

INTEGRABILITY OF ABSOLUTELY CONTINUOUS TRANSFORMATIONS OF MEASURES AND APPLICATIONS TO OPTIMAL MASS TRANSPORTATION

VLADIMIR I. BOGACHEV, ALEXANDER V. KOLESNIKOV

ABSTRACT. Given two Borel probability measures μ and ν on \mathbb{R}^d such that $d\nu/d\mu = g$, we consider certain mappings of the form $T(x) = x + F(x)$ that transform μ into ν . Our main results give estimates of the form

$$\int_{\mathbb{R}^d} \Phi_1(|F|) d\mu \leq \int_{\mathbb{R}^d} \Phi_2(g) d\mu$$

for certain functions Φ_1 and Φ_2 under appropriate assumptions on μ . Applications are given to optimal mass transportations in the Monge problem.

Keywords: optimal transportation, Gaussian measure, convex measure, logarithmic Sobolev inequality, Poincaré inequality, Talagrand inequality.

AMS Subject Classification: 28C20, 46G12, 49J40, 60B11.

1. INTRODUCTION

In the last decade there has been a growing interest in the study of connections between the two classical optimization problems raised by Monge and Kantorovich. In particular, fruitful relations between these problems, nonlinear transformations of measures, nonlinear differential equations and nonlinear functional inequalities have been revealed (see [1], [3], [4], [8], [9], [10], [11], [12], [13], [14], [15], [17], [18], [20], [21], [23], [24], [26], [28], [29], where one can find additional references). We recall that the general Monge problem deals with measurable mappings T from a given probability space (X, \mathcal{A}, μ) to a probability space (Y, \mathcal{B}, ν) that transform μ into ν and minimize the integrals

$$K(\mu, \nu, h, T) := \int_X h(x, T(x)) \mu(dx)$$

for a given nonnegative measurable function h on $X \times Y$ (called a cost function). The corresponding infimum is denoted by $K(\mu, \nu, h)$. Under broad assumptions, there is a mapping T at which the infimum is attained. Such a mapping is called an optimal transportation between μ and ν . If $X = Y$ is a metric space with a distance d , then typical cost functions are $h(x, y) = d(x, y)^p$. It has been shown by Talagrand [28] that if $X = Y = \mathbb{R}^d$, $h(x, y) = |x - y|^2$, $\mu = \gamma$ is the standard Gaussian measure and $\nu = g \cdot \gamma$, where $g \log g \in L^1(\gamma)$ with the convention $0 \log 0 := 0$, then

$$K(\gamma, g \cdot \gamma, h) \leq 2 \text{Ent}_\gamma(g),$$

where

$$\text{Ent}_\gamma(g) := \int_X g \log g d\gamma.$$

Writing T as $T(x) = x + F(x)$, we represent Talagrand's inequality in the form

$$\int_X |F|^2 d\gamma \leq 2 \int_X g \log g d\gamma.$$

In this work, we consider a more general problem when a mapping $T = I + F$, where $I(x) = x$, transforms a measure μ into a measure $g \cdot \mu$ and one is interested in the integrability of functions of $|F|$ under appropriate integrability assumptions on g . An interesting result in this direction is due to Fernique [16] who considered a Gaussian measure γ on a separable Fréchet space and a probability measure $g \cdot \gamma$ such that $g \in L^p(\gamma)$ for some $p > 1$. It was shown in [16] that there exists a mapping $T = U + S$, where the mapping U preserves the measure γ and S is a mapping with values in the Cameron–Martin space H of γ , such that the function $\exp(\omega|S|_H^2)$ is integrable for sufficiently small ω (however, the mapping T is not necessarily the optimal transportation). Our main result in Section 2 contains Talagrand’s inequality and Fernique’s estimate for optimal transportation as partial cases and applies to more general transformations. All the results in this work have the following form: if a probability measure μ is transformed into a probability measure $\nu = g \cdot \mu$ by a mapping $T(x) = x + F(x)$, then

$$\int_{\mathbb{R}^d} \Phi_1(|F(x)|) \mu(dx) \leq \int_{\mathbb{R}^d} \Phi_2(g(x)) \mu(dx)$$

for certain functions Φ_1 and Φ_2 under appropriate assumptions on μ . A lot of special cases are obtained by varying the class of measures μ , transformations T , and functions Φ_2 . For example, if μ is Gaussian (or satisfies the logarithmic Sobolev inequality) and $\Phi_2(g) = g^p$, then $\Phi_1(|F|) = \exp(\omega|F|^2)$, and if $\Phi_2(g) = g|\log g|^p$, then $\Phi_1(|F|) = |F|^r$. Section 2 deals with the case where μ is a Gaussian or strictly convex measure, T is either a rather general invertible transformation or an optimal transportation for the cost functions $h(x, y) = |x - y|^p$, $p \in (1, 2]$. In Section 3 and Section 4 we consider the optimal transportations for $h(x, y) = |x - y|^2$ in the case of measures satisfying the logarithmic Sobolev inequality and the Poincaré inequality, respectively.

Let us introduce some notation and terminology. Given a measure μ and a μ -integrable function g , we denote by $g \cdot \mu$ the measure with density g with respect to μ . The symbol $W^{p,k}(U)$ stands for the classical Sobolev space of functions on an open set $U \subset \mathbb{R}^d$ that belong to $L^p(U)$ along with their generalized partial derivatives up to order k . The symbol $W^{p,k}(U, \mathbb{R}^d)$ denotes the space of mappings from U to \mathbb{R}^d whose components belong to $W^{p,k}(U)$. The class $W_{loc}^{p,k}$ consists of mappings (or functions) f on \mathbb{R}^d such that $\zeta f \in W^{p,k}(\mathbb{R}^d, \mathbb{R}^d)$ (respectively, $\zeta f \in W^{p,k}(\mathbb{R}^d)$) for every $\zeta \in C_0^\infty(\mathbb{R}^d)$.

Suppose μ is a Borel measure on \mathbb{R}^d with density $\varrho \in W_{loc}^{1,1}$. Given a mapping $F \in W_{loc}^{1,1}$, let

$$\delta_\mu F(x) := \operatorname{div} F(x) + \langle F(x), \nabla \varrho(x) / \varrho(x) \rangle,$$

where we set $\nabla \varrho(x) / \varrho(x) = 0$ whenever $\varrho(x) = 0$. If $|F| \varrho, |F| |\nabla \varrho|, |DF| \varrho \in L_{loc}^1(\mathbb{R}^d)$, then for all $\varphi \in C_0^\infty(\mathbb{R}^d)$ one has

$$\int_{\mathbb{R}^d} \varphi \delta_\mu F d\mu = - \int_{\mathbb{R}^d} \langle \nabla \varphi, F \rangle d\mu.$$

Let $x^2 := |x|^2$ if $x \in \mathbb{R}^d$. Given a linear operator A on \mathbb{R}^d , we set

$$\det_2 A := \det A \exp(\operatorname{Trace}(I - A)).$$

If A is symmetric and has eigenvalues a_1, \dots, a_d , then $\det_2 A = \prod_{i=1}^d a_i e^{1-a_i}$. In particular, if $A \geq 0$, then $0 \leq \det_2 A \leq 1$. If $A = I + B$, then

$$\det_2 A = \det(I + B) \exp(-\operatorname{Trace} B).$$

An absolutely continuous measure μ on \mathbb{R}^d is called convex if its density has the form $\exp(-V)$, where V is a convex function.

A probability measure μ on \mathbb{R}^d is said to satisfy the logarithmic Sobolev inequality if there exists a number $C > 0$ such that the following inequality holds for every function $\varphi \in C_b^\infty(\mathbb{R}^d)$:

$$\int_{\mathbb{R}^d} \varphi^2 \log \varphi^2 d\mu - \left(\int_{\mathbb{R}^d} \varphi^2 d\mu \right) \log \left(\int_{\mathbb{R}^d} \varphi^2 d\mu \right) \leq 2C \int_{\mathbb{R}^d} |\nabla \varphi|^2 d\mu. \quad (1.1)$$

A necessary condition for μ to satisfy (1.1) is the existence of $\varepsilon > 0$ such that $\exp(\varepsilon|x|^2) \in L^1(\mu)$ (in fact, this is true for any $\varepsilon < (2C)^{-1}$, see [22, Ch. 5]). It has been shown in [30] that if μ is a convex measure such that $\exp(\varepsilon|x|^2) \in L^1(\mu)$ for some $\varepsilon > 0$, then μ satisfies (1.1) with some $C(\varepsilon)$. As an example of a measure satisfying (1.1) one can take a probability measure with density $\exp(-V)$, where V is a twice differentiable convex function with $D^2V(x) \geq C^{-1}I$ (see [22]).

A probability measure μ on \mathbb{R}^d is said to satisfy the Poincaré inequality if for all $\varphi \in C_b^\infty(\mathbb{R}^d)$ one has

$$\int_{\mathbb{R}^d} \left(\varphi - \int_{\mathbb{R}^d} \varphi d\mu \right)^2 d\mu \leq C_2 \int_{\mathbb{R}^d} |\nabla \varphi|^2 d\mu. \quad (1.2)$$

According to [2], any convex measure satisfies (1.2). Note that every measure μ satisfying (1.2) satisfies also the inequality

$$\int_{\mathbb{R}^d} \left| \varphi - \int_{\mathbb{R}^d} \varphi d\mu \right|^p d\mu \leq C_p \int_{\mathbb{R}^d} |\nabla \varphi|^p d\mu, \quad \varphi \in C_b^\infty(\mathbb{R}^d), \quad (1.3)$$

for every $p \geq 2$ and $C_p = C_2^{p/2}(1 + p^2/4)^{p/2}$. Indeed, without loss of generality we may assume that $\int_{\mathbb{R}^d} \varphi d\mu = 0$. Let $p = 2 + t$. By the Poincaré and Hölder inequalities one has

$$\begin{aligned} \int_{\mathbb{R}^d} |\varphi|^{2+t} d\mu &\leq \left(\int_{\mathbb{R}^d} \varphi |\varphi|^{t/2} d\mu \right)^2 + C_2(1 + t/2)^2 \int_{\mathbb{R}^d} |\nabla \varphi|^2 |\varphi|^t d\mu \\ &\leq \int_{\mathbb{R}^d} \varphi^2 d\mu \int_{\mathbb{R}^d} |\varphi|^t d\mu + C_2(1 + t/2)^2 \left(\int_{\mathbb{R}^d} |\nabla \varphi|^{2+t} d\mu \right)^{\frac{2}{2+t}} \left(\int_{\mathbb{R}^d} |\varphi|^{2+t} d\mu \right)^{\frac{t}{2+t}} \\ &\leq \left[\int_{\mathbb{R}^d} C_2 |\nabla \varphi|^2 d\mu + C_2(1 + t/2)^2 \left(\int_{\mathbb{R}^d} |\nabla \varphi|^{2+t} d\mu \right)^{\frac{2}{2+t}} \right] \left(\int_{\mathbb{R}^d} |\varphi|^{2+t} d\mu \right)^{\frac{t}{2+t}} \\ &\leq C_2 \left[1 + (1 + t/2)^2 \right] \left(\int_{\mathbb{R}^d} |\nabla \varphi|^{2+t} d\mu \right)^{\frac{2}{2+t}} \left(\int_{\mathbb{R}^d} |\varphi|^{2+t} d\mu \right)^{\frac{t}{2+t}}, \end{aligned}$$

which yields our claim.

Note that if μ satisfies (1.1) or (1.2) with some constants, then any probability measure of the form $\exp U \cdot \mu$ also does provided that U is bounded (see [22, p. 96]).

Finally, we recall that by Brenier's theorem [13] (see [29, p. 66]) for every pair of Borel probability measures μ and ν on \mathbb{R}^d with finite second moments such that μ is absolutely continuous, there is a convex function V on \mathbb{R}^d (which is finite on a convex set of full μ -measure) such that the mapping $T = \nabla V$ transforms μ into ν and is a unique optimal transportation between μ and ν for the cost function $h(x, y) = |x - y|^2$. Moreover, suppose that ν is also absolutely continuous. Let us define the Legendre transform V^* of the function V by

$$V^*(x) = \sup_y [\langle x, y \rangle - V(y)].$$

Then V^* is finite and convex on a convex set of full ν -measure, the mapping $S = \nabla V^*$ transforms ν into μ , is a unique optimal transportation between ν and μ , and for μ -a.e. x

and y satisfies the relations

$$\nabla V^*(\nabla V(x)) = x, \quad V(x) + V^*(\nabla V(x)) = \langle x, \nabla V(x) \rangle, \quad V(x) + V^*(y) \geq \langle x, y \rangle. \quad (1.4)$$

McCann [23] (see also [29]) refined Brenier's theorem as follows: given two Borel probability measures μ and ν on \mathbb{R}^d such that μ is absolutely continuous (no moment condition is imposed), there is a mapping T that is the gradient of a convex function and transforms μ into ν . Such a mapping is unique in the sense that any two such mappings coincide μ -a.e.

2. ESTIMATES OF NONLINEAR TRANSFORMATIONS

First we consider transformations of the standard Gaussian measure γ on \mathbb{R}^d .

Theorem 2.1. *Let $T = I + F$ be a Borel mapping on \mathbb{R}^d that is injective on a set of full γ -measure and transforms γ into a measure $g \cdot \gamma$ and let $F \in W_{loc}^{1,1}$. Suppose that for some $\kappa \geq 0$ almost everywhere one has $\langle DF(x)a, a \rangle \geq -\kappa|a|^2$ for all $a \in \mathbb{R}^d$ and that $\det_2(I + DF) > 0$ almost everywhere (i.e., $\det(I + DF) > 0$ almost everywhere). Let θ be a nonnegative increasing locally Lipschitzian function on $[0, +\infty)$ such that for some $\omega \in (0, (4\kappa)^{-1})$ almost everywhere one has $\theta' \leq \omega\theta$. Suppose also that there are a number $\alpha > 2(1 - 4\omega\kappa)^{-1}$ and a function $\Psi \in L^1(\gamma)$ such that*

$$g|\log g|\theta(\alpha|\log g|) \in L^1(\gamma), \quad \log \det_2(I + DF)\theta(|F|^2) \leq \Psi. \quad (2.1)$$

Then

$$\int_{\mathbb{R}^d} |F|^2 \theta(|F|^2) d\gamma \leq \left(\frac{1}{2} - 2\omega\kappa - \frac{1}{\alpha}\right)^{-1} \left[\int_{\mathbb{R}^d} g|\log g|\theta(\alpha|\log g|) d\gamma + \int_{\mathbb{R}^d} \Psi d\gamma \right]. \quad (2.2)$$

In particular, if $\det_2(I + DF) \leq 1$ a.e. (which is fulfilled if, e.g., the operators $I + DF(x)$ are symmetric nonnegative), then

$$\int_{\mathbb{R}^d} |F|^2 \theta(|F|^2) d\gamma \leq \left(\frac{1}{2} - 2\omega\kappa - \frac{1}{\alpha}\right)^{-1} \int_{\mathbb{R}^d} g|\log g|\theta(\alpha|\log g|) d\gamma. \quad (2.3)$$

Finally, if θ satisfies the indicated hypotheses on the whole real line, then in (2.1), (2.2) and (2.3) one can replace $\theta(\alpha|\log g|)$ by $\theta(\alpha \log g)$.

Proof. Suppose first that there is a number τ such that θ is constant on $[\tau, +\infty)$. Note that $\theta(|F|^2) \in W_{loc}^{1,1}$ and we have the equality

$$\nabla \theta(|F|^2) = 2\theta'(|F|^2)DF \cdot F \quad (2.4)$$

and the estimate

$$\theta'(|F(x)|^2)|DF(x) \cdot F(x)| \leq \sqrt{\tau} \sup_{t \in [0, \tau]} \theta'(t) \|DF(x)\|. \quad (2.5)$$

In order to justify this, we fix a ball U and take a sequence of mappings $F_j \in C_0^\infty$ that converge to F in the norm of $W^{1,1}(U, \mathbb{R}^d)$ and almost everywhere. Then $\theta(|F_j|^2) \rightarrow \theta(|F|^2)$ in $L^1(U)$. In addition, the mappings $\nabla \theta(|F_j|^2) = 2\theta'(|F_j|^2)DF_j \cdot F_j$ converge in $L^1(U, \mathbb{R}^d)$ to the mapping $2\theta'(|F|^2)DF \cdot F$, which proves that $\theta(|F|^2)$ on U belongs to $W^{1,1}(U)$ and there hold (2.4) and (2.5).

Let us fix a function $\zeta \in C_0^\infty(\mathbb{R}^d)$ such that $0 \leq \zeta(x) \leq 1$ and $|\nabla \zeta(x)| \leq M$. It follows from our hypotheses that T has a Borel version that is injective on a full measure set and has Lusin's property (N). Hence there holds the usual change of variables formula

$$\log g(T) = -\delta_\gamma F + \frac{1}{2}|F|^2 - \log \det_2(I + DF),$$

where $\delta_\gamma F(x) = \operatorname{div} F(x) - \langle F(x), x \rangle$ (see [19] and [6, Theorem 5.8.29], [5, §6.3]). Therefore, one has

$$\begin{aligned}
\int_{\mathbb{R}^d} \log g(T) \theta(|F|^2) \zeta \, d\gamma &\geq - \int_{\mathbb{R}^d} \delta_\gamma F \theta(|F|^2) \zeta \, d\gamma + \frac{1}{2} \int_{\mathbb{R}^d} |F|^2 \theta(|F|^2) \zeta \, d\gamma - \int_{\mathbb{R}^d} \Psi \zeta \, d\gamma \\
&= 2 \int_{\mathbb{R}^d} \langle DF \cdot F, F \rangle \theta'(|F|^2) \zeta \, d\gamma + \int_{\mathbb{R}^d} \theta(|F|^2) \langle F, \nabla \zeta \rangle \, d\gamma \\
&\quad + \frac{1}{2} \int_{\mathbb{R}^d} |F|^2 \theta(|F|^2) \zeta \, d\gamma - \int_{\mathbb{R}^d} \Psi \zeta \, d\gamma \\
&\geq -2\kappa \int_{\mathbb{R}^d} |F|^2 \theta'(|F|^2) \zeta \, d\gamma + \frac{1}{2} \int_{\mathbb{R}^d} |F|^2 \theta(|F|^2) \zeta \, d\gamma + \int_{\mathbb{R}^d} \theta(|F|^2) \langle F, \nabla \zeta \rangle \, d\gamma - \int_{\mathbb{R}^d} \Psi \zeta \, d\gamma \\
&\geq \left(\frac{1}{2} - 2\omega\kappa \right) \int_{\mathbb{R}^d} |F|^2 \theta(|F|^2) \zeta \, d\gamma + \int_{\mathbb{R}^d} \theta(|F|^2) \langle F, \nabla \zeta \rangle \, d\gamma - \int_{\mathbb{R}^d} \Psi \zeta \, d\gamma.
\end{aligned}$$

The second equality in the above chain of relationships holds by the integration by parts formula, which is applicable by (2.4). By using the inequality

$$x\theta(y) \leq \alpha^{-1}y\theta(y) + |x|\theta(\alpha|x|), \quad (2.6)$$

which holds for every x and every $y \geq 0$ (one has either $x \leq y/\alpha$ or $x > y/\alpha$), we obtain

$$\int_{\mathbb{R}^d} \log g(T) \theta(|F|^2) \zeta \, d\gamma \leq \frac{1}{\alpha} \int_{\mathbb{R}^d} |F|^2 \theta(|F|^2) \zeta \, d\gamma + \int_{\mathbb{R}^d} |\log g(T)| \theta(\alpha |\log g(T)|) \zeta \, d\gamma.$$

Therefore,

$$\begin{aligned}
\left(\frac{1}{2} - 2\omega\kappa - \frac{1}{\alpha} \right) \int_{\mathbb{R}^d} |F|^2 \theta(|F|^2) \zeta \, d\gamma &\leq \int_{\mathbb{R}^d} \theta(|F|^2) \langle F, \nabla \zeta \rangle \, d\gamma \\
&\quad + \int_{\mathbb{R}^d} |\log g(T)| \theta(\alpha |\log g(T)|) \zeta \, d\gamma + \int_{\mathbb{R}^d} \Psi \zeta \, d\gamma. \quad (2.7)
\end{aligned}$$

Let us take a sequence of functions $\zeta_j \in C_0^\infty(\mathbb{R}^d)$ such that

$$\sup_{j,x} |\nabla \zeta_j(x)| \leq M, \quad 0 \leq \zeta_j(x) \leq 1, \quad \text{and} \quad \zeta_j(x) = 1 \quad \text{if} \quad |x| \leq j.$$

Letting $\zeta = \zeta_j$ in (2.7), we observe that as $j \rightarrow \infty$, the first integral on the right tends to zero, because $\theta(|F|^2)|F|$ is a bounded function, and the sequence $\{|\nabla \zeta_j|\}$ is uniformly bounded and tends to zero pointwise. Hence (2.7) holds with $\zeta = 1$. The change of variables formula yields

$$\left(\frac{1}{2} - 2\omega\kappa - \frac{1}{\alpha} \right) \int_{\mathbb{R}^d} |F|^2 \theta(|F|^2) \, d\gamma \leq \int_{\mathbb{R}^d} g |\log g| \theta(\alpha |\log g|) \, d\gamma + \int_{\mathbb{R}^d} \Psi \, d\gamma,$$

which completes the proof in the case under consideration. In the general case let $\theta_j(t) = \theta(t)$ if $t \leq j$ and $\theta_j(t) = \theta(j)$ if $t \geq j$. Since $\theta_j \leq \theta$, estimate (2.2) for every θ_j in place of θ yields the same estimate for θ .

If one has $\det_2(I + DF) \leq 1$ a.e., then we take $\Psi = 0$. Finally, if θ satisfies the indicated hypotheses on the whole real line, then in inequality (2.6) one can replace $\theta(\alpha|x|)$ by $\theta(\alpha x)$, which will lead to replacing $\theta(\alpha|\log g|)$ by $\theta(\alpha \log g)$ in (2.1), (2.2) and (2.3). \square

Corollary 2.1. *Let $T_j = I + F_j$ be Borel mappings on \mathbb{R}^d that are injective on full measure sets and transform γ into the measures $g_j \cdot \gamma$ such that the densities g_j converge in measure γ to a probability density g with respect to γ and the mappings F_j converge in measure γ to a mapping F . Suppose that $F_j \in W_{loc}^{1,1}$, for some $\kappa \geq 0$ almost everywhere*

one has $\langle DF_j(x)a, a \rangle \geq -\kappa|a|^2$ for all $a \in \mathbb{R}^d$ and also $0 < \det_2(I + DF_j) \leq 1$ almost everywhere. Let θ be a nonnegative increasing locally Lipschitzian function on $[0, +\infty)$ such that for some $\omega \in (0, (4\kappa)^{-1})$ almost everywhere one has $\theta' \leq \omega\theta$. Suppose that $\sup_j \|g|\log g|\theta(\alpha|\log g)|\|_{L^1(\gamma)} < \infty$, where $\alpha > 2(1 - 4\omega\kappa)^{-1}$. Then $\gamma \circ (I + F)^{-1} = g \cdot \gamma$ and

$$\int_{\mathbb{R}^d} |F|^2 \theta(|F|^2) d\gamma \leq \left(\frac{1}{2} - 2\omega\kappa - \frac{1}{\alpha}\right)^{-1} \sup_j \int_{\mathbb{R}^d} g_j |\log g_j| \theta(\alpha |\log g_j|) d\gamma.$$

Corollary 2.2. *Let $T = I + F$ be the optimal transportation between γ and $g \cdot \gamma$ for the cost function $h(x, y) = |x - y|^2$. Let θ be the same as in the theorem, where $\omega < 1/4$. Suppose that $g|\log g|\theta(\alpha|\log g|) \in L^1(\gamma)$, where $\alpha > 2(1 - 4\omega)^{-1}$. Then*

$$\int_{\mathbb{R}^d} |F|^2 \theta(|F|^2) d\gamma \leq \left(\frac{1}{2} - 2\omega - \frac{1}{\alpha}\right)^{-1} \int_{\mathbb{R}^d} g |\log g| \theta(\alpha |\log g|) d\gamma.$$

Proof. Let us take a sequence of smooth positive probability densities $g_j \in C_b^\infty(\mathbb{R}^d)$ with respect to the measure γ such that $g_j \rightarrow g$ a.e. and the integrals of $g_j |\log g_j| \theta(\alpha |\log g_j|)$ against γ converge to the integral of $g |\log g| \theta(\alpha |\log g|)$. The corresponding optimal transportations $T_j = I + F_j$ converge in measure γ to the optimal transportation $T = I + F$ for g according to the lemma below. In addition, one has $I + DF_j > 0$ and $0 < \det_2(I + DF_j) \leq 1$. Finally, by Caffarelli's regularity theory the mappings T_j are smooth (see [29, p. 140]). \square

As in Theorem 2.1, if θ satisfies the indicated conditions on the whole real line, then in both corollaries one can replace $\theta(\alpha|\log g|)$ by $\theta(\alpha \log g)$.

Lemma 2.1. *Let μ be a Borel probability measure on \mathbb{R}^d with a density that is locally separated from zero. Suppose that a sequence of Borel probability measures ν_j on \mathbb{R}^d converges weakly to a Borel probability measure ν . Let V_j and V be finite convex functions such that $\mu \circ (\nabla V_j)^{-1} = \nu_j$ and $\mu \circ (\nabla V)^{-1} = \nu$. Suppose that*

$$\sup_j \int_{\mathbb{R}^d} |\nabla V_j| d\mu < \infty.$$

Then the mappings ∇V_j converge μ -almost everywhere to ∇V . The same is true if the hypotheses are fulfilled on an open convex set in \mathbb{R}^d .

Proof. First we verify convergence of ∇V_j to ∇V in measure μ . Let U_k be the closed ball of radius k centered at the origin. By adding constants we may assume that $\int_{U_1} V_j(x) dx = 0$ for all j . It follows from our hypothesis and the compactness of the embedding $W^{1,1}(U_1) \rightarrow L^1(U_1)$ that the sequence of functions $V_j|_{U_1}$ contains a subsequence convergent in $L^1(U_1)$. We denote this subsequence again by $\{V_j\}$. For the same reason, for every fixed k , the sequence of functions $V_j - c_{k,j}$, where $c_{k,j} = \int_{U_k} V_j(x) dx$, has a subsequence $\{V_{j_n} - c_{k,j_n}\}$ that converges in $L^1(U_k)$. Then the sequence $\{c_{k,j_n}\}$ converges, which yields convergence of $\{V_{j_n}\}$ in $L^1(U_k)$. By a diagonal procedure we can choose a subsequence denoted by $\{W_j\}$ that converges in $L^1(U_k)$ for every k and converges almost everywhere. It is readily seen that if a sequence of convex functions on the real line converges almost everywhere, then it converges pointwise. By applying Fubini's theorem we conclude that the sequence $\{W_j\}$ converges pointwise to a convex function W . It is well-known (see [27, Theorems 24.5 and 25.4]) that then the sequence of mappings ∇W_j , which are defined almost everywhere,

converges almost everywhere to ∇W . Hence this sequence converges in $L^1(U_k)$ for every k , because it is uniformly integrable on every ball U_k with Lebesgue measure by the estimate

$$|\partial_{x_i} \Psi(x)| \leq |\Psi(x + e_i)| + |\Psi(x)| + |\Psi(x - e_i)|,$$

which holds for every convex function Ψ and every i . The measures $\mu \circ (\nabla W_j)^{-1}$ converge weakly to the measure $\mu \circ (\nabla W)^{-1}$. It follows that $\nu = \mu \circ (\nabla W)^{-1}$. Therefore, $\nabla V = \nabla W$ a.e., because a convex function whose gradient transforms μ into ν is uniquely defined up to a constant. Finally, our reasoning applies to every subsequence in the initial sequence, whence our claim about convergence in measure μ follows.

In order to see that we have almost everywhere convergence of the whole sequence, we consider the sequence $\{\nabla V_j\}$ that converges in measure μ to the mapping ∇V (as shown above) and assume that the functions V_j and V have zero integrals over U_1 , which can be achieved by adding constants. Then the sequence $\{V_j\}$ converges to V in measure μ . If the sequence $\{V_j(x_0)\}$ does not converge to $V(x_0)$ at some point x_0 , then it contains a subsequence $\{V_{j_n}(x_0)\}$ separated from $V(x_0)$. By the above reasoning one can find a further subsequence in $\{V_{j_n}\}$, denoted by the same symbol, that converges pointwise to a convex function W that differs from V . This gives a contradiction, because $V(x) = W(x)$ a.e., since $\{V_{j_n}\}$ converges in measure to W and to V . As noted above, pointwise convergence of the functions V_j yields almost sure convergence of their gradients. \square

Corollary 2.3. *Let μ be a Borel probability measure on \mathbb{R}^d with a density that is locally separated from zero. Suppose that a sequence of Borel probability measures ν_j converges weakly to a Borel probability measure ν and that the measures μ, ν_j and ν have finite second moments. Let T_j be the optimal transportation between μ and ν_j for the cost function $h(x, y) = |x - y|^2$. If*

$$\sup_j \int_{\mathbb{R}^d} |T_j| d\mu < \infty,$$

then the mappings T_j converge μ -a.e. to the optimal transportation between μ and ν .

Example 2.1. (i) Suppose that a Borel probability measure μ has a density locally separated from zero and a finite second moment. Let $\mathcal{PD}_2(\mu)$ denote the class of all probability densities g with respect to μ such that the function $g(x)|x|^2$ is μ -integrable. For every $g \in \mathcal{PD}_2(\mu)$, let T_g be the optimal transportation between μ and $g \cdot \mu$ for the cost function $h(x, y) = |x - y|^2$. Suppose also that μ satisfies the Talagrand inequality

$$\int_{\mathbb{R}^d} |T_g(x) - x|^2 \mu(dx) \leq c \text{Ent}_\mu(g)$$

for every $g \in \mathcal{PD}_2(\mu)$. Let $g_j \in \mathcal{PD}_2(\mu)$ converge in measure μ to $g \in \mathcal{PD}_2(\mu)$ and let $\sup_j \text{Ent}_\mu(g_j) < \infty$. Then $|T_{g_j} - T_g| \rightarrow 0$ in $L^r(\mu)$ for every $r < 2$ and $T_{g_j} \rightarrow T_g$ almost everywhere. It is well-known that the hypotheses in this example are fulfilled for any measure μ that has a density locally separated from zero and satisfies the logarithmic Sobolev inequality, e.g., for the standard Gaussian measure.

(ii) Let μ be Lebesgue measure on the cube $K = [0, 1]^d$. For every Borel probability measure ν on K , let $T_\nu: K \rightarrow K$ be the optimal transportation between μ and ν . It is known that there is a Borel (even continuous) mapping $\eta: [0, 1] \rightarrow [0, 1]^d$ that transforms Lebesgue measure λ on $[0, 1]$ to Lebesgue measure on $[0, 1]^d$ (see [6, Ch. 9]). Then the mapping $\nu \mapsto \xi_\nu := T_\nu \circ \eta$ defines a parameterization of Borel probability measures on K by mappings from $[0, 1]$ to K with the following continuity property: $\lambda \circ \xi_\nu^{-1} = \nu$ for every ν and $\xi_{\nu_j} \rightarrow \xi_\nu$ λ -a.e. whenever $\nu_j \rightarrow \nu$ weakly. We recall that such a representation is called Skorohod's parameterization of probability measures on K and can be obtained

for every Polish space. This result for general Polish spaces is due to Blackwell, Dubbins, and Fernique; see the corresponding references and discussion in [7] and [6, Ch. 8], where it is explained how the partial case of $[0, 1]^3$ yields the general result.

The theorem can be used for establishing the integrability of the functions $|F|^r$ and $\exp(\omega|F|^2)$ under appropriate integrability assumptions on $g|\log g|^p$ or g^p .

Example 2.2. (i) Let $\theta(t) = \exp(\omega t)$. Under the assumption that almost everywhere $\langle DF(x)a, a \rangle \geq -|a|^2$ for all $a \in \mathbb{R}^d$ and $0 < \det_2(I + DF) \leq 1$ (which is fulfilled if $I + DF(x)$ is a nonnegative symmetric operator), we obtain that

$$\int_{\mathbb{R}^d} |F|^2 \exp(\omega|F|^2) d\gamma \leq \left(\frac{1}{2} - 2\omega - \frac{1}{\alpha}\right)^{-1} \int_{\mathbb{R}^d} g^{1+\alpha} |\log g| d\gamma \quad (2.8)$$

if $\omega < 1/4 - 1/(2\alpha)$. Hence $|F|^2 \exp(\omega|F|^2) \in L^1(\gamma)$ whenever $\omega < (p-3)(4p-4)^{-1}$ and $g \in L^p(\gamma)$.

(ii) Set $\alpha = 3$ and $\omega = 1/20$. Let $\theta(t) = (20r)^r + t^r$, where $r \geq 1$. Then $\theta' \leq \omega\theta$, and under the same assumptions on DF as in (i), we obtain that there is a number $C(r)$ that depends only on r such that

$$\int_{\mathbb{R}^d} |F|^{2+r} d\gamma \leq C(r) \int_{\mathbb{R}^d} g |\log g|^{r+1} d\gamma.$$

If $0 < r < 1$, then in a similar manner we obtain the estimate

$$\int_{\mathbb{R}^d} |F|^{2+r} d\gamma \leq C(r) \int_{\mathbb{R}^d} g [|\log g|^{r+1} + 1] d\gamma$$

by using the function $\theta(t) = 20r + \max(1, |t|^r)$. Finally, taking $\theta = 1$, we arrive at the estimate

$$\int_{\mathbb{R}^d} |F|^2 d\gamma \leq C \int_{\mathbb{R}^d} g |\log g| d\gamma,$$

which, however, can be easily strengthened to Talagrand's estimate

$$\int_{\mathbb{R}^d} |F|^2 d\gamma \leq C \int_{\mathbb{R}^d} g \log g d\gamma$$

by returning directly to the proof of the theorem with $\theta = 1$.

A closer look at the proof shows that it applies to more general measures. Suppose that μ is a probability measure on \mathbb{R}^d with density $f = \exp(-V)$, where V is a twice differentiable convex function.

Theorem 2.2. *Let $D^2V \geq \sigma I$ with some $\sigma > 0$ and let $T = I + F$ be a Borel mapping on \mathbb{R}^d that is injective on a full measure set and transforms μ into a measure $g \cdot \mu$. Suppose that F and θ satisfy the same hypotheses as in Theorem 2.1. Assume that there are a number $\alpha > 0$ with $1/\alpha < \sigma/2 - 2\omega\kappa$ and a function $\Psi \in L^1(\mu)$ such that*

$$g |\log g| \theta(\alpha |\log g|) \in L^1(\mu), \quad \log \det_2(I + DF) \theta(|F|^2) \leq \Psi.$$

Then

$$\int_{\mathbb{R}^d} |F|^2 \theta(|F|^2) d\mu \leq \left(\frac{\sigma}{2} - 2\omega\kappa - \frac{1}{\alpha}\right)^{-1} \left[\int_{\mathbb{R}^d} g |\log g| \theta(\alpha |\log g|) d\mu + \int_{\mathbb{R}^d} \Psi d\mu \right].$$

If $\det_2(I + DF) \leq 1$ a.e., then

$$\int_{\mathbb{R}^d} |F|^2 \theta(|F|^2) d\mu \leq \left(\frac{\sigma}{2} - 2\omega\kappa - \frac{1}{\alpha}\right)^{-1} \int_{\mathbb{R}^d} g |\log g| \theta(\alpha |\log g|) d\mu.$$

Finally, if θ satisfies the indicated conditions on the whole real line, then one can replace $\theta(\alpha|\log g|)$ by $\theta(\alpha \log g)$.

Proof. The reasoning is similar. Note that $\delta_\mu F = \operatorname{div} F - \langle F, \nabla V \rangle$. The change of variables formula takes the form

$$\log g(T) = -\delta_\mu F + V(T) - V - \langle \nabla V, F \rangle - \log \det_2(I + DF).$$

It follows by our hypotheses that

$$V(x + F(x)) - V(x) - \langle \nabla V(x), F(x) \rangle \geq \sigma |F(x)|^2.$$

Therefore, in the situation of the first step of the proof above a similar reasoning yields the estimate

$$\int_{\mathbb{R}^d} \log g(T) \theta(|F|^2) d\mu \geq \left(\frac{\sigma}{2} - 2\omega\kappa \right) \int_{\mathbb{R}^d} |F|^2 \theta(|F|^2) d\mu - \int_{\mathbb{R}^d} \Psi d\mu.$$

The rest of the proof is completely analogous to that of the previous theorem. \square

Remark 2.1. For certain special functions θ our general estimate (2.2) can be slightly improved with the same proof. For example, suppose that in the situation of Theorem 2.1 one has $\theta(t) = \exp(\omega t)$ (as in Example 2.2(i)). Let us repeat the proof of the theorem by using the inequality $xy \leq e^{x-1} + y \log y$ in place of inequality (2.6). Then, for any fixed $\alpha > 0$ and $\omega < \frac{\alpha}{1+2\alpha\kappa}$, we obtain

$$\begin{aligned} \alpha \log g(T) \theta(|F|^2) &\leq \frac{1}{e} \exp[\alpha \log g(T)] + \theta(|F|^2) \log \theta(|F|^2) \\ &\leq \frac{1}{e} \exp[\alpha \log g(T)] + \omega |F|^2 \theta(|F|^2), \end{aligned}$$

since $\log \theta(t) \leq \omega t$ (in the general case $\log \theta(t) \leq \omega t + \log \theta(0)$ if $\theta(0) > 0$). Finally, we have

$$\int_{\mathbb{R}^d} |F|^2 \exp(\omega |F|^2) d\gamma \leq \frac{2}{e(\alpha - 4\omega\kappa\alpha - 2\omega)} \left[\int_{\mathbb{R}^d} g^{1+\alpha} d\gamma + e\alpha \int_{\mathbb{R}^d} \Psi d\gamma \right], \quad (2.9)$$

which is more precise than (2.8). Indeed, if $\kappa = 1$, then $\exp(\omega |F|^2) \in L^1(\gamma)$ provided that $g \in L^{1+\alpha}(\gamma)$ and $\omega < \alpha/(4\alpha + 2)$, whereas (2.8) requires that $\omega < (\alpha - 2)/(4\alpha)$, in particular, in order to have $\omega > 0$ one must take $\alpha > 2$. However, as $\alpha \rightarrow \infty$ both bounds on ω converge to $1/4$. The same concerns Theorem 2.2.

It is worth noting that in general the function $\exp(\omega |F|^2)$ may fail to be γ -integrable for $\omega > 1/2$ even if g is bounded. It suffices to consider the functions $T(x) := x - sx$ with $s \in (0, 1)$ on the real line with the standard Gaussian measure γ . The restriction $s < 1$ is needed here to ensure our hypothesis that $T' \geq 0$. Without that hypothesis, the function $\exp(\omega |F|^2)$ need not be γ -integrable even for $\omega = 1/8$ in the case of a mapping preserving the measure γ (example: $T(x) := x - 2x = -x$).

Remark 2.2. A closer look at the proof of Theorem 2.1 shows that one can relax the hypotheses on the mapping T that transforms γ into $g \cdot \gamma$ as follows. It suffices to assume only that there holds the change of variables formula

$$G(T(x)) = \exp \left[-\operatorname{Trace} \Lambda(x) + \langle F(x), x \rangle + |F(x)|^2/2 \right] \left(\det_2(I + \Lambda(x)) \right)^{-1},$$

where $x \mapsto \Lambda(x)$ is some locally integrable operator-valued mapping such that $\delta_\gamma F(x) := \operatorname{Trace} \Lambda(x) - \langle F(x), x \rangle$ satisfies the condition

$$\int_{\mathbb{R}^d} \delta_\gamma F \theta(|F|^2) d\gamma \leq -2\kappa \int_{\mathbb{R}^d} \theta'(|F|^2) |F|^2 d\gamma.$$

A realization of this approach might use the inequality

$$\int_{\mathbb{R}^d} \varphi(x) \text{Trace}[I + \Lambda(x)] dx \leq - \int_{\mathbb{R}^d} \langle \nabla \varphi(x), T(x) \rangle dx$$

for all nonnegative $\varphi \in C_0^\infty(\mathbb{R}^d)$ established in [14] for optimal transportations corresponding to sufficiently general cost functions (then Λ is Alexandroff's second derivative of some potential). It would be interesting to extend the results from [14] to more general cost functions.

Remark 2.3. (i) Fernique's result [16] has been established under the same bound on ω as in estimate (2.9). However, as noted above, Fernique considered somewhat different mappings, so that it was not clear whether his estimate would hold for optimal transportations.

(ii) Let us consider the inverse mapping $S = T^{-1}$ which is the optimal transportation between $g \cdot \gamma$ and γ in the case of cost function $h(x, y) = |x - y|^2$. Note that S transforms γ into $\frac{1}{g(T)} \cdot \gamma$. Hence

$$\frac{1}{2} \int_{\mathbb{R}^d} |S(x) - x|^2 \gamma(dx) \leq \int_{\mathbb{R}^d} \frac{1}{g(T)} \log \frac{1}{g(T)} d\gamma = \int_{\mathbb{R}^d} \log \frac{1}{g} d\gamma.$$

In the same way as above one can show that if $1/g \in L^\tau(\gamma)$ for some $\tau > 0$, then there exists $\alpha > 0$ such that $\exp(\alpha|S(x) - x|^2) \in L^1(\gamma)$ and if $\log g \in L^{1+\tau}(\gamma)$, then $|S(x) - x|^{2(1+\beta)} \in L^1(\gamma)$ for some $\beta > 0$.

(iii) Let γ be a centered Radon Gaussian measure on a locally convex space X and let H be its Cameron–Martin space with the norm $|\cdot|_H$ (see [5]). One can assume that γ is the countable product of the standard Gaussian measures on the real line and is defined on \mathbb{R}^∞ , then $H = l^2$ with its usual norm. Suppose that g is a probability density with respect to γ such that $g \log g \in L^1(\gamma)$. As shown in [17], there exists a Borel mapping $T: X \rightarrow X$ of the form $T(x) = x + F(x)$, where $F: X \rightarrow H$, such that $\gamma \circ T^{-1} = g \cdot \gamma$ and

$$\int_X |F(x)|_H^2 \gamma(dx) = \inf \int_X |R(x)|_H^2 \gamma(dx),$$

where inf is taken over all Borel mappings $I + R$ such that $R: X \rightarrow H$ and $\gamma \circ (I + R)^{-1} = g \cdot \gamma$. By using the methods of [17] and [10] one can show that Corollary 2.2 extends to this situation and one has

$$\int_X |F|^2 \theta(|F|_H^2) d\gamma \leq \left(\frac{1}{2} - 2\omega - \frac{1}{\alpha} \right)^{-1} \int_X g |\log g| \theta(\alpha |\log g|) d\gamma.$$

Let us now consider the cost function $h(x, y) = \frac{1}{p}|x - y|^p$, where $1 < p \leq 2$. One can easily prove that the conjugated function for $h_0(x) := \frac{1}{p}|x|^p$ equals

$$h_0^*(x) = \frac{p-1}{p^{\frac{1}{p-1}}} |x|^q,$$

where $\frac{1}{p} + \frac{1}{q} = 1$. It is known that the corresponding optimal transportation T between γ and $g \cdot \gamma$ has the form

$$T(x) = x + \nabla h_0^*(\nabla \Phi(x))$$

for some h -concave locally Lipschitzian potential Φ (see [18] or [29, p. 92]). Here the h -concavity of Φ means that $\Phi = (\Phi^h)^h$, where $\psi^h(x) = \inf_y [h(x, y) - \psi(y)]$. Renormalizing Φ one can assume that

$$T(x) = x + |\nabla \Phi(x)|^{q-2} \nabla \Phi(x).$$

Let $A_p(x) := D[|x|^{q-2}x]$. One has

$$(A_p(x))_{ij} = |x|^{q-2} \left(\delta_{ij} + (q-2) \frac{x_i x_j}{|x|^2} \right).$$

Suppose additionally that $\nabla\Phi \in W_{loc}^{1,1}$. Then one can show that the above mentioned h -concavity of Φ yields

$$A_p^{-1}(\nabla\Phi) + D^2\Phi \geq 0.$$

This is verified in [29, p. 103] under more restrictive assumptions on h (which are not satisfied in our case), but the additional differentiability of $\nabla\Phi$ enables one to arrive at the same conclusion. It follows from the estimate $|x|^{q-2}I \leq A_p(x) \leq (q-1)|x|^{q-2}I$ that

$$D^2\Phi \geq -|\nabla\Phi|^{2-q}I.$$

Theorem 2.3. *Let $T = I + F$ be the optimal transportation between γ and $g \cdot \gamma$ for the cost function $h(x, y) = \frac{1}{p}|x - y|^p$, where $1 < p \leq 2$. Suppose additionally that $\nabla\Phi \in W_{loc}^{1,1}$. Then the following assertions are true.*

(i) *For every $\beta > 0$ one has*

$$\int_{\mathbb{R}^d} |F|^{2(\beta+1)} d\gamma \leq 2^{\beta+1} \int_{\mathbb{R}^d} |2\beta(q-1) + \log g|^{\beta+1} g d\gamma. \quad (2.10)$$

(ii) *Whenever $\alpha > 0$ and $0 < \omega < \frac{\alpha}{2+4(q-1)\alpha}$ one has*

$$\int_{\mathbb{R}^d} |F|^2 \exp(\omega|F|^2) d\gamma \leq \frac{2}{e(\alpha - 4\alpha\omega(q-1) - 2\omega)} \int_{\mathbb{R}^d} g^{1+\alpha} d\gamma. \quad (2.11)$$

In particular, $\exp(\omega|F|^2) \in L^1(\gamma)$ if $g \in L^{1+\alpha}(\gamma)$.

(iii) *If θ is a nonnegative increasing locally Lipschitzian function on $[0, +\infty)$ such that $\theta' \leq \omega\theta$ for some $\omega < (\alpha - 1)(4\alpha(q-1))^{-1}$, where $\alpha > 1$, then*

$$\int_{\mathbb{R}^d} |F|^2 \theta(|F|^2) d\gamma \leq 2 \left(1 - 4\omega q + 4\omega - \frac{1}{\alpha} \right)^{-1} \int_{\mathbb{R}^d} g |\log g| \theta(\alpha |\log g|) d\gamma. \quad (2.12)$$

Proof. The reasoning is much the same as above. We only observe that $F = |\nabla\Phi|^{q-2} \nabla\Phi$, hence $|F| = |\nabla\Phi|^{q-1}$. Therefore, in case (iii) one has

$$\begin{aligned} \langle \nabla\theta(|\nabla\Phi|^{2q-2}), F \rangle &= (2q-2)\theta'(|\nabla\Phi|^{2q-2}) |\nabla\Phi|^{3q-6} \langle D^2\Phi \cdot \nabla\Phi, \nabla\Phi \rangle \\ &\geq -(2q-2)\omega\theta(|\nabla\Phi|^{2q-2}) |\nabla\Phi|^{2q-2} = -(2q-2)\omega |F|^2 \theta(|F|^2). \end{aligned}$$

It is also important that $0 \leq \det_2 DT \leq 1$, since $DT = A_p(\nabla\Phi)(A_p^{-1}(\nabla\Phi) + D^2\Phi)$ is a product of two nonnegative matrices. The rest of the proof repeats the reasoning from Theorem 2.1 applied to the estimation of the integral of the function $\log g(T)\theta(|\nabla\Phi|^{2(q-1)})$, which yields an inequality for the integral of the function $|\nabla\Phi|^{2(q-1)}\theta(|\nabla\Phi|^{2(q-1)}) = |F|^2\theta(|F|^2)$. \square

As noted above, the assumption $\nabla\Phi \in W_{loc}^{1,1}$ is rather restrictive and seems to be unnecessary. It is likely that one can get rid of this assumption by approximating $\nabla\Phi$ in the general case by better differentiable transportations (which, however, still needs justification) or by developing the idea of Remark 2.2 in the present situation.

3. THE CASE OF MEASURES SATISFYING THE LOGARITHMIC SOBOLEV INEQUALITY

In this section we consider only the quadratic cost function. We show that the result of the previous section holds (qualitatively) also for measures satisfying the logarithmic Sobolev inequality. The estimates which we obtain here for an arbitrary measure μ satisfying the logarithmic Sobolev inequality are in general weaker than in the Gaussian case. However, the same relation holds: if $g \in L^{1+\varepsilon}(\mu)$ for some $\varepsilon > 0$, then the function $\exp(\omega|T(x) - x|^2)$ is integrable for sufficiently small ω . A similar result for power estimates is established.

We shall apply the following remarkable result proved by Bobkov and Götze in [4]: for every measure satisfying (1.1) there holds the inequality

$$\int_{\mathbb{R}^d} \exp\left[f - \int_{\mathbb{R}^d} f d\mu\right] d\mu \leq \int_{\mathbb{R}^d} \exp\left(C|\nabla f|^2\right) d\mu. \quad (3.1)$$

In the Gaussian case, this inequality was proved in [25] with a worse constant (see also [5, §5.6]).

For the proof of the main result in this section we need an auxiliary estimate. Suppose μ satisfies (1.1). Let V be a convex function such that $T = \nabla V$ is an optimal transportation between μ and $g \cdot \mu$. Note that $|\nabla V| \in L^2(\mu)$, because μ and ν possess finite second moments, so $|T| \in L^2(\mu)$. The functions V and V^* are convex, hence $V(x) \geq M_1 - M_2|x|$ and $V^*(x) \geq M_1 - M_2|x|$ for some numbers M_1 and M_2 . This shows that the functions $\exp\left[t(x^2/2 - V(x))\right]$ and $\exp\left[t(x^2/2 - V^*(x))\right]$ are μ -integrable if $t < C^{-1}$. Moreover, the following inequality holds true:

$$\int_{\mathbb{R}^d} \exp\left[\frac{1}{C}\left(\frac{x^2}{2} - V^*(x)\right)\right] \mu(dx) \leq \exp\left[\frac{1}{C} \int_{\mathbb{R}^d} \left(V(x) - \frac{x^2}{2}\right) \mu(dx)\right]. \quad (3.2)$$

Justification is similar to the proof of the infimum-convolution inequality in [22, Ch. 6]. For notational simplicity we assume that $C = 1$. Given a real function f on \mathbb{R}^d let

$$\tilde{f}(x) = \inf_y [f(y) + |x - y|^2/2].$$

It is shown in [22, Ch. 6] that for every bounded measurable function f one has

$$\int_{\mathbb{R}^d} \exp \tilde{f} d\mu \leq \exp \int_{\mathbb{R}^d} f d\mu. \quad (3.3)$$

The desired result is formally obtained by letting $f(x) := V(x) - x^2/2$, because then $\tilde{f}(x) - x^2/2 = -V^*(x)$. However, a reduction to bounded functions is not straightforward and it is easier to modify the corresponding reasoning. Thus, given a function $f \in L^1(\mu)$, we have to show that $\exp \tilde{f} \in L^1(\mu)$ and (3.3) holds. We may assume that the integral of f against μ is zero. Let

$$g_n := \exp \min(\tilde{f}, n) \left(\int_{\mathbb{R}^d} \exp \min(\tilde{f}, n) d\mu \right)^{-1}.$$

Then $\nu_n := g_n \cdot \mu$ is a probability measure and $\tilde{f} \in L^1(\nu_n)$, since $\tilde{f} \leq f$ and so $\max(\tilde{f}, 0) \in L^1(\mu)$, and if $\tilde{f}(x) < 0$, then $\tilde{f}(x) \exp \tilde{f}(x) \leq 1$. It is known (see [22, Ch. 6]) that the logarithmic Sobolev inequality yields the estimate $K(\mu, \nu) \leq \text{Ent}_\mu(d\nu/d\mu)$, where $K(\mu, \nu) := K(\mu, \nu, h)$ with $h(x, y) = |x - y|^2/2$. Since $\tilde{f}(x) \leq f(y) + |x - y|^2/2$, we have

by the Kantorovich–Rubinshtein inequality (see [29, p. 19])

$$\int_{\mathbb{R}^d} \tilde{f} d\nu_n = \int_{\mathbb{R}^d} \tilde{f} d\nu_n - \int_{\mathbb{R}^d} f d\mu \leq K(\mu, \nu_n) \leq \text{Ent}_\mu(g_n).$$

Hence

$$\int_{\mathbb{R}^d} \tilde{f} \exp \min(\tilde{f}, n) d\mu + \log \|\exp \min(\tilde{f}, n)\|_{L^1(\mu)} \leq \int_{\mathbb{R}^d} \min(\tilde{f}, n) \exp \min(\tilde{f}, n) d\mu.$$

As $\min(\tilde{f}, n) \leq \tilde{f}$, we conclude that $\|\exp \min(\tilde{f}, n)\|_{L^1(\mu)} \leq 1$. By Fatou's theorem this yields the integrability of $\exp \tilde{f}$ and the estimate $\|\exp \tilde{f}\|_{L^1(\mu)} \leq 1$.

Theorem 3.1. *Let μ satisfy (1.1) and let $T(x) = x + \nabla\Phi(x)$ be the optimal transportation between μ and $g \cdot \mu$. Then*

$$\sqrt{\frac{\omega}{4C}} (1 - \sqrt{16\omega C}) \int_{\mathbb{R}^d} |\nabla\Phi|^2 \exp(\omega|\nabla\Phi|^2) d\mu + \int_{\mathbb{R}^d} \exp(\omega|\nabla\Phi|^2) d\mu \leq \left(\int_{\mathbb{R}^d} g^p d\mu \right)^{\frac{1}{p}},$$

where $\omega \geq 0$, $16\omega C < 1$, and $p = \frac{1}{1-\sqrt{\omega C}}$. In particular, the function $|\nabla\Phi|^2 \exp(\omega|\nabla\Phi|^2)$ belongs to $L^1(\mu)$.

Proof. As we have already noted above, (1.1) yields the inclusion $\exp(\varepsilon x^2) \in L^1(\mu)$ for any $\varepsilon < (2C)^{-1}$. Since $g \in L^p(\mu)$, the function $|x|^2 g(x)$ is μ -integrable. By the change of variables formula we have $|T|^2 \in L^1(\mu)$, i.e., $|\nabla\Phi| \in L^2(\mu)$. It is well-known that on account of (1.1) this yields $\Phi \in L^2(\mu)$. Since the potential Φ is defined up to a constant, one can choose Φ in such a way that $\int_{\mathbb{R}^d} \Phi d\mu = 0$. Recall that $\Phi(x) + \frac{x^2}{2} = V(x)$ for some convex function V . Let us fix a number $r > 0$. Inequality (3.2) yields

$$\int_{\mathbb{R}^d} \exp\left[\frac{1}{C}\left(\frac{x^2}{2} - V^*(x)\right)\right] \mu(dx) \leq \exp\left[\frac{1}{C} \int_{\mathbb{R}^d} \left(V(x) - \frac{x^2}{2}\right) \mu(dx)\right] = 1. \quad (3.4)$$

By using (1.4) we obtain

$$\begin{aligned} \frac{r}{2} \int_{\mathbb{R}^d} |\nabla V(x) - x|^2 \exp(\omega|\nabla V(x) - x|^2) \mu(dx) = \\ r \int_{\mathbb{R}^d} \left[\frac{x^2}{2} - V(x) - V^*(\nabla V(x)) + \frac{|\nabla V(x)|^2}{2} \right] \exp(\omega|\nabla V(x) - x|^2) \mu(dx). \end{aligned}$$

By the inequality $xy \leq x \log x - x + e^y$ the latter does not exceed

$$\begin{aligned} \int_{\mathbb{R}^d} \exp\left[r\frac{x^2}{2} - rV(x)\right] \mu(dx) + \int_{\mathbb{R}^d} \exp\left[r\frac{|\nabla V(x)|^2}{2} - rV^*(\nabla V(x))\right] \mu(dx) \\ + 2\omega \int_{\mathbb{R}^d} |\nabla V(x) - x|^2 \exp(\omega|\nabla V(x) - x|^2) \mu(dx) - 2 \int_{\mathbb{R}^d} \exp(\omega|\nabla V(x) - x|^2) \mu(dx). \end{aligned}$$

Hence

$$\begin{aligned} \left(\frac{r}{2} - 2\omega\right) \int_{\mathbb{R}^d} |\nabla V(x) - x|^2 \exp(\omega|\nabla V(x) - x|^2) d\mu + 2 \int_{\mathbb{R}^d} \exp(\omega|\nabla V(x) - x|^2) d\mu \\ \leq \int_{\mathbb{R}^d} \exp\left[r\frac{x^2}{2} - rV(x)\right] \mu(dx) + \int_{\mathbb{R}^d} \exp\left[r\frac{x^2}{2} - rV^*(x)\right] g(x) \mu(dx). \end{aligned}$$

The expression on the right is finite by Hölder's inequality

$$\int_{\mathbb{R}^d} \exp\left[r\frac{x^2}{2} - rV^*(x)\right] g(x) \mu(dx) \leq \left(\int_{\mathbb{R}^d} \exp\left[rq\frac{x^2}{2} - rqV^*(x)\right] \mu(dx) \right)^{\frac{1}{q}} \left(\int_{\mathbb{R}^d} g^p d\mu \right)^{\frac{1}{p}},$$

where $\frac{1}{p} + \frac{1}{q} = 1$ and $rq \leq \frac{1}{C}$. By (3.4) we have

$$\int_{\mathbb{R}^d} \exp\left[r\frac{x^2}{2} - rV^*(x)\right] g(x) \mu(dx) \leq \|g\|_{L^p(\mu)}.$$

Applying (3.1) we find

$$\int_{\mathbb{R}^d} \exp\left[r\frac{x^2}{2} - rV(x)\right] \mu(dx) \leq \int_{\mathbb{R}^d} \exp\left[Cr^2|x - \nabla V(x)|^2\right] \mu(dx).$$

Choosing $r^2 = \omega/C$ we obtain the desired result. \square

We note that the requirement $\omega < (16C)^{-1}$ is more restrictive than the conditions in Section 2.

Theorem 3.2. *For every $\beta > 0$, there exist numbers $A = A(\beta, C)$ and $B = B(\beta, C)$ depending only on β and C such that if μ satisfies (1.1) and $T(x) = x + \nabla\Phi(x)$ is the optimal transportation between μ and $g \cdot \mu$, then*

$$\int_{\mathbb{R}^d} |\nabla\Phi|^{2(1+\beta)} d\mu \leq A + B \left(\int_{\mathbb{R}^d} |\log g|^{1+\beta} g d\mu \right)^2.$$

Proof. We observe that the function $|T(x)|^{2+2\beta}$ is μ -integrable by the change of variables formula and the μ -integrability of the function $|x|^{2+2\beta}g(x)$, which follows by the estimate

$$|x|^{2+2\beta}g(x) \leq |x|^{2+2\beta} \exp(\varepsilon x^2) + |\log g/\varepsilon|^{1+\beta}g,$$

where $0 < \varepsilon < (2C)^{-1}$. In the same way as in Theorem 3.1 we obtain

$$\begin{aligned} & \frac{1}{2C} \int_{\mathbb{R}^d} |\nabla V(x) - x|^{2(1+\beta)} \mu(dx) \\ &= \frac{1}{C} \int_{\mathbb{R}^d} \left[\frac{x^2}{2} - V(x) - V^*(\nabla V(x)) + \frac{|\nabla V(x)|^2}{2} \right] |\nabla V(x) - x|^{2\beta} \mu(dx) \\ &= \frac{1}{C} \int_{\mathbb{R}^d} \left[\frac{x^2}{2} - V(x) \right] |\nabla V(x) - x|^{2\beta} \mu(dx) + \frac{1}{C} \int_{\mathbb{R}^d} \left[\frac{x^2}{2} - V^*(x) \right] |\nabla V^*(x) - x|^{2\beta} g(x) \mu(dx). \end{aligned}$$

Assuming that $\int_{\mathbb{R}^d} \left(V(x) - \frac{x^2}{2} \right) \mu(dx) = 0$ and using the inequality $xy \leq e^y + x \log x - x$ and the estimate

$$\int_{\mathbb{R}^d} \exp\left[\frac{1}{C}\left(\frac{x^2}{2} - V^*\right)\right] \mu(dx) \leq \exp\left[\frac{1}{C} \int_{\mathbb{R}^d} \left(V(x) - \frac{x^2}{2} \right) \mu(dx)\right],$$

we find

$$\begin{aligned} & \frac{1}{C} \int_{\mathbb{R}^d} \left[\frac{x^2}{2} - V^*(x) \right] |\nabla V^*(x) - x|^{2\beta} g(x) \mu(dx) \leq \int_{\mathbb{R}^d} \exp\left[\frac{1}{C}\left(\frac{x^2}{2} - V^*(x)\right)\right] \mu(dx) \\ &+ \int_{\mathbb{R}^d} g(x) |\nabla V^*(x) - x|^{2\beta} \log(g(x) |\nabla V^*(x) - x|^{2\beta}) \mu(dx) - \int_{\mathbb{R}^d} g(x) |\nabla V^*(x) - x|^{2\beta} \mu(dx) \\ &\leq 1 + \int_{\mathbb{R}^d} |\nabla V(x) - x|^{2\beta} \log(g(\nabla V(x)) |\nabla V(x) - x|^{2\beta}) \mu(dx) - \int_{\mathbb{R}^d} |\nabla V(x) - x|^{2\beta} \mu(dx) \\ &\leq 1 + \int_{\mathbb{R}^d} |\nabla V(x) - x|^{2\beta} \log |\nabla V(x) - x|^{2\beta} \mu(dx) \\ &\quad + \left(\int_{\mathbb{R}^d} |\nabla V(x) - x|^{2(\beta+1)} \mu(dx) \right)^{\frac{\beta}{1+\beta}} \left(\int_{\mathbb{R}^d} |\log g(\nabla V(x))|^{\beta+1} \mu(dx) \right)^{\frac{1}{1+\beta}}. \end{aligned}$$

Note that

$$\begin{aligned} & \frac{1}{2\beta} \int_{\mathbb{R}^d} |\nabla V(x) - x|^{2\beta} \log |\nabla V(x) - x|^{2\beta} \mu(dx) \\ & \leq \int_{\mathbb{R}^d} |\nabla V(x) - x|^{2\beta+1} \mu(dx) \leq \left(\int_{\mathbb{R}^d} |\nabla V(x) - x|^{2(\beta+1)} \mu(dx) \right)^{\frac{\beta+\frac{1}{2}}{1+\beta}}. \end{aligned}$$

In addition, by Hölder's inequality and (1.3) we have

$$\begin{aligned} & \int_{\mathbb{R}^d} \left[\frac{x^2}{2} - V(x) \right] |\nabla V(x) - x|^{2\beta} \mu(dx) \\ & \leq \left(\int_{\mathbb{R}^d} |\nabla V(x) - x|^{2(1+\beta)} \mu(dx) \right)^{\frac{\beta}{1+\beta}} \left(\int_{\mathbb{R}^d} \left| \frac{x^2}{2} - V(x) \right|^{1+\beta} \mu(dx) \right)^{\frac{1}{1+\beta}} \\ & \leq \left(\int_{\mathbb{R}^d} |\nabla V(x) - x|^{2(1+\beta)} \mu(dx) \right)^{\frac{\beta}{1+\beta}} \left(\int_{\mathbb{R}^d} \left| \frac{x^2}{2} - V(x) \right|^{2(1+\beta)} \mu(dx) \right)^{\frac{1}{2(1+\beta)}} \\ & \leq K(\beta, C) \left(\int_{\mathbb{R}^d} |\nabla V(x) - x|^{2(1+\beta)} \mu(dx) \right)^{\frac{\beta+\frac{1}{2}}{1+\beta}}. \end{aligned}$$

Letting $t = \int_{\mathbb{R}^d} |\nabla V(x) - x|^{2(1+\beta)} \mu(dx)$, we obtain

$$\frac{t}{2C} \leq 1 + t^{\frac{\beta}{1+\beta}} \left(\int_{\mathbb{R}^d} |\log g|^{1+\beta} g d\mu \right)^{\frac{1}{1+\beta}} + (2\beta + K(\beta, C)) t^{\frac{\beta+\frac{1}{2}}{1+\beta}}.$$

This relation implies our claim. \square

4. THE CASE OF MEASURES SATISFYING THE POINCARÉ INEQUALITY

In general $|\nabla\Phi|^2$ is not exponentially integrable for measures satisfying the classical Poincaré inequality. Indeed, consider the following measures on the real line: $\mu = \frac{1}{2}e^{-|x|} dx$ and $\nu = e^{-2|x|} dx$. In this case $T(x) = 2x$ and $T(x) - x = x$. Clearly, $\frac{d\nu}{d\mu} \in L^p(\mu)$ for every $p > 0$, but $\exp(\omega|T(x) - x|^2)$ is not μ -integrable if $\omega > 0$. However, it is known (see, e.g., [21], [31]) that if μ satisfies the classical Poincaré inequality, then $K(\mu, g \cdot \mu, h)$ with $h(x, y) = |x - y|^2$ can be estimated via $\|g\|_{L^2(\mu)}^2$. Here we prove L^p -estimates for optimal transportations in the case of measures satisfying the Poincaré inequality (1.2). The idea of proof is seen from the following simple reasoning in the case $p = 2$.

Keeping the notation from the previous section and choosing V in such a way that

$$\int_{\mathbb{R}^d} \left(V(x) - \frac{x^2}{2} \right) \mu(dx) = 0,$$

we obtain

$$\begin{aligned}
\frac{1}{2} \int_{\mathbb{R}^d} |\nabla V(x) - x|^2 \mu(dx) &= \int_{\mathbb{R}^d} \left[\frac{x^2}{2} - V(x) - V^*(\nabla V(x)) + \frac{|\nabla V(x)|^2}{2} \right] \mu(dx) \\
&= \int_{\mathbb{R}^d} \left[\frac{x^2}{2} - V^*(x) \right] g(x) \mu(dx) \\
&\leq \int_{\mathbb{R}^d} \left[V(x) - \frac{x^2}{2} \right] g(x) \mu(dx) = \int_{\mathbb{R}^d} \left[V(x) - \frac{x^2}{2} \right] (g(x) - 1) \mu(dx) \\
&\leq \left(\int_{\mathbb{R}^d} (g - 1)^2 d\mu \right)^{1/2} \left(\int_{\mathbb{R}^d} \left[V(x) - \frac{x^2}{2} \right]^2 \mu(dx) \right)^{1/2} \\
&= \left(\int_{\mathbb{R}^d} (g^2 - 1) d\mu \right)^{1/2} \left(C_2 \int_{\mathbb{R}^d} |\nabla V(x) - x|^2 \mu(dx) \right)^{1/2}.
\end{aligned}$$

Hence

$$\int_{\mathbb{R}^d} |\nabla V(x) - x|^2 \mu(dx) \leq 4C_2 \int_{\mathbb{R}^d} (g^2 - 1) d\mu.$$

Theorem 4.1. *Let μ satisfy (1.2). Then for every $\beta > 0$ one has*

$$\frac{1}{2} \left[\int_{\mathbb{R}^d} |\nabla V(x) - x|^{2(\beta+1)} \mu(dx) \right]^{\frac{1}{2(\beta+1)}} \leq C_{2(1+\beta)}^{\frac{1}{2(1+\beta)}} \left(\int_{\mathbb{R}^d} g^2 d\mu \right)^{\frac{1}{2(1+\beta)}} + C_{2(1+\beta)}^{\frac{1}{2(1+\beta)}}.$$

Proof. First we observe that the function $|T(x)|$ belongs to all $L^p(\mu)$. This is easily seen by the change of variables formula and the μ -integrability of the function $|x|^p g(x)$, which follows by the Cauchy inequality and the μ -integrability of the function $|x|^{2p}$ implied by the Poincaré inequality. In the same way as above we estimate

$$\begin{aligned}
\frac{1}{2} \int_{\mathbb{R}^d} |\nabla V(x) - x|^{2(\beta+1)} \mu(dx) \\
= \int_{\mathbb{R}^d} |\nabla V(x) - x|^{2\beta} \left[\frac{x^2}{2} - V(x) - V^*(\nabla V(x)) + \frac{|\nabla V(x)|^2}{2} \right] \mu(dx).
\end{aligned}$$

By using Hölder's inequality, the estimate $V(x) + V^*(x) \geq x^2$, and the change of variables formula, we obtain

$$\begin{aligned}
&\int_{\mathbb{R}^d} |\nabla V(x) - x|^{2\beta} \left[\frac{|\nabla V(x)|^2}{2} - V^*(\nabla V(x)) \right] \mu(dx) \\
&= \int_{\mathbb{R}^d} |\nabla V^*(x) - x|^{2\beta} \left[\frac{x^2}{2} - V^*(x) \right] g(x) \mu(dx) \leq \int_{\mathbb{R}^d} |\nabla V^*(x) - x|^{2\beta} \left[V(x) - \frac{x^2}{2} \right] g(x) \mu(dx) \\
&\leq \left(\int_{\mathbb{R}^d} |\nabla V^*(x) - x|^{2(\beta+1)} g(x) \mu(dx) \right)^{\frac{\beta}{1+\beta}} \left(\int_{\mathbb{R}^d} \left| V(x) - \frac{x^2}{2} \right|^{1+\beta} g(x) \mu(dx) \right)^{\frac{1}{1+\beta}} \\
&\leq \left(\int_{\mathbb{R}^d} |\nabla V(x) - x|^{2(\beta+1)} \mu(dx) \right)^{\frac{\beta}{1+\beta}} \left(\int_{\mathbb{R}^d} \left| V(x) - \frac{x^2}{2} \right|^{2(1+\beta)} \mu(dx) \right)^{\frac{1}{2(1+\beta)}} \left(\int_{\mathbb{R}^d} g^2 d\mu \right)^{\frac{1}{2(1+\beta)}} \\
&\leq C_{2(1+\beta)}^{\frac{1}{2(1+\beta)}} \left(\int_{\mathbb{R}^d} |\nabla V(x) - x|^{2(\beta+1)} \mu(dx) \right)^{\frac{\beta+\frac{1}{2}}{1+\beta}} \left(\int_{\mathbb{R}^d} g^2 d\mu \right)^{\frac{1}{2(1+\beta)}}.
\end{aligned}$$

In addition,

$$\begin{aligned}
& \int_{\mathbb{R}^d} |\nabla V(x) - x|^{2\beta} \left[\frac{x^2}{2} - V(x) \right] \mu(dx) \\
& \leq \left(\int_{\mathbb{R}^d} |\nabla V(x) - x|^{2(\beta+1)} \mu(dx) \right)^{\frac{\beta}{1+\beta}} \left(\int_{\mathbb{R}^d} \left| V(x) - \frac{x^2}{2} \right|^{1+\beta} \mu(dx) \right)^{\frac{1}{1+\beta}} \leq \\
& \leq C_{2(1+\beta)}^{\frac{1}{2(1+\beta)}} \left(\int_{\mathbb{R}^d} |\nabla V(x) - x|^{2(\beta+1)} \mu(dx) \right)^{\frac{\beta}{1+\beta}} \left(\int_{\mathbb{R}^d} |\nabla V(x) - x|^{2(1+\beta)} \mu(dx) \right)^{\frac{1}{2} \frac{1}{1+\beta}} \\
& \leq C_{2(1+\beta)}^{\frac{1}{2(1+\beta)}} \left(\int_{\mathbb{R}^d} |\nabla V(x) - x|^{2(\beta+1)} \mu(dx) \right)^{\frac{\beta+\frac{1}{2}}{1+\beta}}.
\end{aligned}$$

Our claim follows from these estimates. \square

This work was completed during our visit to the University of Bielefeld. We thank M. Röckner for useful discussions. Our research was supported by the RFBR project 04-01-00748, the Scientific Schools Grant 1758.2003.1, the DFG Grant 436 RUS 113/343/0(R), the INTAS project 03-51-5018, and the DFG–Forschergruppe “Spectral Analysis, Asymptotic Distributions, and Stochastic Dynamics”. The second author gratefully acknowledges the support of the Centro di Ricerca Matematica Ennio De Giorgi, Scuola Normale Superiore di Pisa.

REFERENCES

- [1] *Agueh M., Ghoussoub N., Kang X.* Geometric inequalities via a general comparison principle for interacting gases. — *Geom. Funct. Anal.*, 2004, v. 14, p. 215–244.
- [2] *Bobkov S.G.* Isoperimetric and analytic inequalities for log-concave probability measures. — *Ann. Probab.*, 1999, v. 27, n 4, p. 1902–1921.
- [3] *Bobkov S.G., Gentil I., Ledoux M.* Hypercontractivity of Hamilton–Jacobi equations. — *J. Math. Pures Appl.*, 2001, v. 80, n 7, p. 669–696.
- [4] *Bobkov S.G., Götze F.* Exponential integrability and transportation cost related to logarithmic Sobolev inequality. — *J. Funct. Anal.*, 1999, v. 163, p. 1–28.
- [5] *Bogachev V.I.* Gaussian measures. Amer. Math. Soc., Rhode Island, 1998.
- [6] *Bogachev V.I.* Foundations of measure theory. V. 1,2. Regular and Chaotic Dynamics, Moscow (in Russian), 2003.
- [7] *Bogachev V.I., Kolesnikov A.V.* Open mappings of probability measures and the Skorohod representation theorem. — *Teor. Veroyatn. i ee Primen.*, 2001, v. 46, n 1, p. 3–27 (in Russian); English transl.: *Theory Probab. Appl.*, 2001, v. 46, n 1, p. 20–38.
- [8] *Bogachev V.I., Kolesnikov A.V.* Nonlinear transformations of convex measures and entropy of the Radon–Nikodym densities. — *Dokl. Russian Acad. Sci.*, 2004, v. 397, n 2, p. 155–159 (in Russian); English transl.: *Dokl. Math.*, 2004, v. 70, n 1, p. 524–528.
- [9] *Bogachev V.I., Kolesnikov A.V.* Nonlinear transformations of convex measures. — *Teor. Veroyatn. i ee Primen.*, 2005, v. 50, n 1, p. 27–51 (in Russian); English transl.: *Theory Probab. Appl.*, 2005, v. 50, n 1.
- [10] *Bogachev V.I., Kolesnikov A.V.* On the Monge–Ampère equation in infinite dimensions. — *BiBoS Preprint N 05–01–175*, Universität Bielefeld, 2005.
- [11] *Bogachev V.I., Kolesnikov A.V., Medvedev K.V.* On triangular transformations of measures. — *Dokl. Russian Acad. Sci.*, 2004, v. 396, n 6, p. 727–732 (in Russian); English transl.: *Dokl. Math.*, 2004, v. 69, n 3, p. 438–442.
- [12] *Bogachev V.I., Kolesnikov A.V., Medvedev K.V.* Triangular transformations of measures. — *Matem. Sbornik*, 2005, v. 196, n 3, p. 3–30 (in Russian); English transl.: *Sbornik Math.*, 2005, v. 196, n 3.
- [13] *Brenier Y.* Polar factorization and monotone rearrangement of vector valued functions. — *Comm. Pure Appl. Math.*, 1991, v. 44, p. 375–417.

- [14] *Cordero-Erausquin D.* Non-smooth differential properties of optimal transport. In: Recent advances in the theory and applications of mass transport, Contemporary Mathematics, v. 353, p. 61–71, American Math. Soc., Providence, Rhode Island, 2004.
- [15] *Cordero-Erausquin D., Nazaret B., Villani C.* A mass-transportation approach to sharp Sobolev and Gagliardo–Nirenberg inequalities. — Adv. Math., 2004, v. 182, n 2, p. 307–332.
- [16] *Fernique X.* Extension du théorème de Cameron–Martin aux translations aléatoires. II. Intégrabilité des densités. High dimensional probability, III (Sandjberg, 2002), p. 95–102, Progr. Probab., v. 55, Birkhäuser, Basel, 2003.
- [17] *Feyel D., Üstünel A.S.* Monge–Kantorovitch measure transportation and Monge–Ampère equation on Wiener space. — Probab. Theor. Relat. Fields, 2004, v. 128, n 3, p. 347–385.
- [18] *Gangbo W., McCann R.J.* The geometry of optimal transportation. — Acta Math., 1996, v. 177, p. 113–161.
- [19] *Hajlasz P.* Change of variables formula under minimal assumptions. — Colloq. Math., 1993, v. 64, n 1, p. 93–101.
- [20] *Kolesnikov A.V.* Convexity inequalities and nonlinear transformations of measures. — Russian Math. Dokl., 2004, v. 396, n 3, p. 300–304 (in Russian); English transl.: Dokl. Math., 2004, v. 69, n 3, p. 368–372.
- [21] *Kolesnikov A.V.* Convexity inequalities and optimal transport of infinite-dimensional measures. — J. Math. Pures Appl., 2004, v. 83, n 11, p. 1373–1404.
- [22] *Ledoux M.* The concentration of measure phenomenon. Amer. Math. Soc., Providence, Rhode Island, 2001.
- [23] *McCann R.J.* Existence and uniqueness of monotone measure-preserving maps. — Duke Math. J., 1995, v. 80, p. 309–323.
- [24] *Otto F., Villani C.* Generalization of an inequality by Talagrand, and links with the logarithmic Sobolev inequality. — J. Funct. Anal., 2000, v. 173, p. 361–400.
- [25] *Pisier G.* Probabilistic methods in the geometry of Banach spaces. Lecture Notes in Math., v. 1206, p. 167–241, Springer-Verlag, Berlin, 1986.
- [26] *Rachev S.T., Rüschendorf L.* Mass transportation problems. V. I, II, Springer, New York, 1998.
- [27] *Rockafellar T.* Convex analysis. Princeton University Press, 1970.
- [28] *Talagrand M.* Transportation cost for Gaussian and other product measures. — Geom. Funct. Anal., 1996, v. 6, p. 587–600.
- [29] *Villani C.* Topics in optimal transportation. Amer. Math. Soc., Providence, Rhode Island, 2003.
- [30] *Wang F.Y.* Logarithmic Sobolev inequalities on non-compact Riemannian manifolds. — Probab. Theory Relat. Fields, 1997, v. 109, p. 417–424.
- [31] *Wang F.Y.* Probability distance inequalities on Riemannian manifolds and path spaces. — J. Funct. Anal., 2004, v. 206, p. 167–190.

DEPARTMENT OF MECHANICS AND MATHEMATICS, MOSCOW STATE UNIVERSITY, 119992 MOSCOW, RUSSIA

SCUOLA NORMALE SUPERIORE, CENTRO DI RICERCA MATEMATICA ENNIO DE GIORGI, I-56100 PISA, ITALY, E-MAIL: SASCHA77@MAIL.RU