mu.$$

Finally,

$$\int V_{ii}\Phi_{ii}^p\eta\, d\mu + \frac{4}{\varepsilon}\int \frac{\langle (D^2\Phi)^{-1}\nabla\eta, \nabla\eta\rangle}{\eta}\Phi_{ii}^{p+1}d\mu$$
$$\geq (p-\varepsilon)\int \Phi_{ii}^{p-1}\langle (D^2\Phi)^{-1}\nabla\Phi_{ii}, \nabla\Phi_{ii}\rangle\eta d\mu + \int \Phi_{ii}^p\Phi_i^{jk}\Phi_{ijk}\eta d\mu.$$

Assume that η has the form $\eta = \xi(\nabla\Phi)$, where ξ is compactly supported. We get

$$\int V_{ii}\Phi_{ii}^p\eta\, d\mu + \frac{4C^{p+2}}{\varepsilon}\int \frac{|\nabla\xi|^2}{\xi}\, d\nu \geq (p-\varepsilon)\int \Phi_{ii}^{p-1}\langle (D^2\Phi)^{-1}\nabla\Phi_{ii}, \nabla\Phi_{ii}\rangle\eta d\mu + \int \Phi_{ii}^p\Phi_i^{jk}\Phi_{ijk}\eta d\mu.$$

19

It remains to construct a sequence of functions $1 \geq \xi_N \geq 0$ satisfying $\lim_N \xi_N(x) = 1$ for ν-a.e. x and $\lim_N \int |\nabla \xi_N|^2 / \xi_N \, d\nu = 0$. Then applying the Fatou lemma we justify (54).

It is helpful to have in mind that $\Phi_i^{jk} \Phi_{ijk} = \text{Tr} \left[(D^2\Phi)^{-\frac{1}{2}} D^2 \Phi_i (D^2\Phi)^{-\frac{1}{2}} \right]^2 \geq 0$. From (54),

$$\int V_{ii} \Phi_{ii}^p \, d\mu \geq p \int \Phi_{ii}^{p-1} \langle (D^2\Phi)^{-1} \nabla \Phi_{ii}, \nabla \Phi_{ii} \rangle d\mu.$$

Let us integrate by parts the left-hand side $\int V_{ii} \Phi_{ii}^p \, d\mu = \int V_i^2 \Phi_{ii}^p \, d\mu - p \int V_i \Phi_{ii}^{p-1} \Phi_{iii} \, d\mu$. The justification of this integration by parts is much easier, since $D^2\Phi$ and D^2V are bounded. Applying

$$2|\Phi_{iii} V_i| \leq 2|V_i| \sqrt{\Phi_{ii} \cdot \langle (D^2\Phi)^{-1} \nabla \Phi_{ii}, \nabla \Phi_{ii} \rangle} \leq V_i^2 \Phi_{ii} + \langle (D^2\Phi)^{-1} \nabla \Phi_{ii}, \nabla \Phi_{ii} \rangle$$

one obtains

$$\int V_i^2 \Phi_{ii}^p \, d\mu \geq \int \Phi_{ii}^{p-1} \langle (D^2\Phi)^{-1} \nabla \Phi_{ii}, \nabla \Phi_{ii} \rangle d\mu. \tag{55}$$

Let us show that the right-hand side controls powers of the second derivative Φ_{ii}. Indeed, for every $q \geq 2$ and $\varepsilon > 0, 0 \leq \tau \leq 1$ the following estimate holds

$$\int \Phi_{ii}^q \, d\mu = -(q-1) \int \Phi_i \Phi_{iii} \Phi_{ii}^{q-2} \, d\mu + \int \Phi_i V_i \Phi_{ii}^{q-1} \, d\mu$$

$$\leq \varepsilon \int \Phi_i^2 \Phi_{ii}^{q-\tau} \, d\mu + \frac{(q-1)^2}{4\varepsilon} \int \Phi_{ii}^{q-3+\tau} \langle (D^2\Phi)^{-1} \nabla \Phi_{ii}, \nabla \Phi_{ii} \rangle d\mu$$

$$+ \frac{q-1}{q} \int \Phi_{ii}^q \, d\mu + \frac{1}{q} \int |\Phi_i V_i|^q \, d\mu.$$

Finally,

$$\int \Phi_{ii}^q \, d\mu \leq \int |\Phi_i V_i|^q \, d\mu + q\varepsilon \int \Phi_i^2 \Phi_{ii}^{q-\tau} \, d\mu + \frac{q(q-1)^2}{4\varepsilon} \int \Phi_{ii}^{q-3+\tau} \langle (D^2\Phi)^{-1} \nabla \Phi_{ii}, \nabla \Phi_{ii} \rangle d\mu$$

$$\leq \int |\Phi_i V_i|^q \, d\mu + q\varepsilon \int \Phi_i^2 \Phi_{ii}^{q-\tau} \, d\mu + \frac{q(q-1)^2}{4\varepsilon} \int \Phi_{ii}^{q-2+\tau} V_i^2 d\mu.$$

Applying Hölder inequalities

$$\Phi_i^2 \Phi_{ii}^{q-\tau} \leq \frac{q-\tau}{q} \Phi_{ii}^q + \frac{\tau}{q} |\Phi_i|^{\frac{2q}{\tau}},$$

$$\Phi_{ii}^{q-2+\tau} V_i^2 \leq \varepsilon \Phi_{ii}^q + C(\varepsilon, q, \tau) |V_i|^{\frac{2q}{2-\tau}},$$

$$|\Phi_i V_i|^q \leq \frac{2-\tau}{2} |V_i|^{\frac{2q}{2-\tau}} + \frac{\tau}{2} |\Phi_i|^{\frac{2q}{\tau}},$$

choosing sufficiently small ε, and applying the change of variables formula $\int |\Phi_i|^q d\mu = \int |x_i|^q d\nu$ we easily get the claim.

Finally, let us get rid of the assumption $D^2W \geq \frac{1}{C} \cdot \text{Id}$, $D^2V \leq C \cdot \text{Id}$. To this end we approximate μ and ν by measures with smooth potentials satisfying $D^2W_N \geq \frac{1}{C_N} \cdot \text{Id}$, $D^2V_N \leq C_N \cdot \text{Id}$ satisfying $\lim_N \int |(V_N)_i|^{2q} \, d\mu_N = \int |V_i|^{2q} \, d\mu$ and $\lim_N \int |x_i|^{2q} \, d\nu_N =$

$\int |x_i|^{2q} \, d\nu$. It remains to show that the weak $L^q(\mu)$-limit of $(\Phi_N)_{ii}$ coincides with Φ_{ii}. The latter can be easily shown with the help of integration-by-parts and identifications of the poinwise limit $\lim_N \nabla \Phi_N$ with $\nabla \Phi$ (see the proof of Lemma 5.2). $\qquad \square$

Remark 5.6. The conclusion of Proposition 5.5 holds without any additional smoothness assumptions. This can be verified by smooth approximations (see again [16] for details). Finally we get that (53) holds for every log-concave measure ν and measure μ satisfying $\int |V_i|^{\frac{2q}{2-\tau}} \, d\mu < \infty$, where V_i is the logarithmic derivative of μ along x_i.

Second proof of Lemma 5.2: Let us show how Proposition 5.5 implies (48) above, without appealing to the works by Forzani and Maldonado [9, 10] and Gutierrez and Huang [12] related to Caffarelli's $C^{1,\alpha}$-regularity theory. We use that $\sup_N \int |\nabla V_N|^p e^{-V_N} \, dx < \infty$, $p > n$. Since ν is log-concave, all the moments of ν are finite. Thus Proposition 5.5 implies

$$\sup_N \int \|D^2 \Phi_N\|_{HS}^{p'} e^{-V_N} \, dx < \infty$$

for any $n < p' < p$. Applying that V_N are uniformly locally bounded from below, we get that $\sup_N \int_{B_R} \|D^2 \Phi_N\|_{HS}^{p'} \, dx < \infty$ for every R. Then the result follows from the Morrey embedding theorem. $\qquad \square$

6 Corollaries to Theorem 5.1

Proof of Theorem 1.2. For $A \in M_n^+(\mathbb{R})$ define

$$F(A) = f(\log \lambda_1(A), \ldots, \log \lambda_n(A))$$

where $0 < \lambda_1(A) \leq \ldots \leq \lambda_n(A)$ are the eigenvalues of A. According to Lemma 2.4, for any $A \in M_n^+(\mathbb{R})$,

$$|\nabla F|(A) \leq |\nabla f| (\log \lambda_1(A), \ldots, \log \lambda_n(A)). \tag{56}$$

Since f is locally-Lipschitz and the eigenvalues vary continuously with the matrix A, then (56) implies that also F is locally-Lipschitz. Denote by θ the push-forward of the probability measure μ under the map $x \mapsto D^2 \Phi(x)$. Since $\mathbb{E}|f(\Lambda(X))| < \infty$ then $F \in L^1(\theta)$. Since $\mathbb{E}|\nabla f|^2(\Lambda(X)) < \infty$, then $\int |\nabla F|^2 d\theta < \infty$. We may apply Theorem 5.1 and conclude that

$$\int_{M_n^+(\mathbb{R})} F^2 d\theta - \left(\int_{M_n^+(\mathbb{R})} F d\theta \right)^2 \leq 4 \int_{M_n^+(\mathbb{R})} |\nabla F|^2 d\theta.$$

The left-hand side equals $Var[f(\Lambda(X))]$. Glancing at (56), we thus obtain

$$Var[f(\Lambda(X))] \leq 4\mathbb{E}|\nabla f|^2(\Lambda(X)),$$

and the proof is complete. $\qquad \square$

Proof of Theorem 1.1. Plug in $f(x) = x_i$ in Theorem 1.2. Then f is a 1-Lipschitz function, by Remark 1.4 we have $\mathbb{E}\left|f(\Lambda(X))\right| < \infty$. Thus the application of Theorem 1.2 is legitimate, and Theorem 1.1 follows. $\qquad\square$

Proof of Theorem 1.5. The argument is almost identical to the proof of Theorem 1.1, with Lemma 2.3 replacing the role of Lemma 2.4. $\qquad\square$

Let us end this paper with a few remarks concerning future research. If we make further assumptions regarding the log-concave measures in question, it is possible to prove concentration inequalities for the eigenvalues of $D^2\Phi$ themselves, and not only for their logarithms. The analysis of the weighted Riemannian manifold $M_{\mu,\nu}$ leads to such concentration inequalities. Additionally, there is a soft argument which shows that when $\nabla\Phi$ is the Brenier map between the uniform measure on K and the uniform measure on T,

$$\fint_K \Delta\Phi \le nV(K, \dots, K, T),$$

where V stands for mixed volume. The details will be discussed elsewhere. Another possible research direction is to investigate whether phenomena similar to Theorem 1.1 occur also in a non-linear setting, when transporting measures with convexity properties supported on Riemannian manifolds.

References

[1] Alesker, S., Dar, S., Milman, V., *A remarkable measure preserving diffeomorphism between two convex bodies in $\mathbb{R}^n$.* Geom. Dedicata, Vol. 74, No. 2, (1999), 201-212.

[2] Bakry, D., Émery, M., *Diffusions hypercontractives.* Séminaire de probabilités, XIX, 1983/84, Springer Lecture Notes in Math., Vol. 1123, (1985), 177–206.

[3] Bakry, D., Gentil, I., Ledoux, M., *Analysis and Geometry of Markov Diffusion Operators.* A series of comprehensive studies in Math., Vol. 348, Springer, 2014.

[4] Brenier, Y., *Polar factorization and monotone rearrangement of vector-valued functions.* Comm. Pure Appl. Math. 44, No. 4, (1991), 375-417.

[5] Caffarelli, L., *The regularity of mappings with a convex potential.* J. Amer. Math. Soc., Vol. 5, (1992), 99–104.

[6] Caffarelli, L., *A localization property of viscosity solutions to the Monge-Ampère equation and their strict convexity.* Ann. of Math. (2), Vol. 131, No. 1, (1990), 129-134.

[7] Caffarelli, L., *Monotonicity properties of optimal transportation and the FKG and related inequalities.* Comm. Math. Phys., Vol. 214, No. 3, (2000), 547-563. Erratum in Comm. Math. Phys., Vol. 225, No. 2, (2002), 449-450.

[8] Evans, L. C., Gariepy, R. F., *Measure theory and fine properties of functions.* Studies in Advanced Mathematics. CRC Press, Boca Raton, FL, 1992.

[9] Forzani, L., Maldonado, D., *Properties of the solutions to the Monge-Ampère equation.* Nonlinear Anal., Vol. 57, No. 5-6, (2004), 815-829.

[10] Forzani, L., Maldonado, D., *Recent progress on the Monge-Ampère equation.* The p-harmonic equation and recent advances in analysis, Contemp. Math., 370, Amer. Math. Soc., Providence, RI, (2005), 189-198.

[11] Grigor'yan, A., *Heat kernel and analysis on manifolds.* AMS/IP Studies in Advanced Mathematics, Vol. 47, Amer. Math. Soc., Providence, RI; International Press, Boston, MA, 2009.

[12] Gutierrez, C. E., Huang, Q., *Geometric properties of the sections of solutions to the Monge-Ampère equation.* Trans. Amer. Math. Soc., Vol. 352, No. 9, (2000), 4381-4396.

[13] Gromov, M., Milman, V. D., *A topological application of the isoperimetric inequality.* Amer. J. Math., Vol. 105, No. 4, (1983), 843-854.

[14] Klartag, B., *Poincaré inequalities and moment maps.* Ann. Fac. Sci. Toulouse Math., Vol. 22, No. 1, (2013), 1–41.

[15] Klartag, B., *Logarithmically-concave moment measures I.* Preprint, arXiv:1309.2767

[16] Kolesnikov, A. V., *Sobolev regularity of transportation of probability measures and transportation inequalities.* Teor. Veroyatnost. i Primenen., Vol. 57, No. 2, (2012), 296–321.

[17] Kolesnikov, A. V., *Hessian metrics, CD(K,N)-spaces, and optimal transportation of log-concave measures.* Discrete and Continuous Dynamical Systems - Series A., Vol. 34, No. 4, (2014), 1511–1532.

[18] Kolesnikov, A. V., *Mass transportation and contractions.* MIPT Proc. (2010), 2(4), 90–99. arXiv: 1103.1479

[19] McCann, R., *Existence and uniqueness of monotone measure-preserving maps.* Duke Math. J., Vol. 80, No. 2, (1995), 309-323.

[20] Polya, G., *Remark on Weyl's note "Inequalities between the two kinds of eigenvalues of a linear transformation."* Proc. Nat. Acad. Sci. U. S. A., Vol. 36, (1950), 49-51.

[21] Reed, M., Simon, B., *Methods of Modern Mathematical Physics IV: Analysis of Operators.* Academic Press, New York-London, 1978.

[22] Rockafellar, R. T., *Convex analysis.* Princeton Mathematical Series, No. 28, Princeton University Press, Princeton, NJ, 1970.

[23] Strichartz, R. S., *Analysis of the Laplacian on the complete Riemannian manifold.* J. Funct. Anal., Vol. 52, No. 1, (1983), 48-79.

[24] Trudinger, N. S., Wang, X.-J., *The Monge-Ampère equation and its geometric applications.* Handbook of geometric analysis. Adv. Lect. Math. (ALM), Vol. 7, Int. Press, Somerville, MA, (2008), 467-524.

[25] Weyl, H., *Inequalities between the two kinds of eigenvalues of a linear transformation.* Proc. Nat. Acad. Sci. U. S. A., Vol. 35, (1949), 408-411.