

# Relations between Sobolev and Kantorovich norms on manifolds with curvature conditions

V. I. Bogachev, A.V. Shaposhnikov, F.-Yu. Wang

## Abstract

We prove several multiplicative inequalities relating the Kantorovich norm with the Sobolev norm for functions on a Riemannian manifold satisfying certain curvature conditions.

Keywords: Kantorovich norm, Sobolev norm, Riemannian manifold, Ricci tensor, Bakry–Emery tensor

AMS Subject Classification: 58C35, 47J20

## 1. INTRODUCTION AND NOTATION

In the last decade, there has been increasing interest in functional inequalities relating Sobolev norms with certain other norms or similar quantities such as entropy and transportation cost (see [18], [20]). In papers [12], [6], [7], some multidimensional analogs were proved of the classic Hardy–Landau–Littlewood inequality

$$\|f'\|_{L^1}^2 \leq C \|f\|_{L^1} \|f''\|_{L^1}$$

for functions with two derivatives on the real line, having the form

$$\|f\|_{L^1} \leq C \|\nabla f\|_{L^1} \|f\|_K$$

for integrable functions with zero integral from the usual Sobolev class  $W^{1,1}(\mathbb{R}^d)$ , where  $\|f\|_K$  is the Kantorovich norm (all definitions are given below). The indicated Hardy–Landau–Littlewood inequality in the one-dimensional case is equivalent to this inequality applied to  $f'$  in place of  $f$  (then, letting  $g = f'$ , for  $\|f\|_{L^1}$  we obtain  $\|g\|_K$ ). This estimate was extended in [6] to the case of a Riemannian manifold  $M$  equipped with a probability Borel measure  $\mu$  in the following form:

$$\|f - g\|_{L^1(\mu)}^2 \leq C(\mu) \|f\mu - g\mu\|_K \|\nabla f - \nabla g\|_{L^1(\mu)}, \quad (1.1)$$

where  $f$  and  $g$  are probability densities with respect to  $\mu$ ,  $\|\nabla f - \nabla g\|_{L^1(\mu)}$  is the integral of  $|\nabla f - \nabla g|$  with respect to the measure  $\mu$ , and, for any function  $f$  integrable with respect to the measure  $\mu$ , we denote by  $f\mu$  the measure given by density  $f$  with respect to  $\mu$ . The assumptions about  $M$  and  $\mu$  employed in [6] are expressed in terms of inequalities for the heat semigroup on  $M$ . In the paper [7], inequality (1.1) was strengthened (in particular, it was proved for certain infinite-dimensional manifolds), moreover, certain geometric conditions were found for its validity in terms of the Ricci tensor. In particular, it was shown that the number  $C(\mu)$  can be taken in such a way that its dependence on  $\mu$  will have the following character. If  $\mu$  has a density  $e^{-V}$  with respect to the Riemannian volume, where  $V \in C^2$ , and for some number  $k \geq 0$  we have the inequality  $\text{Ric}_V \geq k \cdot \text{I}$ , in which

$$\text{Ric}_V(U, U) = \text{Ric}(U, U) + \text{Hess}_V(U, U), \quad U \in TM,$$

where  $\text{Ric}$  is the Ricci curvature on  $M$  (see [8, p. 60]),  $TM$  is the tangent bundle,  $\text{Hess}_V$  is the Hessian of  $V$ , i.e., the operator corresponding to the second derivative (see [8, p. 41]), then  $C(\mu)$  can be taken depending only on  $k$ .

In this paper we prove a more general inequality with the Kantorovich distance of order 1 and also strengthen a number of inequalities from the recent paper [13] with the use of the Kantorovich distance of order 2.

We recall (see, e.g., [5]) that the Kantorovich distance between two probability Borel measures on a metric space  $M$  with a metric  $\varrho$  is given by

$$\|\mu - \nu\|_K := \sup \left\{ \int_M \varphi d(\mu - \nu), \varphi \in \text{Lip}_1 \right\},$$

where sup is taken over the class  $\text{Lip}_1$  of functions on  $M$  that are Lipschitzian with constant 1. The same formula defines the Kantorovich norm of a bounded signed measure with zero value on the whole space.

The introduced distance is also called the Kantorovich distance of order 1 (exactly this distance was introduced by Kantorovich himself [10], [11]), because for  $p = 1$  it coincides with the value of the Kantorovich distance  $W_p$  of order  $p$  (sometimes erroneously called the Wasserstein distance) defined by the formula

$$W_p(\mu, \nu)^p = \inf_{\sigma \in \Pi(\mu, \nu)} \int_{M \times M} \varrho(x, y)^p \sigma(dx dy),$$

where inf is taken over all measures  $\sigma$  from the set  $\Pi(\mu, \nu)$  of Borel probability measures on  $M \times M$  having projections  $\mu$  and  $\nu$  onto the first and second factors.

Let us also recall that the Bakry–Emery tensor of a Riemannian manifold  $M$  of dimension  $d$  equipped with a measure  $\mu = e^{-V} dx$ , where  $dx$  is the Riemannian volume and  $V$  is a function of class  $C^2$ , is defined by

$$\text{Ric}_{N, \mu} = \text{Ric} + \text{Hess}V - \frac{\nabla V \otimes \nabla V}{N - d}, \quad N \geq d$$

where for  $N = d$  we set  $\frac{\nabla V \otimes \nabla V}{N - d} = \infty \cdot \langle \cdot, \cdot \rangle$  if  $\nabla V \neq 0$  and  $= 0 \cdot \langle \cdot, \cdot \rangle$  if  $\nabla V = 0$ . We say that the pair  $(M, \mu)$  satisfies the curvature condition  $CD(K, N)$ , where  $K \in \mathbb{R}$  and  $N \in [d, +\infty]$ , if

$$\text{Ric}_{N, V} \geq K \cdot I.$$

Throughout the symbol  $\|f\|_p$  denotes the  $L^p$ -norm of a function  $f$  with respect to the measure  $\mu$  involved in the formulations of the theorems, i.e., the norm in the space  $L^p(\mu)$ . The integral of  $f$  with respect to the measure  $\mu$  will be denoted by the symbol  $\mu(f)$ .

## 2. MAIN RESULTS

Here is the first result of this paper.

**Theorem 2.1.** *Suppose that  $M$  is a Riemannian manifold with a probability measure  $\mu(dx) := e^{-V(x)} dx$  satisfying condition  $CD(0, N)$  for some  $N \in [d, \infty]$ .*

- (1) *If  $N < \infty$ , then for any  $p, q \geq 1$  there exists a constant  $C > 0$  depending only on  $p, q, N$  such that*

$$\|f - 1\|_r^\theta \leq C(\|\nabla f\|_q W_p(f\mu, \mu) + \{\|\nabla f\|_1 W_1(f\mu, \mu)\}^{\frac{\theta}{2r}})$$

*holds for any  $f \geq 0$  with  $\mu(f) = 1$ , where  $\theta = 1 + \frac{1}{p} + \frac{1}{N}$ ,  $r = \frac{pq\theta}{p+q}$ . In particular, with  $p = q = 1$  this becomes*

$$\|f - 1\|_{1 + \frac{1}{2N}}^{2 + \frac{1}{N}} \leq C\|\nabla f\|_1 W_1(f\mu, \mu), \quad f \geq 0, \quad \mu(f) = 1.$$

- (2) *If  $N = \infty$ , then for any  $q \in (1, 2]$  there exists a constant  $C > 0$  such that*

$$\|f - 1\|_{\frac{3q}{3q+(q+2)}}^{\frac{3}{2}} \leq C(\|\nabla f\|_q W_2(f\mu, \mu) + \{\|\nabla f\|_1 W_1(f\mu, \mu)\}^{\frac{q+2}{4q}})$$

*holds for any  $f \geq 0$  with  $\mu(f) = 1$ .*

*Proof.* By [13, Theorem 1], there exists a constant  $C_1 > 1$  such that

$$\|(f - C_1)^+\|_r^\theta \leq C_1 \|\nabla f\|_q W_p(f\mu, \mu), \quad f \geq 0, \mu(f) = 1. \quad (2.1)$$

Moreover, by [7, Theorem 1], we have

$$\|f - 1\|_1 \leq \sqrt{2\|\nabla f\|_1 W_1(f\mu, \mu)}, \quad f \geq 0, \mu(f) = 1. \quad (2.2)$$

Since  $(f - 1)^+ = (f - C_1)^+ + (f - 1)^+ \wedge (C_1 - 1)$ , we have

$$\{(f - 1)^+\}^r \leq 2^{r-1}\{(f - C_1)^+\}^r + 2^{r-1}(C_1 - 1)^{r-1}(f - 1)^+.$$

Therefore,

$$\begin{aligned} |f - 1|^r &= \{(f - 1)^+\}^r + \{(1 - f)^+\}^r \leq \{(f - 1)^+\}^r + (1 - f)^+ \\ &\leq 2^{r-1}\{(f - C_1)^+\}^r + (2C_1)^{r-1}|f - 1|. \end{aligned}$$

Combining this with (2.1) and (2.2) we prove the first assertion.

The second assertion can be proved in the same way by using [13, Theorem 5] to replace (2.1), which says that if  $C(0, \infty)$  holds, then for any  $q \in (1, 2]$  there exists a constant  $C_1 > 0$  such that

$$\|(f - C_1)^+\|_{3q/(q+2)}^{3/2} \leq C_1 \|\nabla f\|_q W_2(f\mu, \mu), \quad f \geq 0, \mu(f) = 1,$$

so our claim follows.  $\square$

The next theorem strengthens [13, Proposition 4.2] (see also [9]).

**Theorem 2.2.** *Let  $(M, \mu)$ , where  $\mu(dx) := e^{-V(x)} dx$  is a probability measure, satisfy the curvature condition  $CD(0, \infty)$ . Then for every  $q \in [1, 2]$  there is a number  $C > 0$  depending only on  $q$  such that for every probability measure  $\nu = f\mu$  with a smooth density  $f$  with respect to  $\mu$  one has*

$$\sup_{u \geq C} [u^{3/2}(\ln u)^{1/2}] \mu(f \geq u)^{3/2r} \leq C \|\nabla f\|_q W_2(f\mu, \mu), \quad r = \frac{3q}{q+2}.$$

*Proof.* As in the proof from [13], for every  $u > 0$  and  $t > 0$  we have

$$\mu(f \geq 2u) \leq \frac{(2t)^{q/2}}{u^q} \|\nabla f\|_q^q + \mu(P_t f \geq u).$$

The next step is to apply the well-known entropy inequality

$$\int_M h P_t f d\mu \leq \int_M P_t f \log P_t f d\mu + \log \int_M e^h d\mu.$$

Now let us choose the following  $h$ :

$$h := I_F \log P_t f,$$

where

$$F = \{P_t f \geq u\}.$$

Then

$$\int_M I_F P_t f \log P_t f d\mu \leq \int_M P_t f \log P_t f d\mu + \log \int_M [1 + I_F(P_t f - 1)] d\mu$$

and also

$$\int_F P_t f \log P_t f d\mu \leq \int_M P_t f \log P_t f d\mu + \log \left[ \int_M 1 d\mu + \int_F (P_t f - 1) d\mu \right].$$

Since for every  $x \geq 0$  one has

$$\log(1 + x) \leq x$$

and

$$\int_M 1 d\mu = 1,$$

we obtain

$$\begin{aligned} \int_F P_t f \log P_t f d\mu &\leq \int_M P_t f \log P_t f d\mu + \int_F (P_t f - 1) d\mu, \\ \int_F 1 + P_t f [\log P_t f - 1] d\mu &\leq \int_M P_t f \log P_t f d\mu. \end{aligned}$$

Hence, assuming that  $u \geq 8$ , we have

$$\begin{aligned} \mu(F)u \log u &\leq 2 \int_M P_t f \log P_t f d\mu \leq \frac{1}{2t} W_2^2(f\mu, \mu), \\ \mu(P_t f \geq u) &\leq \frac{1}{2tu \log u} W_2^2(f\mu, \mu). \end{aligned}$$

Finally, we have the following estimate:

$$\mu(f \geq 2u) \leq \frac{(2t)^{q/2}}{u^q} \|\nabla f\|_q^q + \frac{1}{2tu \log u} W_2^2(f\mu, \mu).$$

Optimizing in  $t > 0$ , we obtain the desired inequality.  $\square$

**Theorem 2.3.** *If  $(M, \mu)$  satisfies the hypothesis of Theorem 2.2, then for every  $q \in (1, 2]$  there is a number  $C > 0$  depending only on  $q$  such that for every probability measure  $\nu = f\mu$  with a smooth density  $f$  with respect to  $\mu$  one has*

$$\int (f - C)_+^r (\ln[1 + (f - C)_+])^\alpha d\mu \leq C \|\nabla f\|_q^s W_2^s(\nu, \mu), \quad (2.3)$$

where

$$r = \frac{3q}{q+2}, \quad \alpha = \frac{q}{q+2}, \quad s = \frac{2q}{q+2}.$$

The proof will be given below.

The next lemma is a reinforcement of a remark made in [13] in relation to formula (9) there.

**Lemma 2.4.** *Let  $a > 1, \alpha > 0$ . Then there is a number  $C$  depending only on  $a$  and  $\alpha$  such that, for every collection  $\{F\}_{k=1}^\infty$  of subsets of  $M$ , every interval of nonnegative integer numbers  $I = [k_0, k_1] = \{k_0, k_0 + 1, \dots, k_1\}$  and every  $x \in M$  one has*

$$\sum_{k \in I} (1+k)^\alpha a^k I_{F_k}(x) \leq C \sup_{k \in I} (1+k)^\alpha a^k I_{F_k}(x).$$

*Proof.* For each  $k \geq 0$  we have the following chain of inequalities:

$$\begin{aligned} \sum_{i=0}^k (1+i)^\alpha a^i &\leq \int_0^{k+1} (1+t)^\alpha a^t dt = \int_0^{k+1} (1+t)^\alpha \frac{1}{\log a} \frac{d}{dt} a^t dt \\ &= \frac{1}{\log a} (k+2)^\alpha a^{k+1} - \frac{1}{\log a} - \int_0^{k+1} \alpha (1+t)^{\alpha-1} \frac{1}{\log a} a^t dt. \leq C(a, \alpha) (1+k)^\alpha a^k. \end{aligned}$$

Now it is easy to complete the proof.  $\square$

The next lemma is also based on the reasoning from [13].

**Lemma 2.5.** *Let  $a > 1, \alpha > 0, p > 1$ . Then there is a number  $C = C(a, \alpha)$  such that for every collection  $\{F\}_{k=1}^\infty$  of Borel sets in  $M$ , every interval of nonnegative integer numbers  $I = [k_0, k_1] = \{k_0, k_0 + 1, \dots, k_1\}$  and every  $\varepsilon > 0$  one has*

$$\int \varphi f d\mu \leq \frac{1}{\varepsilon} W_p^p(f\mu, \mu) + C \int \psi_\varepsilon d\mu,$$

where

$$\begin{aligned} \varphi(x) &= \sum_{k \in I} (1+k)^\alpha a^k I_{F_k}(x), \\ \psi_\varepsilon(x) &= \sum_{k \in I} \sup_{y \in M_{k,\varepsilon}(x)} (1+k)^\alpha a^k I_{F_k}(y), \\ M_{k,\varepsilon}(x) &= \left\{ y \in M : d(x, y)^p \leq C\varepsilon(1+k)^\alpha a^k \right\}. \end{aligned}$$

*Proof.* By the Kantorovich duality for every  $\varepsilon > 0$  we have

$$\int \varphi f d\mu \leq \frac{1}{\varepsilon} W_p^p(f\mu, \mu) + \int Q_\varepsilon \varphi d\mu.$$

where

$$Q_\varepsilon \varphi(x) = \sup_{y \in M} \left[ \varphi(y) - \frac{1}{\varepsilon} d(x, y)^p \right].$$

Applying Lemma 2.4, we obtain the estimate

$$\begin{aligned} Q_\varepsilon \varphi(x) &= \sup_{y \in M} \left[ \sum_{k \in I} (1+k)^\alpha a^k I_{F_k}(y) - \frac{1}{\varepsilon} d(x, y)^p \right] \\ &\leq \sup_{y \in M} \left[ C \sup_{k \in I} (1+k)^\alpha a^k I_{F_k}(y) - \frac{1}{\varepsilon} d(x, y)^p \right] \\ &= C \sup_{y \in M} \sup_{k \in I} \left[ (1+k)^\alpha a^k I_{F_k}(y) - \frac{1}{C\varepsilon} d(x, y)^p \right] \\ &= C \sup_{k \in I} \sup_{y \in M} \left[ (1+k)^\alpha a^k I_{F_k}(y) - \frac{1}{C\varepsilon} d(x, y)^p \right] \\ &= C \sup_{k \in I} \sup_{y \in M_{k,\varepsilon}(x)} (1+k)^\alpha a^k I_{F_k}(y) \leq C \sum_{k \in I} \sup_{y \in M_{k,\varepsilon}(x)} (1+k)^\alpha a^k I_{F_k}(y), \end{aligned}$$

which completes the proof.  $\square$

Now we are ready to prove Theorem 2.3. We will use Ledoux's strategy with some modifications. For the reader's convenience all the details are presented. The main difference with the proof by Ledoux of his result is our choice of the "multipliers" and application of Harnack's inequality from [16], where the exponential term is no longer constant.

*Proof of Theorem 2.3.* Let  $f_k = \min((f - 2^k)_+, 2^k)$ ,  $u = 2^{k-1}$ , where  $k \in \mathbb{Z}_+$ , and let  $t_k > 0$ , where the value of  $t_k$  will be chosen later. We will use the following inequality proven in [13] (see the proof of Theorem 1, p. 7):

$$\mu(f_k \geq 2^k) \leq \frac{C(q)t_k^{q/2}}{2^{qk}} \int_{A_k} |\nabla f|^q d\mu + \frac{1}{2^k} \int I_{F_k} f d\mu,$$

$$A_k = \{2^k \leq f < 2^{k+1}\}, \quad F_k = \{P_{t_k} f_k \geq 2^{k-1}\}.$$

Let  $r = \frac{3q}{q+2}$ . Let  $\eta = \frac{C}{\log 2}$ , where  $C$  is the constant from Lemma 2.5. We now choose  $t_k$  such that

$$(1+k)^\alpha 2^{(r-q)k} t_k^{q/2} = \eta^{q/2} \varepsilon^{q/2} \Leftrightarrow t_k = \eta(1+k)^{-2\alpha/q} 2^{2(1-r/q)k} \varepsilon.$$

Let  $I$  be a fixed interval of integers  $[k_0, k_1]$ . Note that  $\mu(f \geq 2^{k+1}) = \mu(f_k \geq 2^k)$ . Then

$$\begin{aligned} S_I &:= \sum_{k \in I} (1+k)^{\alpha} 2^{rk} \mu(f \geq 2^{k+1}) \\ &\leq C(q) \sum_{k \in I} (1+k)^{\alpha} 2^{(r-q)k} t_k^{q/2} \int_{A_k} |\nabla f|^q d\mu + \int \varphi f d\mu, \end{aligned}$$

where

$$\varphi = \sum_{k \in I} (1+k)^{\alpha} 2^{(r-1)k} I_{F_k}.$$

Taking into account the values of  $\{t_k\}$  we have

$$S_I \leq C(q, \eta) \varepsilon^{q/2} \int |\nabla f|^q d\mu + \int \varphi f d\mu.$$

Applying Lemma 2.5, we obtain

$$\int \varphi f d\mu \leq \frac{1}{\varepsilon} W_2^2(f\mu, \mu) + C \int \psi_\varepsilon d\mu,$$

where

$$\psi_\varepsilon(x) = \sum_{k \in I} \sup_{y \in M_{k,\varepsilon}(x)} (1+k)^{\alpha} 2^{(r-1)k} I_{F_k}(y),$$

$$M_{k,\varepsilon}(x) = \left\{ y \in M : d(x, y)^2 \leq C\varepsilon(1+k)^{\alpha} 2^{(r-1)k} \right\}.$$

The  $CD(0, \infty)$ -condition implies Harnack's inequality obtained in [16], according to which, under  $CD(0, \infty)$ , for any nonnegative measurable function  $f$  on  $M$  and all  $t > 0$ ,  $x, y \in M$  one has

$$P_t f(y)^2 \leq P_t(f^2)(x) e^{d(x,y)^2/2t}.$$

Since

$$\begin{aligned} t_k &= \eta(1+k)^{-2\alpha/q} 2^{2(1-r/q)k} \varepsilon, \\ (r-1) - 2(1-r/q) &= r + 2r/q - 3 = \frac{3q+6}{q+2} - 3 = 0, \\ \alpha(1+2/q) &= 1, \\ \eta &= \frac{C}{\log 2}, \end{aligned}$$

for every  $y \in M_k(x, \varepsilon)$  we have

$$\begin{aligned} I_{F_k}(y) &\leq 2^{-2k+2} P_{t_k}(f_k)(y)^2 \\ &\leq 2^{-2k+2} P_{t_k} f_k^2(x) \exp \left[ \frac{C}{2\eta} (1+k)^{\alpha(1+2/q)} \right] \leq 2^{-2k+2+(1+k)/2} P_{t_k} f_k^2(x). \end{aligned}$$

Hence

$$\begin{aligned} \int f_k^2 d\mu &\leq 2^{2k} \mu(f \geq 2^k), \\ \int \psi_\varepsilon d\mu &\leq \sum_{k \in I} (1+k)^{\alpha} 2^{(r-1)k} 2^{-2k+2+(1+k)/2+2k} \mu(f \geq 2^k), \end{aligned}$$

and for sufficiently large  $k_0$  (the lower bound of the interval  $I$ ) one has

$$C \int \psi_\varepsilon d\mu \leq \frac{1}{2} \sum_{k \in I} 2^{rk - \frac{k}{4}} \mu(f \geq 2^k).$$

Then

$$\begin{aligned} S_I &= \sum_{k \in I} (1+k)^\alpha 2^{rk} \mu(f \geq 2^{k+1}) \\ &\leq C \varepsilon^{q/2} \int |\nabla f|^q d\mu + \frac{1}{\varepsilon} W_2^2(f\mu, \mu) + \frac{1}{2} \sum_{k \in I} 2^{rk - \frac{k}{4}} \mu(f \geq 2^k), \\ S_I &\leq C \varepsilon^{q/2} \int |\nabla f|^q d\mu + \frac{1}{\varepsilon} W_2^2(f\mu, \mu) + \frac{1}{2} S_I + \frac{1}{2} 2^{rk_0 - \frac{k_0}{4}} \mu(f \geq 2^{k_0}). \end{aligned}$$

The last term can be estimated using the following weak-type inequality from [13] (see Proposition 6):

$$\mu(f \geq 2^{k_0}) \leq C(k_0) \|\nabla f\|_q^v W_2^v(f d\mu, \mu).$$

Optimizing in  $\varepsilon$ , choosing sufficiently large  $k_0$  and passing to the limit we easily obtain the bound

$$\sum_{k=k_0+1}^{\infty} (1+k)^\alpha 2^{rk} \mu(f \geq 2^k) \leq C \|\nabla f\|_q^v W_2^v(f d\mu, \mu),$$

where the constant  $C$  depends on  $q, \alpha$  and on our choice of  $k_0$ , but does not depend on  $f$ . Thus,

$$\begin{aligned} \int (f-C)_+^r \log^\alpha[1+(f-C)_+] d\mu &= \int_C^\infty \frac{d}{dt} \left( (t-C)^r \log^\alpha[1+(t-C)] \right) \mu(f > t) dt \\ &\leq C' \sum_{k=k_0+1}^{\infty} (1+k)^\alpha 2^{kr} \mu(f \geq 2^k) \leq C'' \|\nabla f\|_q^v W_2^v(f d\mu, \mu), \end{aligned}$$

which completes the proof.  $\square$

**Remark 2.6.** (i) The assumption  $q > 1$  has been used to ensure that

$$r > 1 \Leftrightarrow a = 2^{r-1} > 1.$$

It enables us to apply Lemma 2.4.

(ii) The assumption  $q \leq 2$  has been used implicitly in the inequality

$$\mu(f_k \geq 2^k) \leq \frac{C(q)t_k^{q/2}}{2^{qk}} \int_{A_k} |\nabla f|^q d\mu + \frac{1}{2^k} \int I_{F_k} f d\mu,$$

$$A_k = \{2^k \leq f < 2^{k+1}\}, \quad F_k = \{P_{t_k} f_k \geq 2^{k-1}\}.$$

For the proof, see [13, Proposition 3 and Theorem 1].

**Theorem 2.7.** *Let  $(M, g, \mu)$  be a weighted Riemannian manifold equipped with a probability measure  $\mu$  satisfying the curvature condition  $CD(0, \infty)$ . Then there is a constant  $C > 0$  such that for every probability measure  $d\nu = f d\mu$  with a smooth density  $f$  with respect to  $\mu$  one has*

$$\|f - 1\|_1^2 \leq C \|\nabla f\|_1 W_1(\nu, \mu).$$

*Proof.* For each smooth function  $g \geq 0$  with  $\|g\|_\infty \leq 1$  we have

$$\begin{aligned} \left| \int (f-1)g d\mu \right| &= \left| \int (f-1)P_t g d\mu - \int (f-1) \int_0^t \frac{d}{ds} P_s g d\mu \right| \\ &\leq W_1(\nu, \mu) \|\nabla P_t g\|_\infty + \int_0^t \int |\nabla f| |\nabla P_s g| d\mu dt \\ &\leq W_1(\nu, \mu) \|\nabla P_t g\|_\infty + \|\nabla f\|_1 \int_0^t \|\nabla P_s g\|_\infty dt. \end{aligned}$$

Applying the reverse isoperimetric-type inequality from [1] (for the proof, see also [3, Proposition 2.1])

$$[I(P_t g)]^2 - [P_t(I(g))]^2 \geq 2t|\nabla P_t g|^2$$

one readily obtains the estimate

$$\left| \int (f-1)g \, d\mu \right| \leq C \left[ W_1(\nu, \mu) \frac{1}{\sqrt{t}} + \|\nabla f\|_1 \int_0^t \frac{1}{\sqrt{s}} \, ds \right].$$

Now it is easy to complete the proof.  $\square$

### 3. EXTENSIONS TO THE NEGATIVE CURVATURE CASE

Throughout this section we assume the following:

(H) the bound  $\text{Ric}_V \geq -K$  holds for some constant  $K > 0$ , and the probability measure  $\mu(dx) := e^{-V(x)} dx$  satisfies the log-Sobolev inequality

$$\mu(f^2 \log f^2) \leq \frac{2}{\lambda} \mu(|\nabla f|^2), \quad f \in C_0^\infty(M), \mu(f^2) = 1 \quad (3.1)$$

for some constant  $\lambda > 0$ .

According to [17], under the curvature condition  $\text{Ric} - \text{Hess}_V \geq -K$ , the log-Sobolev inequality holds provided  $\mu(e^{\varepsilon \rho_o^2}) < \infty$  for some constant  $\varepsilon > \frac{K}{2}$ , where  $\rho_o$  is the Riemannian distance to a fixed point in  $M$ .

**Theorem 3.1.** *Assume (H). Then, for any  $q \in (1, 2]$ , there exists a constant  $C \geq 1$  depending only on  $K, \lambda$  and  $q$  such that for any probability density  $f$  with respect to  $\mu$  one has*

$$\sup_{u \geq C} \{u^{\frac{3}{2}} \ln^{\frac{1}{2}} u\} \mu(f \geq u)^{\frac{q+2}{2q}} \leq C \|\nabla f\|_q W_2(f\mu, \mu), \quad (3.2)$$

$$\int_M (f-C)_+^r (\ln[1+(f-C)_+])^\alpha \, d\mu \leq C \|\nabla f\|_q^s W_2^s(\nu, \mu). \quad (3.3)$$

*Proof.* According to [19], the curvature condition is equivalent to the log-Harnack inequality

$$P_t \log g(x) \leq \log P_t g(y) + \frac{K \rho(x, y)^2}{2(1 - e^{-2Kt})}, \quad t > 0, x, y \in M$$

for any positive measurable function  $g$ . This implies (see [15, Corollary 1.2]) that

$$\mu((P_t f) \log P_t f) \leq \frac{K W_2(f\mu, \mu)^2}{2(1 - e^{-2Kt})}, \quad t > 0, f \geq 0, \mu(f) = 1. \quad (3.4)$$

Thus, there exists a constant  $C_1$  depending only on  $K$  such that

$$\mu((P_t f) \log P_t f) \leq \frac{C_1}{t} W_2(f\mu, \mu)^2, \quad t \in (0, 1], f \geq 0, \mu(f) = 1. \quad (3.5)$$

Next, it is well known that the log-Sobolev inequality (3.1) implies that

$$\mu((P_t f) \log P_t f) \leq e^{-\lambda t} \mu(f \log f), \quad f \geq 0, \mu(f) = 1.$$

Combining this with (3.5) and the semigroup property, we obtain for  $t > 1$  that

$$\mu((P_t f) \log P_t f) \leq e^{-\lambda(t-1)} \mu((P_1 f) \log P_1 f) \leq C_1 e^{-\lambda(t-1)} W_2(f\mu, \mu)^2.$$

Taking into account (3.5) we arrive at the estimate

$$\mu((P_t f) \log P_t f) \leq \frac{C_2}{t} W_2(f\mu, \mu)^2, \quad t > 0, f \geq 0, \mu(f) = 1 \quad (3.6)$$

for some constant  $C_2 > 0$ . Let  $q' = \frac{q}{q-1}$ . It is known (see, e.g., [16, Corollary 4.2]) that for each  $p > 1$  there exists a constant  $C(p) > 0$  depending only on  $K$  and  $p$  such that

$$|\nabla P_t h| \leq \frac{C(p)}{\sqrt{t \wedge 1}} (P_t |h|^{q'})^{\frac{1}{q'}}, \quad h \in C_b(M), t > 0. \quad (3.7)$$

This implies that

$$\|\nabla P_t h\|_{q'} \leq \frac{C(p)}{\sqrt{t \wedge 1}} \|h\|_{q'}, \quad t > 0, h \in C_b(M). \quad (3.8)$$

Since the log-Sobolev inequality implies that  $L$  has a spectral gap larger than  $\lambda$ , we have

$$\|\nabla P_t h\|_2 \leq e^{-\lambda t} \|\nabla h\|_2, \quad t \geq 0, h \in W^{1,2}(\mu).$$

This along with (3.8) yields that

$$\|\nabla P_t h\|_2 \leq \frac{C_3 e^{-\lambda t}}{\sqrt{t \wedge 1}} \|h\|_2, \quad t > 0, h \in C_b(M)$$

for some constant  $C_3 > 0$  depending only on  $\lambda$  and  $K$ . Combining this with (3.8) and using the interpolation theorem, we find that

$$\|\nabla P_t h\|_{q'} \leq \frac{C_4}{\sqrt{t}} \|h\|_{q'}, \quad t > 0, h \in C_b(M) \quad (3.9)$$

for some constant  $C_4 > 0$  depending only on  $\lambda, K$  and  $q$ . Noting that for  $h \in C_0^2(M)$  we have

$$|\mu(h(f - P_t f))| = \left| \int_0^t \mu(f L P_s f) ds \right| = \left| \int_0^t \mu(\langle \nabla P_s h, \nabla f \rangle) ds \right|, \quad t > 0,$$

this implies

$$\|P_t f - f\|_q \leq \|f\|_1 \int_0^t \frac{C_4}{\sqrt{s}} ds = 2C_4 \sqrt{t} \|\nabla f\|_q, \quad t > 0.$$

Therefore,

$$\mu(f \geq 2u) \leq \mu(|P_t f - f| \geq u) + \mu(P_t f \geq u) \leq \frac{C_5 t^{q/2}}{u^q} \|\nabla f\|_q^q + \mu(f \geq u).$$

According to the reasoning in the proofs of Theorems 2.2 and 2.3, this and (3.6) imply (3.2) and (3.3).  $\square$

**Theorem 3.2.** *Suppose that the hypotheses of Theorem 3.1 hold. If the diffusion process is strong ergodic, that is,*

$$\|P_t h - \mu(h)\|_\infty \leq c e^{-\lambda t} \|h\|_\infty, \quad t \geq 0, h \in L^\infty(\mu) \quad (3.10)$$

*holds for some constants  $c, \lambda > 0$ , then there exists a constant  $C > 0$  such that for any probability density  $f$  with respect to  $\mu$ ,*

$$\|f - 1\|_1^2 \leq C \|\nabla f\|_1 W(f\mu, \mu). \quad (3.11)$$

*Proof.* By (3.7), there exists a constant  $C_1 > 0$  such that

$$\|\nabla P_t h\|_\infty \leq \frac{C_1}{\sqrt{t}} \|h\|_\infty, \quad t > 0, h \in \mathcal{B}_b(M). \quad (3.12)$$

Combining this with (3.10) we obtain the estimate

$$\|\nabla P_{1+t} h\|_\infty = \|\nabla P_{1+t}(h - \mu(h))\|_\infty \leq C_1 \|P_t h - \mu(h)\|_\infty \leq 2C_1 c e^{-\lambda t} \|h\|_\infty, \quad t \geq 0.$$

This estimate and (3.12) imply that

$$\|\nabla P_{1+t}h\|_\infty \leq \frac{C_2}{\sqrt{t}}\|h\|_\infty, \quad t > 0, h \in \mathcal{B}_b(M)$$

for some constant  $C_2 > 0$ . Then, according to the proof of Theorem 2.7, we obtain (3.11).  $\square$

According to [17], the log-Sobolev inequality and the strong ergodicity are incomparable, but both follow from the ultraboundedness:  $\|P_t\|_{1 \rightarrow \infty} < \infty$  for  $t > 0$ . Since  $\text{Ric}_V$  is bounded below, by [14],  $P_t$  is ultracontractive if and only if

$$\|P_t e^{\lambda \rho_o^2}\|_\infty < \infty, \quad t > 0.$$

For instance, it is the case when  $\text{Ric}$  is bounded from below and  $V \approx -\rho_o^p$  with  $p > 2$ , see [14] for details.

#### 4. THE GAUSSIAN CASE

We now consider the partial case where  $\gamma$  is the standard Gaussian measure on the space  $\mathbb{R}^d$ . Let  $\{T_t\}_{t \geq 0}$  denote the corresponding Ornstein–Uhlenbeck semigroup defined on  $L^1(\gamma)$  by the formula

$$T_t f(x) = \int f(e^{-t}x - \sqrt{1 - e^{-2t}}y) \gamma(dy).$$

Let  $W^{q,1}(\gamma)$  denote the weighted Sobolev class of functions  $f \in L^q(\gamma)$  possessing a generalized gradient  $\nabla f$  with  $|\nabla f| \in L^q(\gamma)$ .

**Theorem 4.1.** *For any probability measure  $f d\gamma$  with  $f \in W^{q,1}(\gamma)$  and any  $u \geq 8$  one has*

$$\gamma(f \geq 2u) \leq \inf_{t > 0} \left[ \frac{K_q^q \arccos^q(e^{-t})}{u^q} \|\nabla f\|_q^q + \frac{2}{(e^{2t} - 1)u \ln u} W_2^2(f\gamma, \gamma) \right].$$

The last theorem with the same formulation extends to Gaussian measures on infinite-dimensional spaces, provided that one employs the Sobolev classes associated with the Cameron–Martin space of the given measure (see [4]). In this case the Poincaré-type estimate holds for all  $q \geq 1$ ; this is why we do not assume that  $q \in [1, 2]$  as we did in Theorem 2.2. For the proof of the theorem we need a lemma.

**Lemma 4.2.** *Let  $f \in W^{q,1}(\gamma)$ . Then*

$$\|T_t f - u\|_q \leq K_q c_t \|\nabla f\|_q,$$

where

$$c_t = \int_0^t \frac{e^{-s}}{\sqrt{1 - e^{-2s}}} ds = \arccos(e^{-t}),$$

$$K_q^q = \frac{1}{\sqrt{2\pi}} \int |x|^q e^{-\frac{x^2}{2}} dx.$$

*Proof.* Standard approximation arguments show that it is sufficient to prove the estimate only when  $X$  is a finite-dimensional space and  $u$  is a smooth function. Since

$$\begin{aligned} u(e^{-t}x + \sqrt{1 - e^{-2t}}y) - u(x) &= \int_0^1 \frac{d}{d\tau} u(e^{-t\tau}x + \sqrt{1 - e^{-2t\tau}}y) d\tau \\ &= t \int_0^1 \nabla u(e^{-t\tau}x + \sqrt{1 - e^{-2t\tau}}y) \cdot \left( -e^{-t\tau}x + \frac{e^{-2t\tau}}{\sqrt{1 - e^{-2t\tau}}}y \right) d\tau, \end{aligned}$$

we have

$$\begin{aligned}
\|T_t u - u\|_q^q &\leq \int_{\mathbb{R}^d \times \mathbb{R}^d} |u(e^{-t}x + \sqrt{1 - e^{-2t}}y) - u(x)|^q \gamma(dx) \gamma(dy) \\
&\leq c_t^{p-1} \int_{\mathbb{R}^d \times \mathbb{R}^d} \int_0^1 \frac{te^{-t\tau}}{\sqrt{1 - e^{-2t\tau}}} \\
&\quad \times \left| \nabla u \left( e^{-t\tau}x + \sqrt{1 - e^{-2t\tau}}y \right) \cdot \left( -\sqrt{1 - e^{-2t\tau}}x + e^{-t\tau}y \right) \right|^q d\tau \gamma(dx) \gamma(dy) \\
&= c_t^{q-1} \int_0^1 \frac{te^{-t\tau}}{\sqrt{1 - e^{-2t\tau}}} \\
&\quad \int_{\mathbb{R}^d \times \mathbb{R}^d} \left| \nabla u \left( e^{-t\tau}x + \sqrt{1 - e^{-2t\tau}}y \right) \cdot \left( -\sqrt{1 - e^{-2t\tau}}x + e^{-t\tau}y \right) \right|^q \gamma(dx) \gamma(dy) \\
&= c_t^{q-1} \int_0^1 \frac{te^{-t\tau}}{\sqrt{1 - e^{-2t\tau}}} \int_{\mathbb{R}^d \times \mathbb{R}^d} |\nabla u(x) \cdot y|^q \gamma(dy) \gamma(dx) d\tau \leq K_q^q c_t^q \int_{\mathbb{R}^d} |\nabla u(x)|^q \gamma(dx),
\end{aligned}$$

which completes the proof.  $\square$

Now we can prove Theorem 4.1.

*Proof.* As in Theorem 2.2 (see also Proposition 3.1 in [13]) we have

$$\begin{aligned}
\gamma(f \geq 2u) &\leq \gamma(|f - T_t f| \geq u) + \gamma(T_t f \geq u) \\
&\leq \frac{K_q^q \arccos^q(e^{-t})}{u^q} \|\nabla f\|_q^q + \frac{2}{u \log u} \int_{\mathbb{R}^d} T_t f \log T_t f d\gamma \\
&\leq \frac{K_q^q \arccos^q(e^{-t})}{u^q} \|\nabla f\|_q^q + \frac{2}{(e^{2t} - 1)u \log u} W_2^2(f\gamma, \gamma)
\end{aligned}$$

Now it is trivial to complete the proof.  $\square$

Our work has been supported by the RFBR project 14-01-91158 of the Russian–Chinese collaboration. The last author has been also supported by the grants NNSFC 11131003 and 11431014, the 985 project and the Laboratory of Mathematical and Complex Systems.

## REFERENCES

- [1] Bakry D., Ledoux M. Levy–Gromov’s isoperimetric inequality for an infinite-dimensional diffusion generator. *Invent. Math.* 1996. V. 123. P. 259–281.
- [2] Bakry D., Gentil I., Ledoux M. *Analysis and geometry of Markov diffusion operators.* Springer, Berlin, 2013.
- [3] Bakry D., Gentil I., Ledoux M. On Harnack inequalities and optimal transportation. *Annali Scuola Norm. Super. Pisa.* 2015. V. 14, N 3. P. 705–727.
- [4] Bogachev V.I. *Gaussian measures.* Amer. Math. Soc., Providence, Rhode Island, 1997.
- [5] Bogachev V.I., Kolesnikov A.V. The Monge–Kantorovich problem: achievements, connections, and perspectives. *Uspehi Matem. Nauk.* 2012. V. 67, N 5. P. 3–110 (in Russian); English transl.: *Russian Math. Surveys.* 2012. V. 67, N 5. P. 785–890.
- [6] Bogachev V.I., Shaposhnikov A.V. Lower bounds for the Kantorovich distance. *Dokl. Akad. Nauk.* 2015. V. 460, N 6. P. 631–633 (in Russian); English transl.: *Doklady Math.* 2015. V. 91, N 1. P. 91–93.
- [7] Bogachev V.I., Wang F.-Y., Shaposhnikov A.V. Estimates of the Kantorovich norm on manifolds *Dokl. Akad. Nauk.* 2015. V. 463, N 6. P. 633–638 (in Russian); English transl.: *Doklady Math.* 2015. V. 92, N 1. P. 1–6.
- [8] Chavel I. *Riemannian geometry. A modern introduction.* 2nd ed. Cambridge Univ. Press, Cambridge, 2006.
- [9] Cinti E., Otto F. Interpolation inequalities in pattern formation. Arxiv: 1509.01189v1 (2015).

- [10] Kantorovitch L.V. On the translocation of masses. Dokl. Akad. Nauk SSSR. 1942. V. 37, N 7–8. P. 227–229 (in Russian); English transl.: C. R. (Doklady) Acad. Sci. URSS. 1942. V. 37. P. 199–201.
- [11] Kantorovich L.V., Rubinshtein G.Sh. On a functional space and certain extremum problems. Dokl. Akad. Nauk SSSR. 1957. V. 115, N 6. P. 1058–1061 (in Russian).
- [12] Kohn R. V., Otto F. Upper bounds on coarsening rates. Commun. Math. Phys. 2002. V. 229, N 3. P. 375–395.
- [13] Ledoux M. Sobolev–Kantorovich inequalities. Anal. Geom. Metr. Spaces. 2015. V. 3. P. 157–166.
- [14] Röckner M., Wang F.-Y. Supercontractivity and ultracontractivity for (non-symmetric) diffusion semigroups on manifolds, Forum Math. 2003. V. 15, P. 893–921.
- [15] Röckner M., Wang F.-Y. Log-Harnack Inequality for Stochastic differential equations in Hilbert spaces and its consequences, Inf. Dimen. Anal. Quant. Probab. Rel. Top. 2010. V. 13. P. 27–37.
- [16] Wang F.-Y. Logarithmic Sobolev inequalities on noncompact Riemannian manifolds. Probab. Theory Related Fields. 1997. V. 109. P. 417–424.
- [17] Wang F.-Y.  $L^1$ -convergence and hypercontractivity of diffusion semigroups on manifolds, Studia Math. 2004. V. 162. P. 219–227.
- [18] Wang F.-Y. Functional inequalities, Markov properties and spectral theory. Science Press, 2005.
- [19] Wang, F.-Y. Harnack inequalities on manifolds with boundary and applications, J. Math. Pures Appl. 2010. V. 94. P. 304–321.
- [20] Wang F.-Y. Analysis for diffusion processes on Riemannian manifolds. World Sci., Singapore, 2013.

V.I. Bogachev: Department of Mechanics and Mathematics, Moscow State University, 119991 Moscow, Russia, and Orthodox S.-Tikhon’s Humanitarian University, Moscow, Russia;

Email: vibogach@mail.ru

F.-Yu. Wang: School of Mathematical Sciences, Beijing Normal University, Beijing 100875, China, and Department of Mathematics, Swansea University, Singleton Park, SA2 8PP, UK

Email: wangfy@bnu.edu.cn; F.Y.Wang@swansea.ac.uk

A.V. Shaposhnikov: Department of Mechanics and Mathematics, Moscow State University, 119991 Moscow, Russia