DE GRUYTER OPEN Anal. Geom. Metr. Spaces 2015; 3:157–166

Research Article Open Access

Michel Ledoux*

Sobolev-Kantorovich Inequalities

DOI 10.1515/agms-2015-0011
Received March 7, 2015; accepted May 20, 2015

Abstract: In a recent work, E. Cinti and F. Otto established some new interpolation inequalities in the study of pattern formation, bounding the $\mathrm{L}^r(\mu)$-norm of a probability density with respect to the reference measure μ by its Sobolev norm and the Kantorovich-Wasserstein distance to μ. This article emphasizes this family of interpolation inequalities, called Sobolev-Kantorovich inequalities, which may be established in the rather large setting of non-negatively curved (weighted) Riemannian manifolds by means of heat flows and Harnack inequalities.

Keywords: Interpolation inequality; Sobolev norm; Kantorovich distance; heat flow; Harnack inequality

MSC: Primary: 35K08, 60J60, 58J60, 53C21 Secondary: 46E35, 35B65

1 Introduction

In the recent contribution [6], E. Cinti and F. Otto established a number of interpolation inequalities which arise in the analysis of pattern formation in physics (more precisely in the study of branching in superconductors). Within this framework, they showed in particular that for any positive periodic smooth function $f : [0,1]^n \to \mathbb{R}$ such that $\int_{[0,1]^n} f dx = 1$,

$$\left\|(f - C)_+\right\|_r^\theta \le C \, \|\nabla f\|_1 \; \mathrm{W}_2(f, 1) \tag{1.1}$$

where

$$r = \frac{3n + 2}{3n},$$

$\theta = \frac{3n+2}{2n}$ and $C > 0$ only depends on n. Here $(f - C)_+ = \max(f - C, 0)$ and $\mathrm{W}_2(f, 1)$ is the Kantorovich-Wasserstein distance between the probability measures fdx and $1\,dx$ (see below). The proof of this inequality in [6] relies on a specific geometric construction of cut-off functions first put forward in [5]. The exponent r is optimal as verified on peak functions (and, as explained in [6], from this point of view the Kantorovich distance behaves like a negative fractional Sobolev norm). Note furthermore that (1.1) is invariant under the change $\int_{[0,1]^n} dx \mapsto \frac{1}{\Lambda^n} \int_{[0,\Lambda]^n} dx$ for any $\Lambda > 0$.

The purpose of this work is to show that the inequality (1.1) is actually one specimen in a all family of interpolation inequalities that we call Sobolev-Kantorovich inequalities and that actually hold in a wide generality.

One suitable framework for such Sobolev-Kantorovich inequalities is the one of (weighted) Riemannian manifolds with non-negative Ricci curvature, covering in particular the previous model example of the torus. Let $M = (M, g)$ be a complete connected n-dimensional Riemannian manifold with Riemannian measure dx, and let, as reference measure, μ be a probability measure on M (with smooth density with respect to dx). The counterpart of the Euclidean geometric decomposition in [6] will be achieved via non-negative curvature conditions, specifically curvature-dimension $CD(0, N)$ of the underlying diffusion operator invariant with respect to μ, or metric measure space (M, g, μ), in the sense of [1, 10, 13] (see Section 2). In case of an n-dimensional

***Corresponding Author: Michel Ledoux:** Institut de Mathématiques de Toulouse, Université de Toulouse–Paul-Sabatier, F-31062 Toulouse, France, and Institut Universitaire de France, E-mail: ledoux@math.univ-toulouse.fr

Riemannian manifold (with finite, normalized, volume element $dx = d\mu$), the $CD(0, N)$ condition with $N = n$ amounts to a non-negative lower bound on the Ricci curvature.

Given two probability measures μ and ν on the Borel sets of M, and $p \geq 1$, the Kantorovich-Wasserstein distance $\mathrm{W}_p(\nu, \mu)$ between ν and μ is defined by

$$\mathrm{W}_p(\nu, \mu) = \inf \left(\int_{M \times M} d(x, y)^p \, d\pi(x, y) \right)^{1/p}$$

where the infimum is taken over all couplings π on the product space $M \times M$ with respective marginals ν and μ and where $d(x, y)$ denotes the Riemannian distance between x and y in M (cf. [13]).

Given the probability μ, the norm of the Lebesgue space $\mathrm{L}^r(\mu)$, $r \geq 1$, is denoted by $\|\cdot\|_r$.

With these notation, the general family of Sobolev-Kantorovich interpolation inequalities emphasized in this work take the following form.

Theorem 1.1. *Let (M, g, μ) be a weighted Riemannian manifold with weighted probability measure μ satisfying the curvature-dimension condition $CD(0, N)$ for some $N \geq 1$. Given $p, q \geq 1$, there is a constant $C > 0$ only depending on p, q, N such that for any probability $d\nu = f d\mu$ with smooth density f with respect to μ,*

$$\big\|(f - C)_+\big\|_r^\theta \leq C \, \|\nabla f\|_q \, \mathrm{W}_p(\nu, \mu) \tag{1.2}$$

where

$$r = \frac{1 + \frac{1}{p} + \frac{1}{N}}{\frac{1}{p} + \frac{1}{q}} \quad (> 1)$$

and $\theta = r\big(\frac{1}{p} + \frac{1}{q}\big) = 1 + \frac{1}{p} + \frac{1}{N}$.

The Cinti-Otto inequality (1.1) on the flat n-dimensional torus corresponds to $q = 1$, $p = 2$ and $N = n$ for which therefore $r = \frac{3n+2}{3n}$. Other values of interests are $q = p = 2$ for which $r = \frac{3N+2}{2N}$. In this case, the inequality may be compared to the Poincaré-Sobolev inequality (see Section 2)

$$\|f\|_r \leq C \, D(M) \, \|\nabla f\|_2 \tag{1.3}$$

for mean-zero smooth functions f, where $r = \frac{2N}{N-2}$ $(N > 2)$ and $D(M)$ is the diameter of the manifold, yielding a stronger embedding under a stronger interpolation factor $D(M) \geq \mathrm{W}_2(\nu, \mu)$.

In a somewhat weaker form, (1.2) may also be expressed by

$$\|f\|_r \leq C + C \, \|\nabla f\|_q^{1/\theta} \, \mathrm{W}_p(\nu, \mu)^{1/\theta}.$$

It is natural to expect that $C = 1 = \int_M f d\mu$ on the left-hand side of (1.2) but the proof below is far from such a claim. The exponents in Theorem 1.1 are optimal on peak functions (in dimension one).

The proof of Theorem 1.1 will be based on heat flow arguments on the semigroup $(P_t)_{t\geq 0}$ with invariant measure μ underlying the weighted manifold (M, g, μ), and slicing methods in the spirit of [6] and [8]. The $CD(0, N)$ condition will allow for the use of heat kernel and Harnack inequalities, consequences of the Li-Yau parabolic inequality in this framework. These tools are actually also available in the more general context of Markov Triples (E, μ, Γ) in the sense of [1]. For the specific torus example, it is plausible that the Euclidean arguments of [6] also allow for the full range of parameters.

Infinite dimensional examples of curvature-dimension $CD(0, \infty)$ may be addressed similarly. For $CD(0, \infty)$ models, the classical Li-Yau parabolic inequality is however no more available and has to be supplemented by Wang's Harnack inequality (cf. [1, 15]). The resulting Sobolev-Kantorovich inequalities (Theorem 4.1) are then dimension free, and the values $q = p = 2$ read in particular

$$\big\|(f - C)_+\big\|_{3/2}^{3/2} \leq C \, \|\nabla f\|_2 \, \mathrm{W}_2(\nu, \mu) \tag{1.4}$$

for some numerical $C > 0$. Under a diameter bound, this result turns into a Nash-type inequality of possible independent interest,

$$\|f\|_{3/2} \le C\big[\|f\|_1^2 + D(M)^2\|\nabla f\|_2^2\big]^{1/3}\|f\|_1^{1/3}$$

for every (smooth) $f : M \to \mathbb{R}$.

In the form (1.4), Sobolev-Kantorovich interpolation inequalities are also in the spirit of the Otto-Villani HWI inequality [1, 12, 13] stating in this context that for any probability $d\nu = f d\mu$,

$$\int_M f \log f \, d\mu \le \bigg(\int_M \frac{|\nabla f|^2}{f}\, d\mu\bigg)^{1/2} \mathrm{W}_2(\nu, \mu).$$

To outline the content of the article, we present in the next section the framework of investigation and the basic geometric tools which will be used in the proof of Theorem 1.1. These rely on the curvature-dimension condition $CD(0, N)$ in the form of Harnack and pseudo-Poincaré inequalities. Section 3 is then devoted to the proof of Theorem 1.1 which adapts, with the preceding geometric tools and a slicing argument, the scheme emphasized in [6]. The subsequent section addresses the infinite-dimensional case under $CD(0, \infty)$ with the tool of Wang's Harnack inequality.

2 Framework and geometric tools

For simplicity in the exposition, the results of this work are presented in the weighted Riemannian setting, that is a complete connected Riemannian manifold (M, g) equipped with a weighted probability measure $d\mu = \mathrm{e}^{-V}dx$ where $V : M \to \mathbb{R}$ is a smooth potential, to which is naturally attached the diffusion operator $\mathrm{L} = \Delta - \nabla V \cdot \nabla$ with invariant and reversible measure μ and associated heat semigroup $(P_t)_{t\ge 0}$.

The curvature-dimension $CD(0, N)$, for some $N \ge 1$, condition on L, or the weighted Riemannian manifold (M, g, μ), is described by the Bochner-type inequality

$$\frac{1}{2}\,\mathrm{L}\big(|\nabla f|^2\big) - \nabla f \cdot \nabla \mathrm{L} f \ge \frac{1}{N}\,(\mathrm{L}f)^2$$

holding for all smooth $f : M \to \mathbb{R}$. For example, the Laplace-Beltrami operator Δ on an n-dimensional Riemannian manifold with non-negative Ricci curvature satisfies $CD(0, n)$. The curvature condition $CD(0, \infty)$ covers infinite dimensional models in the sense of the operator L, such as for example log-concave measures on $\mathbb{R}^n$. We refer to [1, 10, 11, 13] for standard references on such a framework.

As mentioned in the introduction, the Sobolev-Kantorovich interpolation inequalities emphasized here may actually be developed in the more general setting of a Markov diffusion Triple (E, μ, Γ) in the sense of [1], consisting of a state space E equipped with a diffusion semigroup $(P_t)_{t\ge 0}$ with infinitesimal generator L and carré du champ operator Γ, and invariant and reversible probability measure μ. In the weighted Riemannian context, $\Gamma(f, f) = |\nabla f|^2$ for smooth functions f. The curvature condition $CD(0, N)$ in this context stems from the abstract form of the Bochner identity in terms of the Γ_2 operator. The state space E may be endowed with an intrinsic distance d for which Lipschitz functions f are such that $\Gamma(f)$ is bounded (μ-almost everywhere), which coincides with the Riemannian distance in the manifold case.

One basic tool in the heat kernel estimates necessary in this investigation will be the Li-Yau parabolic inequality [9] in the form of its following consequences. These results, established first on Riemannian manifolds with non-negative curvature, extend to the case of weighted Riemannian manifolds and more generally Markov Triples under the curvature-dimension condition $CD(0, N)$ (cf. [1]).

The first main property of interest is the following Harnack inequality (cf. [1, 7, 9]).

Proposition 2.1. *Let (M, g, μ) be a weighted Riemannian manifold satisfying the curvature-dimension condition $CD(0, N)$. For any non-negative (measurable) function f on M, and any $t, s > 0$, $x, y \in M$,*

$$P_t f(y) \le P_{t+s} f(x) \Big(\frac{t+s}{t}\Big)^{N/2} \mathrm{e}^{d(x,y)^2/4s}.$$

The second main tool is a pseudo-Poincaré inequality.

Proposition 2.2. *Let (M, g, μ) be a weighted Riemannian manifold satisfying the curvature-dimension condition $CD(0, N)$. For any smooth function f on M, and any $t > 0$,*

$$\|f - P_t f\|_q \leq B\sqrt{t}\,\|\nabla f\|_q$$

where $B > 0$ is numerical (for example $B = \sqrt{2}$) for $1 \leq q \leq 2$ and only depends on N for $q \geq 2$.

About the proof of this proposition (see e.g. Lemma 3 in [8]), it should be pointed out that the case $1 \leq q \leq 2$ indeed does not require the dimensional Li-Yau inequality and rather relies on Poincaré inequalities for heat kernel measures valid under the only curvature condition $CD(0, \infty)$. Specifically, the argument involves reverse Poincaré inequalities (cf. [1]) which express, under this curvature, that for any (measurable) $h : M \to \mathbb{R}$, and any $s > 0$, pointwise

$$2s\,|\nabla P_s h|^2 \leq P_s(h^2) - (P_s h)^2. \tag{2.1}$$

Provided indeed with this result, for any (say bounded) h,

$$\int_M h(f - P_t f)d\mu = -\int_0^t \int_M h\,\mathrm{L}P_s f\,d\mu ds = \int_0^t \int_M \nabla P_s h \cdot \nabla f\,d\mu ds.$$

If $2 \leq q' \leq \infty$, by Jensen's inequality on (2.1),

$$\|\nabla P_s h\|_{q'} \leq \frac{1}{\sqrt{2s}}\,\|h\|_{q'}$$

so that by duality, if $1 \leq q \leq 2$,

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

To complete this section, we provide a direct proof of the Poincaré-Sobolev inequality (1.3) relying on the preceding tools as we could not locate a suitable reference. For a given $f : M \to \mathbb{R}$, by Proposition 2.2 with $q = 2$, for any $t \geq 0$,

$$\|f\|_2 \leq \sqrt{2t}\,\|\nabla f\|_2 + \|P_t f\|_2.$$

On the other hand, by Proposition 2.1 with $s = D^2 = D(M)^2$, for every non-negative f on M, and every $x, y \in M$ and $t \geq 0$,

$$P_t f(y) \leq 2\left(1 + \frac{D^2}{t}\right)^{N/2} P_{t+D^2} f(x).$$

Hence, after integration in the x variable, $\|P_t f\|_\infty \leq 2\big(1 + \frac{D^2}{t}\big)^{N/2}\|f\|_1$ and, by interpolation, $\|P_t f\|_2 \leq \sqrt{2}\big(1 + \frac{D^2}{t}\big)^{N/4}\|f\|_1$.

Combining the preceding, for any (smooth) f and every $t \geq 0$,

$$\|f\|_2 \leq \sqrt{2t}\,\|\nabla f\|_2 + \sqrt{2}\left(1 + \frac{D^2}{t}\right)^{N/4}\|f\|_1.$$

Optimizing in $t > 0$ yields the Nash-type inequality

$$\|f\|_2 \leq C\big[\|f\|_2 + D\,\|\nabla f\|_2\big]^{N/(N+2)}\|f\|_1^{2/(N+2)}$$

where $C > 0$ only depends on N. By a standard slicing procedure (see [1, Chap. 6]), the latter amounts to the Sobolev inequality

$$\|f\|_r \leq C'\big[\|f\|_2 + D\,\|\nabla f\|_2\big]$$

for $r = \frac{2N}{N-2}$ ($N > 2$). Now, it is also classical in this context (cf. [1, 3]) that a Poincaré inequality for mean-zero functions f holds true with constant proportional to the diameter, namely

$$\|f\|_2 \leq C'' D\,\|\nabla f\|_2$$

for $C'' > 0$ only depending on the dimension N. Hence, the announced Poincaré-Sobolev inequality (1.3) follows.

3 Proof of Theorem 1.1

This section is devoted to the proof of the main result, Theorem 1.1. The general scheme follows the approach developed by E. Cinti and F. Otto [6]. To get a better feeling about the argument, we start with the simpler weak-type version of the statement. Actually, as put forward and achieved in [6], a main step will be to pass from the weak formulation to the strong one.

Proposition 3.1. *In the setting of Theorem 1.1, there is constant $C > 0$ only depending on p, q, N such that*

$$\sup_{u \geq C} u^\theta \, \mu(f \geq u)^{\theta/r} \leq C \, \|\nabla f\|_q \, \mathrm{W}_p(\nu, \mu).$$

Proof. Let thus f be a smooth probability density on M, and recall $d\nu = f d\mu$. By Proposition 2.2, for every $u > 0$ and $t > 0$,

$$\mu(f \geq 2u) \leq \mu\big(|f - P_t f| \geq u\big) + \mu(P_t f \geq u) \leq \frac{B^q t^{q/2}}{u^q} \, \|\nabla f\|_q^q + \mu(P_t f \geq u). \tag{3.1}$$

Now, by Markov's inequality and symmetry of $(P_t)_{t \geq 0}$ with respect to μ,

$$\mu(P_t f \geq u) \leq \frac{1}{u} \int_M \mathbb{1}_F P_t f \, d\mu = \frac{1}{u} \int_M P_t(\mathbb{1}_F) d\nu$$

where $F = \{P_t f \geq u\}$. By the Kantorovich duality (cf. [13]), for every $\varepsilon > 0$,

$$\int_M P_t(\mathbb{1}_F) d\nu \leq \frac{1}{\varepsilon} \mathrm{W}_p^p(\nu, \mu) + \int_M \widehat{Q}_\varepsilon P_t(\mathbb{1}_F) d\mu$$

where $\widehat{Q}_\varepsilon$ is the sup-convolution

$$\widehat{Q}_\varepsilon P_t(\mathbb{1}_F)(x) = \sup_{y \in M} \Big[P_t(\mathbb{1}_F)(y) - \frac{1}{\varepsilon} \, d(x, y)^p \Big].$$

Clearly, for any $x \in M$, the supremum may be restricted to those y for which $d(x, y)^p \leq \varepsilon$.

Use next the Harnack inequality from Proposition 2.1 with s replaced by ts to get that for every $s > 0$ and $x, y \in M$,

$$P_t(\mathbb{1}_F)(y) \leq P_{t(s+1)}(\mathbb{1}_F)(x)(1 + s)^{N/2} \, \mathrm{e}^{d(x,y)^2/4ts}.$$

In particular, for every y such that $d(x, y)^p \leq \varepsilon$ with $\varepsilon = (ts)^{p/2}$,

$$P_t(\mathbb{1}_F)(y) \leq 2(1 + s)^{N/2} P_{t(s+1)}(\mathbb{1}_F)(x)$$

so that, with this choice of ε,

$$\widehat{Q}_\varepsilon P_t(\mathbb{1}_F)(x) \leq 2(1 + s)^{N/2} P_{t(s+1)}(\mathbb{1}_F)(x).$$

Therefore, after integration with respect to μ,

$$\mu(F) = \mu(P_t f \geq u) \leq \frac{1}{u(ts)^{p/2}} \, \mathrm{W}_p^p(\nu, \mu) + \frac{2}{u} \, (1 + s)^{N/2} \mu(F).$$

Choose $s = 2^{-5} u^{2/N}$ so that $\frac{2}{u} \, (1 + s)^{N/2} \leq \frac{1}{2}$ provided that $u \geq C = C_N$ large enough. Hence, together with (3.1),

$$\mu(f \geq 2u) \leq \frac{B^q t^{q/2}}{u^q} \, \|\nabla f\|_q^q + \frac{2^{(5p/2)+1}}{t^{p/2} u^{1+(p/N)}} \, \mathrm{W}_p^p(\nu, \mu).$$

Optimizing in $t > 0$ yields the conclusion. □

On the basis of the preceding weak-type estimate, we address the proof of Theorem 1.1 together with a standard slicing argument as extensively presented in [1] for related functional inequalities.

Starting as above, for every $u > 0$ and $t > 0$, by Proposition 2.2,

$$\begin{aligned}\mu(f \geq 2u) &\leq \mu(f \geq 2u, P_t f \leq u) + \mu(f \geq 2u, P_t f \geq u)\\ &\leq \mu\big(|f - P_t f| \geq u\big) + \frac{1}{2u}\int_M \mathbb{1}_F f\, d\mu\\ &\leq \frac{B^q\, t^{q/2}}{u^q}\,\|\nabla f\|_q^q + \frac{1}{2u}\int_M \mathbb{1}_F f\, d\mu\end{aligned}$$

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

Apply the preceding (which holds true for any smooth – Lipschitz – positive function, not only probability densities) to $f_k = \min((f - 2^k)_+, 2^k)$, $k \in \mathbb{Z}$, and $u = 2^{k-1}$, $t = t_k > 0$, so to get

$$\begin{aligned}\mu(f_k \geq 2^k) &\leq \frac{(2B)^q\, t_k^{q/2}}{2^{qk}}\int_{A_k} |\nabla f|^q d\mu + \frac{1}{2^k}\int_M \mathbb{1}_{F_k} f_k d\mu\\ &\leq \frac{(2B)^q\, t_k^{q/2}}{2^{qk}}\int_{A_k} |\nabla f|^q d\mu + \frac{1}{2^k}\int_M \mathbb{1}_{F_k} d\nu\end{aligned}$$

where $A_k = \{2^k \leq f \leq 2^{k+1}\}$, $F_k = \{P_{t_k} f_k \geq 2^{k-1}\}$, $k \in \mathbb{Z}$, and where we used that $f_k \leq f$ in the last step. Note that $\mu(f \geq 2^{k+1}) = \mu(f_k \geq 2^k)$, and sum the preceding inequalities, after multiplication by 2^{rk}, over the set of integers $I = \{k_0, k_0 + 1, \ldots, k_1\}$ for some $k_0 \geq 0$ to be determined below (and where k_1 is arbitrarily large, tending to infinity at the end of the argument). Hence,

$$\sum_{k\in I} 2^{rk}\mu(f \geq 2^{k+1}) \leq (2B)^q \sum_{k\in I} 2^{(r-q)k}\, t_k^{q/2}\int_{A_k} |\nabla f|^q d\mu + \int_M \phi\, d\nu \tag{3.2}$$

where $\phi = \sum_{k\in I} 2^{(r-1)k}\,\mathbb{1}_{F_k}$.

By the Kantorovich duality, for any $\varepsilon > 0$,

$$\int_M \phi\, d\nu \leq \frac{1}{\varepsilon}\,\mathrm{W}_p^p(\nu, \mu) + \int_M \widehat{Q}_\varepsilon \phi\, d\mu \tag{3.3}$$

where

$$\widehat{Q}_\varepsilon P_t \phi(x) = \sup_{y\in M}\Big[\phi(y) - \frac{1}{\varepsilon}\, d(x,y)^p\Big] = \sup_{y\in M}\Big[\sum_{k\in I} 2^{(r-1)k}\,\mathbb{1}_{F_k}(y) - \frac{1}{\varepsilon}\, d(x,y)^p\Big]_+ .$$

It is an elementary, yet crucial, observation emphasized in Claim B in the proof of Proposition 1.3 of [6], that

$$\sum_{k\in I} 2^{(r-1)k}\,\mathbb{1}_{F_k}(y) \leq c\,\sup_{k\in I} 2^{(r-1)k}\,\mathbb{1}_{F_k}(y) \tag{3.4}$$

for some $c = c_r > 0$ only depending on $r > 1$. Therefore, for every fixed $x \in M$,

$$\begin{aligned}\widehat{Q}_\varepsilon \phi(x) &\leq c \sup_{y\in M}\sup_{k\in I}\Big[2^{(r-1)k}\,\mathbb{1}_{F_k}(y) - \frac{1}{c\varepsilon}\, d(x,y)^p\Big]_+\\ &\leq c \sup_{k\in I}\sup_{y\in M}\Big[2^{(r-1)k}\,\mathbb{1}_{F_k}(y) - \frac{1}{c\varepsilon}\, d(x,y)^p\Big]_+\\ &\leq c \sum_{k\in I}\sup_{y\in M_k(x)} 2^{(r-1)k}\,\mathbb{1}_{F_k}(y)\end{aligned} \tag{3.5}$$

where $M_k(x) = \{y \in E;\, d(x,y)^p \leq c\varepsilon\, 2^{(r-1)k}\}$, $k \in I$.

Apply next the Harnack inequality from Proposition 2.1 with $t = t_k > 0$ to be specified and $s = t_k s_k = (c\varepsilon)^{2/p} 2^{2(r-1)k/p}$. Recalling that $F_k = \{P_{t_k} f_k \geq 2^{k-1}\}$, for every $y \in M_k(x)$,

$$\mathbb{1}_{F_k}(y) \leq 2^{-k+1} P_{t_k} f_k(y) \leq 2^{-k+2}\,(1 + s_k)^{N/2} P_{t_k(s_k+1)} f_k(x).$$

As a consequence of (3.5),

$$\widehat{Q}_\varepsilon \phi(x) \leq 4c \sum_{k\in I} 2^{(r-2)k}(1 + s_k)^{N/2} P_{t_k(s_k+1)} f_k(x).$$

Integrating with respect to μ,

$$\int_M \widehat{Q}_\varepsilon \phi \, d\mu \leq 4c \sum_{k \in I} 2^{(r-1)k} (1+s_k)^{N/2} \mu(f \geq 2^k)$$

where we used that $\int_M P_{t_k(s_k+1)} f_k d\mu = \int_M f_k d\mu \leq 2^k \mu(f \geq 2^k)$.

Choose $t_k = \lambda\, 2^{-2(r-q)k/q}$ and replace ε by $\lambda^{p/2}\varepsilon$ where $\lambda > 0$. Therefore

$$\int_M \widehat{Q}_{\lambda^{p/2}\varepsilon} \phi \, d\mu \leq 4c \sum_{k \in I} 2^{(r-1)k} \left[1 + (c\varepsilon)^{2/p}\, 2^{r'k}\right]^{N/2} \mu(f \geq 2^k)$$

where

$$r' = \frac{2}{p}(r-1) + \frac{2}{q}(r-q).$$

On the other hand

$$\sum_{k \in I} 2^{(r-q)k}\, t_k^{q/2} \int_{A_k} |\nabla f|^q d\mu \leq \lambda^{q/2} \sum_{k \in I} \int_{A_k} |\nabla f|^q d\mu \leq \lambda^{q/2} \int_M |\nabla f|^q d\mu$$

(since for a smooth $f : M \to \mathbb{R}$, $\nabla f = 0$ on every level set $\{f = a\}$).

Summarizing together with (3.2) and (3.3),

$$\sum_{k \in I} 2^{rk} \mu(f \geq 2^{k+1}) \leq (2B)^q \lambda^{q/2} \, \|\nabla f\|_q^q + \frac{1}{\lambda^{p/2}\varepsilon} \, \mathrm{W}_p^p(\nu, \mu) + 4c \sum_{k \in I} 2^{(r-1)k} \left[1 + (c\,\varepsilon)^{2/p}\, 2^{r'k}\right]^{N/2} \mu(f \geq 2^k).$$

The final step will be to absorb the sum over $k \in I$ on the right-hand side by the one on the left-hand side. To this task, note that $r = (r-1) + \frac{N}{2} r'$ by the very definition of r. For any $\eta > 0$, there exist $\varepsilon = \varepsilon(\eta) > 0$ and $k_0 = k_0(\eta) \geq 0$ large enough (with further dependence only on p, r, N) so that

$$2^{(r-1)k} \left[1 + (c\varepsilon)^{2/p}\, 2^{r'k}\right]^{N/2} \leq \eta\, 2^{rk}$$

for every $k \geq k_0$. Hence,

$$\sum_{k \in I} 2^{rk} \mu(f \geq 2^{k+1}) \leq (2B)^q \lambda^{q/2} \, \|\nabla f\|_q^q + \frac{1}{\lambda^{p/2}\varepsilon} \, \mathrm{W}_p^p(\nu, \mu) + 4c\eta \sum_{k \in I} 2^{rk} \mu(f \geq 2^k).$$

If $\eta > 0$ is chosen so that $2^{r+3} c\eta \leq 1$, then

$$\sum_{k \in I} 2^{rk} \mu(f \geq 2^{k+1}) \leq 2(2B)^q \lambda^{q/2} \, \|\nabla f\|_q^q + \frac{2}{\lambda^{p/2}\varepsilon} \, \mathrm{W}_p^p(\nu, \mu) + c\eta\, 2^{rk_0+3} \mu(f \geq 2^{k_0}).$$

To conclude, optimize in $\lambda > 0$ and make use of the weak-type Proposition 3.1 to control the last term on the right-hand side of this inequality. It follows that, for k_0 large enough, there is a constant $C_1 > 0$ depending only on p, q, N such that 5

$$\sum_{k \in I} 2^{rk} \mu(f \geq 2^{k+1}) \leq C_1 \, \|\nabla f\|_q^{r/\theta} \, \mathrm{W}_p(\nu, \mu)^{r/\theta}. \tag{3.6}$$

Finally, it is an easy exercise to check that

$$\begin{aligned} \left\|(f-C)_+\right\|_r^r &= \int_0^\infty \mu(f \geq C + t) d(t^r) \\ &\leq 2^{(k_0+1)r} \mu(f \geq C) + (2^{2r} - 2^r) \sum_{k \geq k_0} 2^{rk} \mu(f \geq 2^{k+1}) \\ &\leq 2^{2r} \sum_{k \geq k_0} 2^{rk} \mu(f \geq 2^{k+1}) \end{aligned}$$

with for example $C = 2^{k_0+1}$. Therefore, letting $k_1 \to \infty$ in (3.6) then yields the conclusion. The proof of Theorem 1.1 is complete.

4 The case of the $CD(0, \infty)$ condition

As discussed in the introduction, the dimensional Harnack inequality from Proposition 2.1 is no more available under the infinite-dimensional non-negative curvature $CD(0, \infty)$ condition and has to be substituted by Wang's Harnack inequality [1, 14, 15]. The latter expresses, under thus $CD(0, \infty)$, that for any positive (measurable) f on M, and all $t > 0$, $x, y \in M$,

$$P_t f(y)^2 \leq P_t(f^2)(x)\, \mathrm{e}^{d(x,y)^2/2t}. \tag{4.1}$$

With this tool in hand, we may state the main result in the $CD(0, \infty)$ case. Only $p = 2$ is considered. Besides, values of q have to be restricted to the interval $[1, 2]$ by Proposition 2.2. In addition, since (3.4) does not hold for $r = 1$, the value $q = 1$ has also to be excluded (although a weak-type estimate holds in this case).

Theorem 4.1. *Let (M, g, μ) be a weighted Riemannian manifold with weighted probability measure μ satisfying the curvature condition $CD(0, \infty)$. Given $1 < q \leq 2$, there is a constant $C > 0$ only depending on q such that for any probability $d\nu = f d\mu$ with smooth density f with respect to μ,*

$$\big\|(f - C)_+\big\|_r^{3/2} \leq C\, \|\nabla f\|_q\, \mathrm{W}_2(\nu, \mu) \tag{4.2}$$

where

$$r = \frac{3q}{q+2}.$$

As in the preceding section, we start with a weak-type estimate, actually valid for any $q \in [1, 2]$. The main argument relies here on a somewhat different use of the Harnack inequality, namely a reverse transportation cost inequality along the flow (cf. [1, 2]).

Proposition 4.2. *In the setting of Theorem 4.1, there is a constant $C > 0$ only depending on $q \in [1, 2]$ such that*

$$\sup_{u \geq C} u^{3/2} \mu(f \geq u)^{3/2r} \leq C\, \|\nabla f\|_q\, \mathrm{W}_2(\nu, \mu).$$

Proof. As in the proof of Proposition 3.1, for every $u > 0$ and $t > 0$,

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

Recall now the entropic inequality stating that for any (say bounded) measurable h on M,

$$\int_M h P_t f\, d\mu \leq \int_M P_t f \log P_t f\, d\mu + \log \int_M \mathrm{e}^h d\mu.$$

Applying it to $h = \mathbb{1}_F$ where $F = \{P_t f \geq u\}$ yields

$$u\mu(F) \leq \int_M \mathbb{1}_F P_t f\, d\mu \leq \int_M P_t f \log P_t f\, d\mu + \log\big(1 + (\mathrm{e} - 1)\mu(F)\big)$$

Hence, provided that $u \geq 4$ (for example),

$$u\, \mu(P_t f \geq u) = u\mu(F) \leq 2 \int_M P_t f \log P_t f d\mu.$$

Now, as a consequence of Wang's Harnack inequality (4.1), the entropy of $P_t f$ is bounded from above by the Kantorovich distance $\mathrm{W}_2(\nu, \mu)$ as (see e.g. (1.11) in [2])

$$\int_M P_t f \log P_t f\, d\mu \leq \frac{1}{4t}\, \mathrm{W}_2^2(\nu, \mu).$$

Therefore, for any $u \geq 4$,

$$\mu(P_t f \geq u) \leq \frac{1}{2tu}\, \mathrm{W}_2^2(\nu, \mu).$$

As a conclusion,

$$\mu(f \geq 2u) \leq \frac{(2t)^{q/2}}{u^q} \, \|\nabla f\|_q^q + \frac{1}{2tu} \, \mathrm{W}_2^2(\nu, \mu)$$

which is the result of the proposition after optimization in $t > 0$. □

Proof of Theorem 4.1. We follow the proof of Theorem 1.1 until (3.5) at which point we make use of (4.1). Recall that $F_k = \{P_{t_k} f_k \geq 2^{k-1}\}$. With $t_k = c\varepsilon\, 2^{(r-1)k}$, for every $y \in M_k(x)$,

$$\mathbb{1}_{F_k}(y) \leq 2^{-2k+2} P_{t_k}(f_k)(y)^2 \leq 2^{-2k+3}\, P_{t_k}(f_k^2)(x).$$

Hence,

$$\int_M \widehat{Q}_\varepsilon \phi \, d\mu \leq 8c \sum_{k \in I} 2^{(r-1)k} \mu(f \geq 2^k)$$

where we used that $\int_M f_k^2 d\mu \leq 2^{2k} \mu(f \geq 2^k)$.

Now, going back to (3.2), the value of r ensures that $2^{(r-q)k}\, t_k^{q/2} = (c\varepsilon)^{q/2}$ so that

$$\sum_{k \in I} 2^{rk} \mu(f \geq 2^{k+1}) \leq 2^{2q} (c\varepsilon)^{q/2} \, \|\nabla f\|_q^q + \frac{1}{\varepsilon} \, \mathrm{W}_2^2(\nu, \mu) + 8c \sum_{k \in I} 2^{(r-1)k} \mu(f \geq 2^k).$$

For k_0 large enough (only depending on r),

$$\sum_{k \in I} 2^{rk} \mu(f \geq 2^{k+1}) \leq 2^{2q+1} (c\varepsilon)^{q/2} \, \|\nabla f\|_q^q + \frac{2}{\varepsilon} \, \mathrm{W}_2^2(\nu, \mu) + 16c\, 2^{(r-1)k_0} \mu(f \geq 2^{k_0}).$$

The last term on the right-hand of this inequality may then be handled by the weak-type estimate of Proposition 4.2, concluding the proof of Theorem 4.1 after optimization in $\varepsilon > 0$. □

In addition to the preceding statements, there is a version of Proposition 4.2 for $p = q = 1$. This result shares some similarities with the recent [4].

Proposition 4.3. *In the setting of Theorem 4.1, under the curvature condition $CD(0, \infty)$, there is a numerical constant $C > 0$ such that for any probability $d\nu = f d\mu$ with smooth density f with respect to μ,*

$$\sup_{u \geq C} u^2 \, \mu(f \geq u)^2 \leq C \, \|\nabla f\|_1 \, \mathrm{W}_1(\nu, \mu).$$

The proof of this proposition follows the standard pattern. For any $u > 0$ and $t > 0$,

$$\mu(f \geq 2u) \leq \frac{\sqrt{2t}}{u} \, \|\nabla f\|_1 + \mu(P_t f \geq u).$$

Now, by (2.1), $\sqrt{2t}\, P_t(\mathbb{1}_F) f$ is 1-Lipschitz (where as usual $F = \{P_t f \geq u\}$) so that, for the W_1 metric,

$$\int_M P_t(\mathbb{1}_F) f \, d\mu \leq \int_M P_t(\mathbb{1}_F) d\mu + \frac{1}{\sqrt{2t}} \, \mathrm{W}_1(\nu, \mu) = \mu(F) + \frac{1}{\sqrt{2t}} \, \mathrm{W}_1(\nu, \mu).$$

Hence

$$\mu(P_t f \geq u) \leq \frac{1}{u} \int_M \mathbb{1}_F P_t f d\mu = \frac{1}{u} \int_M P_t(\mathbb{1}_F) f d\mu \leq \frac{1}{u} \, \mu(P_t f \geq u) + \frac{1}{\sqrt{2t}\, u} \, \mathrm{W}_1(\nu, \mu)$$

and for $u \geq 2$,

$$\mu(P_t f \geq u) \leq \frac{2}{\sqrt{2t}\, u} \, \mathrm{W}_1(\nu, \mu).$$

Therefore

$$\mu(f \geq 2u) \leq \frac{\sqrt{2t}}{u} \, \|\nabla f\|_1 + \frac{2}{\sqrt{2t}\, u} \, \mathrm{W}_1(\nu, \mu)$$

from which the claim follows after optimization in $t > 0$.

Proposition 4.3 has some interesting consequences to isoperimetric-type inequalities. Indeed, applied to f smoothly approaching $\frac{\mathbb{1}_A}{\mu(A)}$ for say a closed set A in M, it yields that whenever $\mu(A) \le \frac{1}{C}$,

$$\frac{1}{C^2\,\mathrm{W}_1(A)} \le \mu^+(A) \tag{4.3}$$

where

$$\mu^+(A) = \liminf_{\delta\to 0} \frac{1}{\delta}\left[\mu(A_\delta) - \mu(A)\right]$$

is the Minkowski content (surface measure) of A and, to ease the notation, $\mathrm{W}_1(A) = \mathrm{W}_1(\nu, \mu)$ where $d\nu = f d\mu = \frac{\mathbb{1}_A}{\mu(A)} d\mu$. The inequality (4.3) may be seen as a local version of the links between concentration and isoperimetric-type inequalities under non-negative curvature put forward by E. Milman in [10, 11]. Indeed, concentration at some scale amounts to an upper bound on $\mathrm{W}_1(A)$ in terms of a function of $\frac{1}{\mu(A)}$ which in turn, by (4.3), ensures that for sets with small measure the surface measure $\mu^+(A)$ is (uniformly) bounded away from zero. Due to the concavity of the isoperimetric profile under non-negative curvature, the latter always yields a comparison to the exponential isoperimetric profile (see [10, 11] for details).

Acknowledgement: I am grateful to E. Cinti and F. Otto for informing me about their paper [6] and for useful exchanges on the topic of this work.

References

[1] D. Bakry, I. Gentil, M. Ledoux. *Analysis and geometry of Markov diffusion operators.* Grundlehren der mathematischen Wissenschaften 348. Springer (2014).

[2] D. Bakry, I. Gentil, M. Ledoux. On Harnack inequalities and optimal transportation *Ann. Scuola Norm. Sup. Pisa,* to appear (2015).

[3] D. Bakry, Z. Qian. Some new results on eigenvectors via dimension, diameter, and Ricci curvature. *Adv. Math.* 155, 98–153 (2000).

[4] V. I. Bogachev, F.-Y. Wang, A. V. Shaposhnikov. Estimates of the Kantorovich norm on manifolds (2015).

[5] R. Choksi, S. Conti, R. Kohn, F. Otto. Ground state energy scaling laws during the onset and destruction of the intermediate state in a type I superconductor. *Comm. Pure Appl. Math.* 6, 595–626 (2008).

[6] E. Cinti, F. Otto. Interpolation inequalities in pattern formation (2014).

[7] E. B. Davies. *Heat kernels and spectral theory.* Cambridge Tracts in Mathematics 92. Cambridge (1989).

[8] M. Ledoux. On improved Sobolev embedding theorems. *Math. Res. Letters* 10, 659–669 (2003).

[9] P. Li, S.-T. Yau. On the parabolic kernel of the Schrödinger operator. *Acta Math.* 156, 153–201 (1986).

[10] E. Milman. On the role of convexity in isoperimetry, spectral gap and concentration. *Invent. Math.* 177, 1–43 (2009).

[11] E. Milman. Isoperimetric and concentration inequalities: equivalence under curvature lower bound. *Duke Math. J.* 154, 207–239 (2010).

[12] F. Otto, C. Villani. Generalization of an inequality by Talagrand, and links with the logarithmic Sobolev inequality. *J. Funct. Anal.* 173, 361–400 (2000).

[13] C. Villani. *Optimal transport. Old and new.* Grundlehren der mathematischen Wissenschaften 338. Springer (2009).

[14] F.-Y. Wang. Logarithmic Sobolev inequalities on noncompact Riemannian manifolds. *Probab. Theory Related Fields* 109, 417–424 (1997).

[15] F.-Y. Wang. *Functional inequalities, Markov properties and spectral theory.* Science Press (2005).