

# ANALYSIS & PDE

Volume 8

No. 1

2015

BO'AZ B. KLARTAG AND ALEXANDER V. KOLESNIKOV

**EIGENVALUE DISTRIBUTION OF OPTIMAL TRANSPORTATION**

# EIGENVALUE DISTRIBUTION OF OPTIMAL TRANSPORTATION

BO'AZ B. KLARTAG AND ALEXANDER V. KOLESNIKOV

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  $\mathbb{R}^n$ . We show that the eigenvalues of the Hessian matrix  $D^2\Phi$  exhibit concentration properties on a multiplicative scale, regardless of the choice of the two measures or the dimension  $n$ .

## 1. Introduction

Let  $\mu$  and  $\nu$  be two absolutely continuous probability measures on  $\mathbb{R}^n$ . It was discovered by Brenier [1991] and McCann [1995] that there exists a convex function  $\Phi$  on  $\mathbb{R}^n$  with  $(\nabla\Phi)_*\mu = \nu$ , i.e.,

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

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

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

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

( $\star$ ) The function  $\Phi$  is  $C^2$ -smooth in  $\text{Supp}(\mu)$ .

It follows from work of Caffarelli [1990; 1992; 1999] that ( $\star$ ) holds true when each of the measures  $\mu$  and  $\nu$  satisfies the following additional condition: either the support of the measure is the entire  $\mathbb{R}^n$  or else the support is a bounded, convex domain and the 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 additional conditions. For related results on the regularity of optimal transportation, see Delanoë [1991] and Urbas [1997].

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

**Theorem 1.1.** *Let  $\mu, \nu$  be absolutely continuous, log-concave probability measures on  $\mathbb{R}^n$ . Let  $\nabla\Phi$  be the Brenier map between  $\mu$  and  $\nu$ , and assume  $(\star)$ . Write  $0 < \lambda_1(x) \leq \dots \leq \lambda_n(x)$  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  $\mu$ . Then, for  $i = 1, \dots, n$ ,*

$$\text{Var}[\log \lambda_i(X)] \leq 4.$$

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  $\mu$  and  $\nu$ , 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  $x \in \text{Supp}(\mu)$  set

$$\Lambda(x) = (\log \lambda_1(x), \dots, \log \lambda_n(x)).$$

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  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  with  $\mathbb{E}|f(\Lambda(X))| < \infty$ ,*

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

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

Set  $\pi = \Lambda_*(\mu)$ , the push-forward of the measure  $\mu$  under the map  $\Lambda$ . [Theorem 1.2](#) is a spectral gap estimate for the metric-measure space  $(\mathbb{R}^n, |\cdot|, \pi)$ . Gromov and Milman [[1983](#)] 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  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  be a 1-Lipschitz function (i.e.,  $|f(x) - f(y)| \leq |x - y|$ ).*

*Write  $A = \mathbb{E}f(\Lambda(X))$ . 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.

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

- (i)  $\text{Var}[Y] \leq 4$ .

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

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

(iii) For any 1-Lipschitz function  $f : \mathbb{R} \rightarrow \mathbb{R}$ , 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 [\[Klartag 2013, Theorem 1.4\]](#). The results in [\[Klartag 2013\]](#) were obtained under a technical assumption dubbed “regularity at infinity”, which we shall 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  $\text{Supp}(\mu)$ . The weighted Riemannian manifold

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

was studied in [\[Kolesnikov 2014\]](#), where it was shown that the associated Ricci–Bakry–Émery tensor is nonnegative when  $\mu$  and  $\nu$  are log-concave. We will also consider the map

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

from  $\text{Supp}(\mu) \subseteq \mathbb{R}^n$  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  $\Gamma_2$ -calculus. [Section 5](#) is devoted to the elimination of these regularity assumptions. In [Section 6](#) we complete the proofs of the theorems formulated above. 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  $C^\infty$ -smooth one. We write  $\log$  for the natural logarithm,  $x \cdot y$  stands for the standard scalar product of  $x, y \in \mathbb{R}^n$ , and  $\text{Tr}(A)$  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  $f : (0, \infty) \rightarrow \mathbb{R}$  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  $v_1, \dots, v_n \in \mathbb{R}^n$  and  $\lambda_1, \dots, \lambda_n > 0$ , where we write  $x \otimes x = (x_i x_j)_{i,j=1,\dots,n}$  for  $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ .

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

$$\|\log(A^{1/2} B A^{1/2})\|_{HS} \leq \|\log(A)\|_{HS} + \|\log(B)\|_{HS}, \quad (2)$$

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

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

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

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

$$D_k(AB) \leq D_k(A)D_k(B) \quad (k = 1, \dots, n). \quad (4)$$

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

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

where  $e^{\alpha_1} \geq \dots \geq e^{\alpha_n} > 0$  are the eigenvalues of  $A$  and  $e^{\beta_1} \geq \dots \geq e^{\beta_n} > 0$  are the eigenvalues of  $B$ . We will next apply a lemma of Weyl [1949]; see also [Polya 1950]. According to the inequality of Weyl and Polya, the inequalities (5) entail that

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

for any convex, nondecreasing function  $h : \mathbb{R} \rightarrow \mathbb{R}$ . For  $t \in \mathbb{R}$  let  $t_+ = \max\{t, 0\}$ . The function  $t \mapsto (t_+)^2$  is convex and nondecreasing; hence, from (6),

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

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

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

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, nondecreasing function  $h$ . By again using  $h(t) = (t_+)^2$ , we get

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

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

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

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

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

$$\text{dist}(A, B) = \|\log(A^{-1/2}BA^{-1/2})\|_{HS}. \quad (10)$$

Equivalently,  $\text{dist}(A, B)$  equals

$$\sqrt{\sum_i \log^2 \lambda_i},$$

where  $\lambda_1, \dots, \lambda_n > 0$  are the eigenvalues of the matrix  $A^{-1}B$  which is conjugate to  $A^{-1/2}BA^{-1/2}$ . The latter equivalent definition of  $\text{dist}$  shows that, for any invertible  $n \times n$  matrix  $T$ ,

$$\text{dist}(A, B) = \text{dist}(T^tAT, T^tBT) \quad (A, B \in M_n^+(\mathbb{R}^n)), \quad (11)$$

where  $T^t$  is the transpose of the matrix  $T$ . Observe too that  $\text{dist}(A, B) = \text{dist}(A^{-1}, B^{-1})$  for any  $A, B \in M_n^+(\mathbb{R})$ . [Lemma 2.1](#) states that, for  $A, B \in M_n^+(\mathbb{R})$ ,

$$\text{dist}(A, B) \leq \text{dist}(A, \text{Id}) + \text{dist}(\text{Id}, B), \quad (12)$$

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

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

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

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

where  $\dot{\gamma}(s) = d\gamma(s)/ds$  is a symmetric  $n \times n$  matrix. Then  $\text{Length}$  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  $B$ ,

$$\lim_{\varepsilon \rightarrow 0} \frac{\text{dist}^2(A + \varepsilon B, A)}{\varepsilon^2} = \|B\|_A^2 = \text{Tr}[(A^{-1}B)^2]. \quad (14)$$

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

$$\gamma_{A,B}(s) = A^{1/2}(A^{-1/2}BA^{-1/2})^s A^{1/2} \quad (0 \leq s \leq 1).$$

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

*Proof.* The invariance property (11) implies that

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

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

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

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  $A = \text{Id}$  by noting that

$$\gamma_{A,B}(s) = A^{1/2} \gamma_{\text{Id}, A^{-1/2}BA^{-1/2}}(s) A^{1/2} \quad (0 \leq s \leq 1).$$

Abbreviate  $\gamma(s) = \gamma_{A,B}(s) = \gamma_{\text{Id}, B}(s)$ . Since  $\gamma(s) = B^s$ , we have  $\dot{\gamma}(s) = B^s \log(B)$  and hence, for any  $0 \leq s \leq 1$ ,

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

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

The right-hand side of (14) depends quadratically on  $B$ , and therefore Lemma 2.2 tells us that our distance function  $\text{dist}$  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

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

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

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

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

From the Cauchy–Schwartz inequality,

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

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

**Lemma 2.4.** *For  $A \in M_n^+(\mathbb{R})$ , let its eigenvalues be  $\lambda_1(A) \geq \dots \geq \lambda_n(A) > 0$ . The map  $\Lambda : M_n^+(\mathbb{R}) \rightarrow \mathbb{R}^n$  defined via*

$$\Lambda(A) = (\log(\lambda_1(A)), \dots, \log(\lambda_n(A))). \quad (16)$$

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

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

Fix  $A_1, A_2 \in \mathcal{F}$  with  $A_1 \neq A_2$ . Consider the curve  $\gamma(s) = \gamma_{A_1, A_2}(s / \text{dist}(A_1, A_2))$ , where  $\gamma_{A_1, A_2}(s)$  is as in Lemma 2.2. Then  $\gamma$  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 the matrix  $\gamma(s)$  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

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

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

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

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

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

The relation (19) is standard; see, e.g., [Reed and Simon 1978, Section XII.1]. Consequently,

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

However, by (18),

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

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



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

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

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

### 3. Bakry–Émery $\Gamma_2$ -calculus

Let  $\mu$  and  $\nu$  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  $V, W : \mathbb{R}^n \rightarrow \mathbb{R}$ . Let  $\nabla\Phi$  be the Brenier map between  $\mu$  and  $\nu$ . Caffarelli's regularity theory states that  $\Phi : \mathbb{R}^n \rightarrow \mathbb{R}$  is a smooth, convex function. Therefore (1) implies that the transport equation

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

holds everywhere in  $\mathbb{R}^n$ . In particular, the matrix  $D^2\Phi(x) = (\Phi_{ij}(x))_{i,j=1,\dots,n}$  is invertible and hence positive-definite for any  $x \in \mathbb{R}^n$ . The inverse to  $D^2\Phi(x)$  is denoted by  $(D^2\Phi(x))^{-1} = (\Phi^{ij}(x))_{i,j=1,\dots,n}$ . 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  $\Phi_{jk}^i = \Phi^{i\ell} \Phi_{j\ell k}$  and  $\Phi_k^{ij} = \Phi^{i\ell} \Phi^{jm} \Phi_{km\ell}$ . Differentiating (22), we obtain

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

Following [Kolesnikov 2014], 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

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

See [Grigor'yan 2009] and Bakry, Gentil and Ledoux [Bakry et al. 2014] for background on weighted Riemannian manifolds and the  $\Gamma_2$ -calculus. For a smooth function  $u : \mathbb{R}^n \rightarrow \mathbb{R}$  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  $u, v : \mathbb{R}^n \rightarrow \mathbb{R}$ , via

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

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

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

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

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

for any smooth functions  $u, v : \mathbb{R}^n \rightarrow \mathbb{R}$ , one of which is compactly supported. The next step is to consider the carré du champ of  $M_{\mu, \nu}$ : As in [Bakry and Émery 1985], for a smooth function  $u : K \rightarrow \mathbb{R}$  we define

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

**Lemma 3.1.** *For any smooth function  $u : \mathbb{R}^n \rightarrow \mathbb{R}$ , 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 [Klartag 2013] 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 involves a tedious but straightforward computation which shows that

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

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

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

by representing the left-hand side of (26) as the trace of the square of the matrix  $B = (b_i^j)_{i,j=1,\dots,n}$ , 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

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

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

Another approach to Lemma 3.1 is to use the notation of Riemannian geometry as in [Kolesnikov 2014] and use the Bochner formula. We first observe that identity (23) in the case  $j = 1$  has the simple form

$$L\Phi_1 = -V_1. \quad (27)$$

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

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

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

$$\Gamma_2(u) = \|D_M^2 u\|_M^2 + \text{Ric}_M(\nabla_M u, \nabla_M u), \quad (29)$$

where  $\|D_M^2 u\|_M^2$  is the Hilbert–Schmidt norm of the Riemannian Hessian of  $u$  and  $\text{Ric}_M$  is the Bakry–Émery–Ricci tensor of the weighted Riemannian manifold  $M = M_{\mu, \nu}$ . Let us analyze the term in (29) involving the Hessian of  $u$ . The Christoffel 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 where  $u = \Phi_1$ , 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

$$\begin{aligned} \frac{1}{2}L(\Phi_{11}) &= -\langle \nabla_M \Phi_1, \nabla_M V_1 \rangle_M + \frac{1}{4}\Phi_{1j}^k \Phi_{1k}^j + \text{Ric}_M(\nabla_M u, \nabla_M u) \\ &= -V_{11} + \frac{1}{4}\Phi_{1j}^k \Phi_{1k}^j + (\text{Ric}_M)_{11}. \end{aligned} \quad (30)$$

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

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

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

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

Since  $D^2 V$  and  $D^2 W$  are positive semidefinite, for any smooth  $u : \mathbb{R}^n \rightarrow \mathbb{R}$  we have

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

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 [Klartag 2013]. The idea is to consider the map

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

Denote by  $(g_{ij}(x))_{i,j=1,\dots,n}$  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

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

Note that the positive semidefinite matrix  $(g_{ij}(x))_{i,j=1,\dots,n}$  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 [\[Klartag 2013\]](#), but in a cumbersome formulation and under an undesired assumption called “regularity at infinity”, which we eliminate here.

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

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

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., [\[Strichartz 1983\]](#)), 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*

$$\|Lu_k - f\|_{L^2(\mu)} \longrightarrow 0 \quad \text{as } k \rightarrow \infty.$$

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

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

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

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

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

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

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  $\eta_R$  be a smooth cutoff function in  $\mathbb{R}^n$  that equals one on a Euclidean ball of radius  $R$  centered at the origin, equals zero outside a Euclidean ball of radius  $2R$ , and satisfies  $|\nabla \eta_R| \leq 2/R$  throughout  $\mathbb{R}^n$ . Then,

$$\int_K |\nabla_M(\eta_R f)|^2 d\mu \leq \int_{\mathbb{R}^n} |\nabla_M \eta_R|^2 f^2 d\mu \leq \varepsilon_0^{-1} \int_{\mathbb{R}^n} |\nabla \eta_R|^2 f^2 d\mu \leq \frac{2}{R\varepsilon_0} \int_{\mathbb{R}^n} f^2 d\mu \longrightarrow 0 \quad \text{as } R \rightarrow \infty,$$

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

Suppose that  $F$  is a locally Lipschitz function on a Riemannian manifold such as  $M_n^+(\mathbb{R})$ . By the Rademacher theorem, the gradient  $\nabla F$  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 note we set

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

where  $\text{dist}$  is the Riemannian distance and  $B(x, \varepsilon) = \{y \mid \text{dist}(x, y) < \varepsilon\}$ . Since  $F$  is locally Lipschitz, the function  $|\nabla F|$  is locally bounded and upper semicontinuous. 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  $\theta$  the push-forward of the measure  $\mu$  under the map (32). Then, for any locally Lipschitz function  $F : M_n^+(\mathbb{R}) \rightarrow \mathbb{R}$  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  $L^2(\theta)$ , the function  $f$  defined via

$$f(x) = F(D^2\Phi(x)) \quad (x \in \mathbb{R}^n)$$

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

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

where  $f_i$  and  $g_{ij}$  are evaluated at the point  $x$ . In the case where the matrix  $(g_{ij}(x))_{i,j=1,\dots,n}$  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

$$h(x) \geq \sup\left\{\Phi^{ij} U_j f_i \mid g_{ij} \Phi^{ki} \Phi^{lj} U_k U_l \leq 1, U_1, \dots, U_n \in \mathbb{R}\right\}. \quad (38)$$

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  $u \in C_c^\infty(\mathbb{R}^n)$ ,

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

To this end we observe that, since  $u$  is compactly supported,

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

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

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

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

$$\begin{aligned} - \int_{\mathbb{R}^n} f(Lu) d\mu &= \int_{\mathbb{R}^n} \Phi^{ij} f_i u_j d\mu \leq \int_{\mathbb{R}^n} h(x) \sqrt{g_{ij} \Phi^{ki} \Phi^{\ell j} u_k u_\ell} d\mu(x) \\ &\leq \sqrt{\int_{\mathbb{R}^n} h^2 d\mu} \sqrt{\int_{\mathbb{R}^n} g_{ij} \Phi^{ki} \Phi^{\ell j} u_k u_\ell d\mu} \leq 2 \sqrt{\int_{\mathbb{R}^n} h^2 d\mu} \sqrt{\int_{\mathbb{R}^n} (Lu)^2 d\mu} \end{aligned}$$

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

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

Hence,

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

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

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

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

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

*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  $\Phi$  by a sequence of functions  $\Phi_N$  that satisfy assumption (34), and to prove the pointwise (and even local uniform) convergence  $D^2\Phi_N(x) \rightarrow D^2\Phi(x)$  as  $N \rightarrow \infty$ . Below we discuss two possible justifications of this convergence, as we believe that both of them may be useful. The first proof occupies Section 5A, and 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 Section 5B, is in fact an alternative approach to Caffarelli’s  $C^{1,\alpha}$ -regularity results in the log-concave case. The argument in Section 5B is more self-contained, and 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  $\mu$ . This integrability condition is rather mild in our opinion, and it is satisfied in many cases of interest.

**5A. First proof of Theorem 5.1.** As before, we write  $e^{-V}$  and  $e^{-W}$  for the densities of  $\mu$  and  $\nu$  respectively. By log-concavity, the functions  $V$  and  $W$  are locally Lipschitz in the open sets  $\text{Supp}(\mu)$  and  $\text{Supp}(\nu)$  respectively. From condition  $(\star)$ , the function  $\Phi$  is  $C^2$ -smooth, and the push-forward equation (1) implies that

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

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

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

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

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

is also  $C^2$ -smooth and strictly convex in  $\text{Supp}(\nu)$ , 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  $\text{Supp}(\mu)$  and  $\text{Supp}(\nu)$ . The reader is referred to [Rockafellar 1970] for the basic properties of the Legendre transform.

We will approximate  $\mu$  and  $\nu$  by sequences of probability measures  $\mu_N$  and  $\nu_N$  with the following properties:

- (i) The probability measures  $\mu_N$  and  $\nu_N$  have densities in  $\mathbb{R}^n$  of the form  $e^{-V_N}$  and  $e^{-W_N}$  respectively.
- (ii) The functions  $V_N, W_N : \mathbb{R}^n \rightarrow \mathbb{R}$  are smooth and, for any  $x \in \mathbb{R}^n$ ,

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

- (iii)  $V_N \rightarrow V$  locally uniformly in  $\text{Supp}(\mu)$  and, similarly,  $W_N \rightarrow W$  locally uniformly in  $\text{Supp}(\nu)$ .

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

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

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*

$$D^2\Phi_{N_j}(x) \rightarrow D^2\Phi(x) \quad \text{as } j \rightarrow \infty$$

*locally uniformly in  $x \in \text{Supp}(\mu)$ .*

*Proof.* Fix  $x_0 \in \text{Supp}(\mu)$ . It suffices to find  $\{N_j\}$  such that  $D^2\Phi_{N_j} \rightarrow D^2\Phi$  uniformly in a neighborhood of  $x_0$ . A standard convexity argument (e.g., [Klartag 2014, Section 2]) based on (iii) and the fact that  $\int e^{-V} = \int e^{-W} = 1$  shows that there exist  $A, B > 0$  with

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

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

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

From the Rellich–Kondrachov compactness theorem (e.g., [Evans and Gariepy 1992, Section 4.6]), we conclude that there exist a subsequence  $\Phi_{N_j}$ , numbers  $C_j \in \mathbb{R}$  and a certain function  $F : \text{Supp}(\mu) \rightarrow \mathbb{R}$  such that, for any compact  $K \subset \text{Supp}(\mu)$ , 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 [Rockafellar 1970, Theorem 10.9], we may assume that  $F$  is convex and that the convergence is locally uniform in  $\text{Supp}(\mu)$ . Thus, from [ibid., Theorem 24.5],

$$\nabla\Phi_N(x) \rightarrow \nabla F(x) \quad \text{as } N \rightarrow \infty \quad (45)$$

for almost any  $x \in \text{Supp}(\mu)$ . 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  $\text{Supp}(\mu)$ . Since  $\Phi$  is  $C^2$ -smooth, we may then apply [ibid., Theorem 25.7] and upgrade (45) to

$$\nabla\Phi_N(x) \rightarrow \nabla\Phi(x) \quad \text{as } N \rightarrow \infty \quad (46)$$

locally uniformly in  $\text{Supp}(\mu)$ . The convexity arguments in [ibid., Section 25] also give that  $\nabla\Phi_N^* \rightarrow \nabla\Phi^*$  locally uniformly in  $\text{Supp}(\nu)$ . As for the modulus of convexity, we have

$$\omega_{\Phi_N}(x; \delta) \rightarrow \omega_{\Phi}(x; \delta) \quad \text{as } N \rightarrow \infty \quad \text{and} \quad \omega_{\Phi_N^*}(y; \delta) \rightarrow \omega_{\Phi^*}(y; \delta) \quad \text{as } N \rightarrow \infty \quad (47)$$

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

We will now invoke the estimates of Gutierrez and Huang [2000] and Forzani and Maldonado [2004; 2005], which are constructive versions of Caffarelli’s  $C^{1,\alpha}$ -regularity theory. Thanks to (iii), (42), (46) and (47), we are allowed to apply [Gutiérrez and Huang 2000, Theorem 2.1] and [Forzani and Maldonado 2004, Theorem 15] locally near  $x_0$ . From the latter result, we learn that there exist  $\alpha, \delta, C > 0$  such that, for any  $x, y \in B(x_0, \delta)$  and  $N \geq 1$ ,

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



The function  $V$  is locally Lipschitz. From (iii) and [Rockafellar 1970, Theorem 24.5], the sequence  $\{V_N\}$  is uniformly locally Lipschitz. This means that, for any compact subset  $K \subset \text{Supp}(\mu)$ , 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 exists  $\widehat{C} > 0$  such that  $u_N(x) = -V_N(x) + W_N(\nabla\Phi_N(x))$  satisfies

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

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

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

We are finally in good shape for applying the  $C^{2,\alpha}$ -estimates from [Trudinger and Wang 2008, 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  $N \geq 1$ ,

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

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

$$\int_K (\Delta\Phi_N)\xi = - \int_K \nabla\Phi_N \cdot \nabla\xi \longrightarrow - \int_K \nabla\Phi \cdot \nabla\xi = \int_K (\Delta\Phi)\xi \quad \text{as } N \rightarrow \infty,$$

where  $\xi$  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, the sequence  $\{\int_K \|D^2\Phi_N\|_{HS}\}_{N \geq 1}$  is also bounded. From (49) and the Arzella–Ascoli theorem, there exists a subsequence, still denoted by  $\{\Phi_N\}$ , such that  $D^2\Phi_N \longrightarrow D^2\Phi$  uniformly on  $K = B(x_0, \delta/2)$ .  $\square$

**Remark 5.3.** Our proof of Lemma 5.2 does not make any use of the log-concavity of  $\mu$  and  $\nu$ . 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  $\theta_N$  the push-forward of the measure  $\mu_N$  under the map  $x \mapsto D^2\Phi_N(x)$ . Then, for any bounded, continuous function  $b : M_n^+(\mathbb{R}) \rightarrow \mathbb{R}$ ,*

$$\int_{M_n^+(\mathbb{R})} b \, d\theta_N \longrightarrow \int_{M_n^+(\mathbb{R})} b \, d\theta \quad \text{as } N \rightarrow \infty. \quad (50)$$

Furthermore, if  $b : M_n^+(\mathbb{R}) \rightarrow \mathbb{R}$  is bounded and upper semicontinuous, then

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

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

$$\int_{\mathbb{R}^n} b(D^2\Phi_N(x))e^{-V_N(x)} \, dx \longrightarrow \int_{\mathbb{R}^n} b(D^2\Phi(x))e^{-V(x)} \, dx \quad \text{as } N \rightarrow \infty.$$

This follows from [Lemma 5.2](#) and from the dominated convergence theorem, since [\(43\)](#) provides an integrable majorant. Next, assume that  $b$  is bounded and upper semicontinuous. Then, for any  $x \in \text{Supp}(\mu)$ ,

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

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

*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 [\[2000\]](#) which appears in [\[Kolesnikov 2010\]](#), and thus obtain from [\(ii\)](#) that, for any  $x \in \mathbb{R}^n$ ,

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

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 semicontinuous and bounded, while  $F$  is continuous and bounded. By taking the limit as  $N \rightarrow \infty$  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 select a smooth function  $\theta_R : M_n^+(\mathbb{R}) \rightarrow [0, 1]$  such that  $\theta_R$  equals one on  $B(\text{Id}, R)$  and vanishes outside  $B(\text{Id}, 2R)$ , 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,  $F_R \rightarrow F$  in  $L^2(\theta)$  as  $R \rightarrow \infty$ . All that remains is to show that

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

The functions  $\theta_R$  and  $F$  are continuous, and we may therefore 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  $F, |\nabla F| \in L^2(\theta)$ , [\(52\)](#) follows in the case where  $F \in L^2(\theta)$ .

Finally, to eliminate the assumption that  $F$  is 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  $\{A \in M_n^+(\mathbb{R}) \mid F(A) = R\}$  has zero  $\theta$ -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.  $\square$

**5B. Second proof: log-concave target measure.** In our second proof we will exploit the fact that  $\nu$  is log-concave, but we will not require the log-concavity of  $\mu$ . Throughout this subsection we make the following additional 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.,

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

By the Morrey embedding theorem (see, e.g., [Evans and Gariepy 1992, Section 4.5]), the function  $V$  is locally Hölder. We will approximate  $\mu$  and  $\nu$  by sequences of probability measures  $\mu_N$  and  $\nu_N$  having properties (i), (ii) and (iii) from Section 5A. We also require a fourth property:

(i) 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 Section 5A — to convolve with a tiny Gaussian and then multiply by the density of a huge Gaussian — also yields property (iv). Recall that the Brenier map  $\nabla \Phi_N$  between  $\mu_N$  and  $\nu_N$  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  $\Phi$  are smooth on the entire  $\mathbb{R}^n$  and that  $\nu$  is a log-concave measure. Then, for every  $q \geq 2$ ,  $0 < \tau < 1$ ,  $i = 1, \dots, n$ , there exists  $C(q, \tau) > 0$  such that*

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

*Proof.* Assume in addition that  $D^2 W \geq (1/C) \cdot \text{Id}$ ,  $D^2 V \leq C \cdot \text{Id}$ . In this case,  $D^2 \Phi \leq C^2 \cdot \text{Id}$ . 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  $\Phi_{ii}^p$ ,  $p \geq 0$ , and formally integrate by parts with respect to  $\mu$ . Using the convexity of  $W$  we obtain

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

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

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

Applying the Cauchy inequality yields

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

Finally,

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

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

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

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

It is helpful to keep in mind that  $\Phi_i^{jk} \Phi_{ijk} = \text{Tr}[(D^2\Phi)^{-1/2} D^2\Phi_i (D^2\Phi)^{-1/2}]^2 \geq 0$ . 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

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

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

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

Finally,

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

Applying the Hölder inequalities

$$\begin{aligned} \Phi_i^2 \Phi_{ii}^{q-\tau} &\leq \frac{q-\tau}{q} \Phi_{ii}^q + \frac{\tau}{q} |\Phi_i|^{2q/\tau}, \\ \Phi_{ii}^{q-2+\tau} V_i^2 &\leq \varepsilon \Phi_{ii}^q + C(\varepsilon, q, \tau) |V_i|^{2q/(2-\tau)}, \\ |\Phi_i V_i|^q &\leq \frac{1}{2}(2-\tau) |V_i|^{2q/(2-\tau)} + \frac{1}{2}\tau |\Phi_i|^{2q/\tau}, \end{aligned}$$

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

Finally, let us get rid of the assumption that  $D^2W \geq (1/C) \cdot \text{Id}$ ,  $D^2V \leq C \cdot \text{Id}$ . To this end we approximate  $\mu$  and  $\nu$  by measures with smooth potentials satisfying  $D^2W_N \geq (1/C_N) \cdot \text{Id}$ ,  $D^2V_N \leq C_N \cdot \text{Id}$  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  $\Phi_{ii}$ . The latter can be easily shown with the help of integration by parts and identifications of the pointwise limit  $\lim_N \nabla \Phi_N$  with  $\nabla \Phi$  (see the proof of [Lemma 5.2](#)).  $\square$

**Remark 5.6.** The conclusion of [Proposition 5.5](#) holds without any additional smoothness assumptions. This can be verified by smooth approximations (again, see [\[Kolesnikov 2013\]](#) for details). Finally, we see that [\(53\)](#) holds for every log-concave measure  $\nu$  and measure  $\mu$  satisfying  $\int |V_i|^{2q/(2-\tau)} d\mu < \infty$ , where  $V_i$  is the logarithmic derivative of  $\mu$  along  $x_i$ .

*Second proof of [Lemma 5.2](#).* Let us demonstrate how [Proposition 5.5](#) implies [\(48\)](#) above without appealing to the works by Forzani and Maldonado [\[2004; 2005\]](#) and Gutierrez and Huang [\[2000\]](#) related to Caffarelli's  $C^{1,\alpha}$ -regularity theory. We know that  $\sup_N \int |\nabla V_N|^p e^{-V_N} dx < \infty$ ,  $p > n$ . Since  $\nu$  is log-concave, all the moments of  $\nu$  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 the fact that the  $V_N$  are uniformly locally bounded from below, we see that  $\sup_N \int_{B_R} \|D^2\Phi_N\|_{HS}^{p'} dx < \infty$  for every  $R$ . The result then follows from the Morrey embedding theorem.  $\square$

## 6. Corollaries to [Theorem 5.1](#)

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

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

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

$$|\nabla F|(A) \leq |\nabla f|(\log \lambda_1(A), \dots, \log \lambda_n(A)). \quad (56)$$

Since  $f$  is locally Lipschitz and the eigenvalues vary continuously with the matrix  $A$ , (56) implies that  $F$  is locally Lipschitz. Denote by  $\theta$  the push-forward of the probability measure  $\mu$  under the map  $x \mapsto D^2\Phi(x)$ . Since  $\mathbb{E}|f(\Lambda(X))| < \infty$ ,  $F \in L^1(\theta)$ . Since  $\mathbb{E}|\nabla f|^2(\Lambda(X)) < \infty$ ,  $\int |\nabla F|^2 d\theta < \infty$ . 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  $\text{Var}[f(\Lambda(X))]$ . Glancing at (56), we thus obtain

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

and the proof is complete. □

*Proof of [Theorem 1.1](#).* Substitute  $f(x) = x_i$  in [Theorem 1.2](#). Then  $f$  is a 1-Lipschitz function, and by [Remark 1.4](#) we have  $\mathbb{E}|f(\Lambda(X))| < \infty$ . Thus, the application of [Theorem 1.2](#) is legitimate, and [Theorem 1.1](#) follows. □

*Proof of [Theorem 1.5](#).* The argument is almost identical to the proof of [Theorem 1.1](#), with [Lemma 2.3](#) replacing [Lemma 2.4](#). □

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 should be possible to prove concentration inequalities for the eigenvalues of  $D^2\Phi$  themselves and not only for their logarithms. For example, 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  $T$ ,

$$\int_K \Delta\Phi \leq nV(K, \dots, K, T),$$

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 nonlinear setting, when transporting measures with convexity properties supported on Riemannian manifolds.

### Acknowledgements

We would like to thank Emanuel Milman for interesting discussions. The research made in Sections 1–2, 4, 6 was supported by a grant from the European Research Council (ERC). The research made in Sections 3, 5 was supported by the Russian Science Foundation grant 14-11-00196.

### References

[Alesker et al. 1999] S. Alesker, S. Dar, and V. Milman, “A remarkable measure preserving diffeomorphism between two convex bodies in  $\mathbf{R}^n$ ”, *Geom. Dedicata* **74**:2 (1999), 201–212. [MR 2000a:52004](#) [Zbl 0927.52007](#)

- [Bakry and Émery 1985] D. Bakry and M. Émery, “Diffusions hypercontractives”, pp. 177–206 in *Séminaire de probabilités, XIX, 1983/84*, edited by J. Azéma and M. Yor, Lecture Notes in Math. **1123**, Springer, Berlin, 1985. [MR 88j:60131](#) [Zbl 0561.60080](#)
- [Bakry et al. 2014] D. Bakry, I. Gentil, and M. Ledoux, *Analysis and geometry of Markov diffusion operators*, Grundlehren der Math. Wiss. **348**, Springer, Cham, 2014. [MR 3155209](#) [Zbl 06175511](#)
- [Brenier 1991] Y. Brenier, “Polar factorization and monotone rearrangement of vector-valued functions”, *Comm. Pure Appl. Math.* **44**:4 (1991), 375–417. [MR 92d:46088](#) [Zbl 0738.46011](#)
- [Caffarelli 1990] L. A. Caffarelli, “A localization property of viscosity solutions to the Monge–Ampère equation and their strict convexity”, *Ann. of Math. (2)* **131**:1 (1990), 129–134. [MR 91f:35058](#) [Zbl 0704.35045](#)
- [Caffarelli 1992] L. A. Caffarelli, “The regularity of mappings with a convex potential”, *J. Amer. Math. Soc.* **5**:1 (1992), 99–104. [MR 92j:35018](#) [Zbl 0753.35031](#)
- [Caffarelli 2000] L. A. Caffarelli, “Monotonicity properties of optimal transportation and the FKG and related inequalities”, *Comm. Math. Phys.* **214**:3 (2000), 547–563. Erratum in **225**:2 (2002), 449–450. [MR 2002c:60029](#) [Zbl 0978.60107](#)
- [Delanoë 1991] P. Delanoë, “Classical solvability in dimension two of the second boundary-value problem associated with the Monge–Ampère operator”, *Ann. Inst. H. Poincaré Anal. Non Linéaire* **8**:5 (1991), 443–457. [MR 92g:35070](#) [Zbl 0778.35037](#)
- [Evans and Gariepy 1992] L. C. Evans and R. F. Gariepy, *Measure theory and fine properties of functions*, CRC Press, Boca Raton, FL, 1992. [MR 93f:28001](#) [Zbl 0804.28001](#)
- [Forzani and Maldonado 2004] L. Forzani and D. Maldonado, “Properties of the solutions to the Monge–Ampère equation”, *Nonlinear Anal.* **57**:5-6 (2004), 815–829. [MR 2005c:35095](#) [Zbl 1137.35361](#)
- [Forzani and Maldonado 2005] L. Forzani and D. Maldonado, “Recent progress on the Monge–Ampère equation”, pp. 189–198 in *The  $p$ -harmonic equation and recent advances in analysis*, edited by P. Poggi-Corradini, Contemp. Math. **370**, Amer. Math. Soc., Providence, RI, 2005. [MR 2005k:35115](#) [Zbl 1134.35334](#)
- [Grigor’yan 2009] A. Grigor’yan, *Heat kernel and analysis on manifolds*, AMS/IP Studies in Advanced Mathematics **47**, American Mathematical Society, Providence, RI, 2009. [MR 2011e:58041](#) [Zbl 1206.58008](#)
- [Gromov and Milman 1983] M. Gromov and V. D. Milman, “A topological application of the isoperimetric inequality”, *Amer. J. Math.* **105**:4 (1983), 843–854. [MR 84k:28012](#) [Zbl 0522.53039](#)
- [Gutiérrez and Huang 2000] C. E. Gutiérrez and Q. Huang, “Geometric properties of the sections of solutions to the Monge–Ampère equation”, *Trans. Amer. Math. Soc.* **352**:9 (2000), 4381–4396. [MR 2000m:35060](#) [Zbl 0958.35043](#)
- [Klartag 2013] B. Klartag, “Poincaré inequalities and moment maps”, *Ann. Fac. Sci. Toulouse Math. (6)* **22**:1 (2013), 1–41. [MR 3247770](#) [Zbl 1279.60036](#)
- [Klartag 2014] B. Klartag, “Logarithmically-concave moment measures I”, pp. 231–260 in *Geometric aspects of functional analysis*, edited by B. Klartag and E. Milman, Lecture Notes in Math. **2116**, Springer International Publishing, 2014.
- [Kolesnikov 2010] A. V. Kolesnikov, “Mass transportation and contractions”, *MIPT Proc.* **2**:4 (2010), 90–99. In Russian; translated at [arXiv 1103.1479](#).
- [Kolesnikov 2013] A. V. Kolesnikov, “Sobolev regularity of mass transportation and transportation inequalities”, *Teor. Veroyatnost. i Primenen.* **57**:2 (2013), 296–321. In Russian; translated in *Theory Probab. Appl.* **57**:2 (2013), 243–264. [MR 3201654](#) [Zbl 1279.60032](#)
- [Kolesnikov 2014] A. V. Kolesnikov, “Hessian metrics,  $CD(K, N)$ -spaces, and optimal transportation of log-concave measures”, *Discrete Contin. Dyn. Syst.* **34**:4 (2014), 1511–1532. [MR 3121630](#) [Zbl 1279.35045](#)
- [McCann 1995] R. J. McCann, “Existence and uniqueness of monotone measure-preserving maps”, *Duke Math. J.* **80**:2 (1995), 309–323. [MR 97d:49045](#) [Zbl 0873.28009](#)
- [Polya 1950] G. Polya, “Remark on Weyl’s note ‘‘Inequalities between the two kinds of eigenvalues of a linear transformation.’’”, *Proc. Nat. Acad. Sci. U. S. A.* **36** (1950), 49–51. [MR 11,526b](#) [Zbl 0041.15402](#)
- [Reed and Simon 1978] M. Reed and B. Simon, *Methods of modern mathematical physics, IV: Analysis of operators*, Academic Press, New York, 1978. [MR 58 #12429c](#) [Zbl 0401.47001](#)
- [Rockafellar 1970] R. T. Rockafellar, *Convex analysis*, Princeton Mathematical Series **28**, Princeton University Press, Princeton, N.J., 1970. [MR 43 #445](#) [Zbl 0193.18401](#)

- [Strichartz 1983] R. S. Strichartz, “Analysis of the Laplacian on the complete Riemannian manifold”, *J. Funct. Anal.* **52**:1 (1983), 48–79. [MR 84m:58138](#) [Zbl 0515.58037](#)
- [Trudinger and Wang 2008] N. S. Trudinger and X.-J. Wang, “The Monge–Ampère equation and its geometric applications”, pp. 467–524 in *Handbook of geometric analysis*, vol. 1, edited by L. Ji et al., Adv. Lect. Math. (ALM) **7**, International Press, 2008. [MR 2010g:53065](#) [Zbl 1156.35033](#)
- [Urbas 1997] J. Urbas, “On the second boundary value problem for equations of Monge–Ampère type”, *J. Reine Angew. Math.* **487** (1997), 115–124. [MR 98f:35057](#) [Zbl 0880.35031](#)
- [Weyl 1949] H. Weyl, “Inequalities between the two kinds of eigenvalues of a linear transformation”, *Proc. Nat. Acad. Sci. U. S. A.* **35** (1949), 408–411. [MR 11,37d](#) [Zbl 0032.38701](#)

Received 13 Feb 2014. Accepted 2 Dec 2014.

BO’AZ B. KLARTAG: [klartagb@post.tau.ac.il](mailto:klartagb@post.tau.ac.il)

*School of Mathematical Sciences, Tel Aviv University, 69978 Tel Aviv, Israel*

ALEXANDER V. KOLESNIKOV: [sascha77@mail.ru](mailto:sascha77@mail.ru)

*Faculty of Mathematics, National Research University Higher School of Economics, 117312, Moscow, Vavilova St., 7, Russia*



# Analysis & PDE

[msp.org/apde](http://msp.org/apde)

## EDITORS

EDITOR-IN-CHIEF

Maciej Zworski  
[zworski@math.berkeley.edu](mailto:zworski@math.berkeley.edu)  
University of California  
Berkeley, USA

## BOARD OF EDITORS

Nicolas Burq	Université Paris-Sud 11, France <a href="mailto:nicolas.burq@math.u-psud.fr">nicolas.burq@math.u-psud.fr</a>	Yuval Peres	University of California, Berkeley, USA <a href="mailto:peres@stat.berkeley.edu">peres@stat.berkeley.edu</a>
Sun-Yung Alice Chang	Princeton University, USA <a href="mailto:chang@math.princeton.edu">chang@math.princeton.edu</a>	Gilles Pisier	Texas A&M University, and Paris 6 <a href="mailto:pisier@math.tamu.edu">pisier@math.tamu.edu</a>
Michael Christ	University of California, Berkeley, USA <a href="mailto:mchrist@math.berkeley.edu">mchrist@math.berkeley.edu</a>	Tristan Rivière	ETH, Switzerland <a href="mailto:riviere@math.ethz.ch">riviere@math.ethz.ch</a>
Charles Fefferman	Princeton University, USA <a href="mailto:cf@math.princeton.edu">cf@math.princeton.edu</a>	Igor Rodnianski	Princeton University, USA <a href="mailto:irod@math.princeton.edu">irod@math.princeton.edu</a>
Ursula Hamenstaedt	Universität Bonn, Germany <a href="mailto:ursula@math.uni-bonn.de">ursula@math.uni-bonn.de</a>	Wilhelm Schlag	University of Chicago, USA <a href="mailto:schlag@math.uchicago.edu">schlag@math.uchicago.edu</a>
Vaughan Jones	U.C. Berkeley & Vanderbilt University <a href="mailto:vaughan.f.jones@vanderbilt.edu">vaughan.f.jones@vanderbilt.edu</a>	Sylvia Serfaty	New York University, USA <a href="mailto:serfaty@cims.nyu.edu">serfaty@cims.nyu.edu</a>
Herbert Koch	Universität Bonn, Germany <a href="mailto:koch@math.uni-bonn.de">koch@math.uni-bonn.de</a>	Yum-Tong Siu	Harvard University, USA <a href="mailto:siu@math.harvard.edu">siu@math.harvard.edu</a>
Izabella Laba	University of British Columbia, Canada <a href="mailto:ilaba@math.ubc.ca">ilaba@math.ubc.ca</a>	Terence Tao	University of California, Los Angeles, USA <a href="mailto:tao@math.ucla.edu">tao@math.ucla.edu</a>
Gilles Lebeau	Université de Nice Sophia Antipolis, France <a href="mailto:lebeau@unice.fr">lebeau@unice.fr</a>	Michael E. Taylor	Univ. of North Carolina, Chapel Hill, USA <a href="mailto:met@math.unc.edu">met@math.unc.edu</a>
László Lempert	Purdue University, USA <a href="mailto:lempert@math.purdue.edu">lempert@math.purdue.edu</a>	Gunther Uhlmann	University of Washington, USA <a href="mailto:gunther@math.washington.edu">gunther@math.washington.edu</a>
Richard B. Melrose	Massachusetts Institute of Technology, USA <a href="mailto:rbm@math.mit.edu">rbm@math.mit.edu</a>	András Vasy	Stanford University, USA <a href="mailto:andras@math.stanford.edu">andras@math.stanford.edu</a>
Frank Merle	Université de Cergy-Pontoise, France <a href="mailto:Frank.Merle@u-cergy.fr">Frank.Merle@u-cergy.fr</a>	Dan Virgil Voiculescu	University of California, Berkeley, USA <a href="mailto:dvv@math.berkeley.edu">dvv@math.berkeley.edu</a>
William Minicozzi II	Johns Hopkins University, USA <a href="mailto:minicozz@math.jhu.edu">minicozz@math.jhu.edu</a>	Steven Zelditch	Northwestern University, USA <a href="mailto:zelditch@math.northwestern.edu">zelditch@math.northwestern.edu</a>
Werner Müller	Universität Bonn, Germany <a href="mailto:mueller@math.uni-bonn.de">mueller@math.uni-bonn.de</a>		

## PRODUCTION

[production@msp.org](mailto:production@msp.org)

Silvio Levy, Scientific Editor

---

See inside back cover or [msp.org/apde](http://msp.org/apde) for submission instructions.

---

The subscription price for 2015 is US \$205/year for the electronic version, and \$390/year (+\$55, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues from the last three years and changes of subscribers address should be sent to MSP.

---

Analysis & PDE (ISSN 1948-206X electronic, 2157-5045 printed) at Mathematical Sciences Publishers, 798 Evans Hall #3840, c/o University of California, Berkeley, CA 94720-3840, is published continuously online. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices.

---

APDE peer review and production are managed by EditFLOW<sup>®</sup> from MSP.

PUBLISHED BY

 **mathematical sciences publishers**  
nonprofit scientific publishing

<http://msp.org/>

© 2015 Mathematical Sciences Publishers

# ANALYSIS & PDE

Volume 8 No. 1 2015

---

Hölder continuity and bounds for fundamental solutions to nondivergence form parabolic equations SEIICHIRO KUSUOKA	1
Eigenvalue distribution of optimal transportation BO'AZ B. KLARTAG and ALEXANDER V. KOLESNIKOV	33
Nonlocal self-improving properties TUOMO KUUSI, GIUSEPPE MINGIONE and YANNICK SIRE	57
Symbol calculus for operators of layer potential type on Lipschitz surfaces with VMO normals, and related pseudodifferential operator calculus STEVE HOFMANN, MARIUS MITREA and MICHAEL E. TAYLOR	115
Criteria for Hankel operators to be sign-definite DIMITRI R. YAFAEV	183
Nodal sets and growth exponents of Laplace eigenfunctions on surfaces GUILLAUME ROY-FORTIN	223