



From concentration to logarithmic Sobolev and Poincaré inequalities

Nathael Gozlan ^{*}, Cyril Roberto ¹, Paul-Marie Samson

*Université Paris Est Marne la Vallée, Laboratoire d'Analyse et de Mathématiques Appliquées (UMR CNRS 8050),
5 bd Descartes, 77454 Marne la Vallée Cedex 2, France*

Received 2 July 2010; accepted 17 November 2010

Available online 26 November 2010

Communicated by C. Villani

Abstract

We give a new proof of the fact that Gaussian concentration implies the logarithmic Sobolev inequality when the curvature is bounded from below, and also that exponential concentration implies Poincaré inequality under null curvature condition. Our proof holds on non-smooth structures, such as length spaces, and provides a universal control of the constants. We also give a new proof of the equivalence between dimension free Gaussian concentration and Talagrand's transport inequality.

© 2010 Elsevier Inc. All rights reserved.

Keywords: Concentration of measure; Transport inequalities; Poincaré inequalities; Logarithmic-Sobolev inequalities; Ricci curvature; Length spaces

Contents

1. Introduction	1492
1.1. Inequalities related to Gaussian concentration	1493
1.2. From concentration to functional inequalities	1494
1.3. Lott–Villani–Sturm curvature	1495
1.4. Main results	1497

^{*} Corresponding author.

E-mail addresses: nathael.gozlan@univ-mlv.fr (N. Gozlan), cyril.roberto@univ-mlv.fr (C. Roberto), paul-marie.samson@univ-mlv.fr (P.-M. Samson).

¹ The author was supported by the European Research Council through the “Advanced Grant” PTRELSS 228032.

1.5. Method of proof	1498
2. Concentration, inf-convolutions and non-tight transport inequalities	1499
2.1. Transport inequalities and their dual forms	1500
2.2. From concentration to non-tight transport inequalities	1503
2.3. A characterization of dimension free Gaussian concentration	1507
3. Log-Sobolev inequality: proof of Theorem 1.13	1508
4. Poincaré inequality: proof of Theorem 1.14	1513
4.1. Non-tight Poincaré inequality	1515
4.2. Weak Poincaré inequality	1517
5. Extensions	1518
Acknowledgments	1519
Appendix A. Technical results	1519
References	1520

This paper deals with the concentration of measure phenomenon and coercive inequalities, namely: transport inequalities, Poincaré and logarithmic Sobolev inequalities, isoperimetry. These notions are briefly introduced in the next section. We refer respectively to the books and surveys [29,27,5,35,38,23,49,2,45,20] for a more general introduction and a complete list of references on these topics.

It is well known that if a probability measure μ on, say, (X, d) a smooth Riemannian manifold, satisfies the logarithmic Sobolev inequality with constant C (see (1.3)), then the following Gaussian concentration property holds: for all subset A with $\mu(A) \geq 1/2$,

$$\mu(A^r) \geq 1 - Me^{-ar^2}, \quad \forall r \geq 0,$$

where $A^r = \{x \in X; d(x, A) \leq r\}$, with $M = 1$ and $a = 1/C$. Conversely, in a smooth Riemannian framework, under some curvature condition, Wang has shown in [46] that the above Gaussian concentration property implies a logarithmic Sobolev inequality. This result has been improved by Milman in a series of papers [34–36], showing that the logarithmic Sobolev constant C only depends on the concentration constants and on the lower bound on the Ricci curvature.

In the present paper, one of the main contribution is to extend Milman’s result to non-smooth structures, such as length spaces (see Theorem 1.13). The curvature condition of the length space is defined in the sense of Lott–Villani–Sturm (see Definition 1.9). In the same spirit, we show that the exponential concentration property implies Poincaré inequality when the curvature of the length space is bounded below by 0 (see Theorem 1.14).

The main ingredient in the proof of these new results is a characterization of the concentration property in terms of non-tight transport inequality (see Section 2). A byproduct of this characterization (see Corollary 2.23) is a new simple proof of the equivalence between dimension free Gaussian concentration and Talagrand’s transport inequality first established by the first named author in [18].

1. Introduction

In this section, we introduce the different inequalities related to concentration of measure and the notion of curvature in the sense of Lott–Villani–Sturm. Then we state our main results and we outline the proof.

1.1. Inequalities related to Gaussian concentration

In the sequel, (\mathcal{X}, d) is a polish space. A probability measure μ on \mathcal{X} enjoys the Gaussian concentration inequality if there are two positive constants M and a such that for all $A \subset \mathcal{X}$ with $\mu(A) \geq 1/2$, the following inequality holds

$$\mu(A^r) \geq 1 - Me^{-ar^2}, \quad \forall r \geq 0, \tag{1.1}$$

where $A^r = \{x \in X; d(x, A) \leq r\}$.

Gaussian concentration can be seen as a weak version of the Gaussian isoperimetric inequality. Namely, one says that μ verifies the Gaussian isoperimetric inequality with a positive constant C (**G.Isop**(C)) if, for all $A \subset \mathcal{X}$,

$$C\mu^+(A) \geq \Phi' \circ \Phi^{-1}(\mu(A)), \tag{1.2}$$

where

$$\mu^+(A) = \liminf_{r \rightarrow 0} \frac{\mu(A^r \setminus A)}{r},$$

and where Φ denotes the cumulative distribution function of the standard Gaussian measure on \mathbb{R} :

$$\Phi(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^t e^{-x^2/2} dx, \quad \forall t \in \mathbb{R}.$$

It can be shown (see e.g. [29, Proposition 2.1]) that μ verifies (1.2), if and only if μ verifies the following concentration property: for all $A \subset \mathcal{X}$,

$$\mu(A^r) \geq \Phi(\Phi^{-1}(\mu(A)) + r/C), \quad \forall r \geq 0.$$

It is not difficult to check that if $\mu(A) \geq 1/2$, then $\Phi(\Phi^{-1}(\mu(A)) + r/C) \geq 1 - e^{-r^2/(2C^2)}$, $r \geq 0$, and so μ verifies the Gaussian concentration property with $M = 1$ and $a = 1/(2C^2)$.

There are two other important functional inequalities giving Gaussian concentration: the logarithmic Sobolev inequality and Talagrand’s transport inequality. One says that μ verifies the logarithmic Sobolev inequality (a notion introduced by Gross [22], see also [42]) with the positive constant C (**LSI**(C)), if

$$\text{Ent}_\mu(f^2) \leq C \int |\nabla^- f|^2 d\mu, \tag{1.3}$$

for all bounded Lipschitz function f , where $|\nabla^- f|$ is defined by

$$|\nabla^- f|(x) = \limsup_{y \rightarrow x} \frac{[f(y) - f(x)]_-}{d(x, y)}, \quad \text{with } [A]_- = \max(0, -A),$$

when x is not isolated in \mathcal{X} and 0 otherwise. If μ verifies **LSI**(C), then by Herbst’s argument (see e.g. [2, Section 7.4.1]), it verifies the Gaussian concentration property (1.1) with $M = 1$ and $a = 1/C$.

On the other hand, one says that μ verifies the transport inequality **T**₂(C) with some positive constant C if

$$\mathcal{T}_2(\nu, \mu) \leq CH(\nu|\mu) \tag{1.4}$$

holds for all probability measure $\nu \in \mathcal{P}(\mathcal{X})$ (the set of all the Borel probability measures on \mathcal{X}), where $\mathcal{T}_2(\nu, \mu)$ denotes the quadratic optimal transport cost between ν and μ and is defined by

$$\mathcal{T}_2(\nu, \mu) = \inf_{\pi} \iint d^2(x, y) d\pi(x, y),$$

where the infimum runs over all probability measures π on $\mathcal{X} \times \mathcal{X}$ having ν and μ as marginal distributions. The quantity $H(\nu|\mu)$ is the relative entropy of ν with respect to μ :

$$H(\nu|\mu) = \begin{cases} \int \log \frac{d\nu}{d\mu} d\nu & \text{if } \nu \ll \mu, \\ +\infty & \text{otherwise.} \end{cases}$$

If μ satisfies **T**₂(C), then by Marton’s argument [33], it verifies (1.1), with $a = u/C$, $u \in (0, 1)$ and some positive constant M depending only on C and u .

There is a natural hierarchy between these inequalities: when (\mathcal{X}, d) is, for example, a complete Riemannian manifold (in this case, $\nabla^- f$ is the usual gradient), then

$$\mathbf{G.Isop}(\sqrt{C/2}) \Rightarrow \mathbf{LSI}(C) \Rightarrow \mathbf{T}_2(C) \Rightarrow (1.1). \tag{1.5}$$

The first implication is due to Ledoux (see [26,28]), the second one to Otto and Villani [37] (see also [11], and [31] or [21] for an extension on metric spaces) and the last one follows from Marton’s argument [33], as already mentioned.

1.2. From concentration to functional inequalities

In recent years, different authors have shown that this hierarchy can be reversed under the curvature condition,

$$\text{Ric} + \text{Hess } V \geq K, \tag{1.6}$$

where $d\mu(x) = e^{-V(x)} dx$ and $K \leq 0$ (for $K > 0$, it is known from [3] that the Gaussian isoperimetric inequality **G.Isop**($1/\sqrt{K}$) holds). Let us recall some of these contributions.

When $|K|/2a < 1$, it was first shown by Wang in [46] that the condition

$$I = \int e^{ad^2(x,x_0)} d\mu(x) < +\infty, \tag{1.7}$$

for some (and thus all) $x_o \in \mathcal{X}$ was enough to ensure that $\mathbf{LSI}(C)$ holds for some constant C . This constant C depends on a , K and I .

It can be deduced from Bakry and Ledoux’s paper [3], that if the probability measure μ verifies $\mathbf{LSI}(C)$ with a constant C such that $|K|C/2 < 1$, then μ verifies the Gaussian isoperimetric inequality with some other constant \tilde{C} depending only on C and K .

In [37], Otto and Villani proved that $\mathbf{T}_2(C)$ implies $\mathbf{LSI}(\tilde{C})$ for some \tilde{C} depending on K and C as soon as $|K|C/2 < 1$.

At this step, it is worth noting an important difference between these three results. In Bakry and Ledoux, or Otto and Villani results, the relation between constants C and \tilde{C} is universal: in both cases, the constant \tilde{C} depends only on K and C . In particular, the dimension of \mathcal{X} does not appear in the expression of \tilde{C} . On the other hand, the constant \tilde{C} in Wang’s result depends on the integral I . Since I always depends on the dimension of the manifold \mathcal{X} , Wang’s result is dimensional.

After these pioneer works, many developments appear to unify and to generalize these observations [7,9,10,6]. We refer to [34] for a detailed bibliography.

In a series of papers [34–36], E. Milman has recently obtained the most general results in this direction. Under the curvature condition (1.6), Milman has shown with a great generality that concentration inequalities imply isoperimetric inequalities. The most remarkable feature of his work is the purely adimensional character of the relations between constants. Let us give a direct corollary of Milman’s study in the context of Gaussian inequalities. According to Milman’s results, if (1.6) holds with $K \leq 0$, and if μ verifies the Gaussian concentration inequality (1.1) with constants a and M such that $|K|/(2a) < 1$, then μ verifies the inequality $\mathbf{G.Isop}(\sqrt{C/a})$, where C is a constant depending solely on a , M and K . In particular μ verifies $\mathbf{LSI}(2C/a)$, and contrary to what happen in Wang’s theorem, the constant appearing in the logarithmic Sobolev inequality is not affected by the dimension of \mathcal{X} .

Milman’s proof uses rather difficult tools of Riemannian geometry. Recently, Ledoux [30], gave a simplified approach to some of Milman’s results relying on semigroup tools and Γ_2 calculus. The purpose of this article is to propose a first step in the challenging problem of extending Milman’s equivalence between concentration and functional inequalities in the framework of metric measured spaces with curvature, in the sense of Lott–Villani–Sturm, bounded from below (see [32,43,44]). In this non-smooth context, the tools used by Milman and Ledoux are no longer available.

In the next subsection, we present the notions of length spaces and curvature.

1.3. Lott–Villani–Sturm curvature

In all what follows, (\mathcal{X}, d) is a complete separable and locally compact metric space. We will further assume that (\mathcal{X}, d) is a length space. This means that the distance between two points equals the infimum of the lengths of the curves joining these points: for all $x, y \in \mathcal{X}$,

$$d(x, y) = \inf \{ \ell(\gamma); \gamma : [0, 1] \rightarrow \mathcal{X}, \text{ continuous, } \gamma(0) = x, \gamma(1) = y \},$$

where

$$\ell(\gamma) = \sup_{N \geq 1} \sup_{0=t_0 < t_1 < \dots < t_N=1} \left\{ \sum_{i=1}^N d(\gamma(t_{i-1}), \gamma(t_i)) \right\}.$$

Let $\mathcal{W}_2 = \sqrt{\mathcal{T}_2}$ denote the Wasserstein distance on the space $\mathcal{P}_2(\mathcal{X})$ of Borel probability measures ν on \mathcal{X} with finite second moment: $\int d(x_o, x)^2 d\nu(x) < +\infty$, for some (and thus all) $x_o \in \mathcal{X}$. The metric space $(\mathcal{P}_2(\mathcal{X}), \mathcal{W}_2)$ is canonically associated to the original metric space (\mathcal{X}, d) . Following Lott and Villani [32], we define \mathcal{DC}_∞ to be the set of continuous convex functions $U : [0, \infty) \rightarrow \mathbb{R}$, with $U(0) = 0$ and such that $\lambda \mapsto e^\lambda U(e^{-\lambda})$ is convex on $(-\infty, \infty)$. For any $U \in \mathcal{DC}_\infty$, let

$$U'(\infty) = \lim_{x \rightarrow \infty} \frac{U(x)}{x} \in \mathbb{R} \cup \{\infty\}.$$

According to our reference probability measure μ , define the function $U_\mu : \mathcal{P}_2(\mathcal{X}) \rightarrow \mathbb{R} \cup \{-\infty\}$ by

$$U_\mu(\nu) = \int U(f) d\mu + U'(\infty)\nu_{\text{sing}}(\mathcal{X}),$$

where $\nu = f\mu + \nu_{\text{sing}}$ is the Lebesgue decomposition of ν with respect to μ into an absolutely continuous part $f\mu$ and a singular part ν_{sing} . For any $U \in \mathcal{DC}_\infty$, let $p(x) = xU'_+(x) - U(x)$ (U'_+ stands for the right derivative of U), and for $K \in \mathbb{R}$ we define

$$\kappa(U) = \inf_{x>0} K \frac{p(x)}{x} \in \mathbb{R} \cup \{-\infty\}. \tag{1.8}$$

Definition 1.9. The space (\mathcal{X}, d, μ) has ∞ -Ricci curvature bounded below by K , $K \in \mathbb{R}$, if for all $\nu_1, \nu_2 \in \mathcal{P}_2(\mathcal{X})$ whose supports are included in the support of μ , there exists a Wasserstein geodesic $\{\nu_t\}_{t \in [0,1]}$ from ν_0 to ν_1 (this means that $\mathcal{W}_2(\nu_s, \nu_t) = |s - t|\mathcal{W}_2(\nu_0, \nu_1)$ for all $s, t \in [0, 1]$) such that for all $U \in \mathcal{DC}_\infty$ and all $t \in [0, 1]$,

$$U_\mu(\nu_t) \leq tU_\mu(\nu_1) + (1 - t)U_\mu(\nu_0) - \frac{1}{2}\kappa(U)t(1 - t)\mathcal{W}_2^2(\nu_0, \nu_1). \tag{1.10}$$

As explained in [32, Theorem 7.3.b], this definition is exactly equivalent to the usual curvature condition (1.6) when \mathcal{X} is a smooth Riemannian manifold (M, g) . A straightforward consequence of this curvature condition is the inequality (1.12) below called *HWI inequality*. This inequality is at the heart of the proofs of our main results.

Proposition 1.11 (*HWI inequality*). (See Proposition 3.36 in [32].) Let $U \in \mathcal{DC}_\infty \cap \mathcal{C}^2$; if (\mathcal{X}, d, μ) has ∞ -Ricci curvature bounded below by K , then for any probability measure ν absolutely continuous with respect to μ , such that $f = d\nu/d\mu$ is a positive Lipschitz function on \mathcal{X} with $\int U(f) d\mu < \infty$, it holds

$$U_\mu(\nu) \leq U(1) + \sqrt{I_{\mu,U}(\nu)}\sqrt{\mathcal{T}_2(\nu, \mu)} - \frac{\kappa(U)}{2}\mathcal{T}_2(\nu, \mu), \tag{1.12}$$

where $I_{\mu,U}$ is the generalized Fisher information associated to U ,

$$I_{\mu,U}(\nu) = \int fU''(f)^2|\nabla^- f|^2 d\mu.$$

The sketch of the proof of this result is to choose $\nu_0 = \nu$, $\nu_1 = \mu$ and let t go to 0 in (1.10) (see [32]).

1.4. Main results

Now we are in position to give our results concerning logarithmic Sobolev and Poincaré inequalities.

Theorem 1.13 (Logarithmic Sobolev inequality). *Suppose (\mathcal{X}, d, μ) has ∞ -Ricci curvature bounded below by $K \leq 0$ and that μ verifies the Gaussian concentration property (1.1) with positive constants a and M . If the constants a , M and K satisfy the relation*

$$\frac{|K|}{2a} < \tau(M) := \frac{\log(2)}{(2\sqrt{M} + \sqrt{\log(2)})^2},$$

then μ verifies the logarithmic Sobolev inequality **LSI**(C) for some C depending only on K , a and M .

In particular, when $K = 0$, one has, for any bounded Lipschitz function $f : \mathcal{X} \rightarrow \mathbb{R}$,

$$\text{Ent}_\mu(f^2) \leq \frac{DM}{a} \int |\nabla^- f|^2 d\mu,$$

where D is some universal constant.

Let us make a few comments on the above theorem. For $K = 0$, the result is as good as possible: the logarithmic Sobolev inequality is (up to numerical factors) equivalent to the Gaussian concentration. When $K < 0$, it is known, in a Riemannian setting, that the Gaussian concentration implies the logarithmic Sobolev inequality, only when $|K|/(2a) < 1$. Here, we recover the qualitative condition that the concentration constant a has to be larger than the curvature constant K , but with a wrong ratio. Moreover this ratio τ depends on the constant M . Though perfectible, this result is the first extension of Wang’s theorem available on this very general non-smooth framework.

Our second main result deals with Poincaré inequality.

Theorem 1.14 (Poincaré inequality). *Suppose (\mathcal{X}, d, μ) has ∞ -Ricci curvature bounded below by 0 and that μ verifies the following exponential concentration property: for all $A \subset \mathcal{X}$, with $\mu(A) \geq 1/2$,*

$$\mu(A^r) \geq 1 - Me^{-ar}, \quad \forall r \geq 0$$

with $M, a > 0$. Then, there exists a constant D that depends only on M such that for any bounded Lipschitz function $f : \mathcal{X} \rightarrow \mathbb{R}$, it holds

$$\text{Var}_\mu(f) \leq \frac{D}{a^2} \int |\nabla^- f|^2 d\mu. \tag{1.15}$$

Since Poincaré inequality gives back exponential concentration, the conclusion of this theorem is quite satisfactory. In [35], Milman has obtained the following striking result on a Riemannian manifold: when $K = 0$, if μ verifies any non-trivial concentration property (not necessarily exponential) then μ verifies Poincaré inequality (and even Cheeger linear isoperimetric inequality). The proof of this extremely powerful result uses as a main ingredient the fact that, in this situation, the isoperimetric profile J_μ of μ , defined by $J_\mu(t) = \inf\{\mu^+(A); \mu(A) = t\}$, $t \in [0, 1]$, is a concave function of t . The difficult proof of the concavity of J_μ uses purely Riemannian geometric tools, and we do not know if it is reasonable to ask for an extension on metric spaces. This is far beyond the scope of the present paper. The question to know if the conclusion of the above mentioned result by Milman holds on metric spaces is open.

Let us end the presentation of our results by saying that during the preparation of this work, we have made the drastic choice to restrict ourselves only to these two functional inequalities. Many Sobolev type inequalities could be covered by our methods: F -Sobolev inequalities, Beckner–Latała–Oleszkiewicz inequalities, super-Poincaré or Nash inequalities. Some results in this direction, without proof, are given in the last section. The reason of this restriction is that we wanted to put in light in the most transparent way the general methodology and the different ingredients entering the proofs.

1.5. Method of proof

Here we briefly describe, in the Gaussian case, the method on which rely our proofs. As we said above, the starting point is the HWI inequality which, in the case where $U(x) = x \log(x)$, takes the form

$$\begin{aligned} H(v|\mu) &\leq \sqrt{\mathcal{T}_2(v, \mu)}\sqrt{I_{\mu,U}(v)} - \frac{K}{2}\mathcal{T}_2(v, \mu), \quad v \in \mathcal{P}(\mathcal{X}) \\ &\leq \frac{\lambda - K}{2}\mathcal{T}_2(v, \mu) + \frac{1}{2\lambda}I_{\mu,U}(v), \end{aligned} \tag{1.16}$$

for all $\lambda > 0$, with $I_{\mu,U}(v) = \int \frac{|\nabla^- f|^2}{f} d\mu$ when $v = f\mu$ with f Lipschitz.

In recent years, several authors have used this inequality (or the corresponding semiconvexity property given by (1.10)) to derive functional inequalities see [14,1,15,7]. It was first noticed by Otto and Villani in [37] (see also [32] for the metric space case), that (1.16) gives back the Bakry–Emery condition when $K > 0$. Namely, taking $\lambda = K$, the right-hand side of (1.16) is in this case smaller than $\frac{1}{2K}I_{\mu,U}(v)$, and so **LSI**($2/K$) holds.

Another application of (1.16) was given in [37] in the case $K \leq 0$. If μ verifies $\mathbf{T}_2(C)$ for some C such that $|K|C/2 < 1$, then plugging the inequality $\mathcal{T}_2(v, \mu) \leq CH(v|\mu)$ into (1.16) leads (for a convenient choice of λ) to the following logarithmic Sobolev inequality

$$H(v|\mu) \leq \frac{C}{(1 + KC/2)^2}I_{\mu,U}(v).$$

This last argument suggests that the HWI inequality is an efficient tool in order to reverse the hierarchy (1.5), when $K \leq 0$. To show that Gaussian concentration implies the logarithmic Sobolev inequality, our idea is to plug into (1.16) a transport inequality weaker than \mathbf{T}_2 . Let us say that μ verifies the (non-tight) transport inequality $\mathbf{T}_2(c_1, c_2)$ for some $c_1, c_2 \geq 0$, if

$$\mathcal{T}_2(v, \mu) \leq c_1 H(v|\mu) + c_2, \tag{1.17}$$

for all $v \in \mathcal{P}(\mathcal{X})$. One of the key ingredients of the proof of Theorem 1.13, is that it is possible to encode the Gaussian concentration property (1.1) by an inequality $\mathbf{T}_2(c_1, c_2)$ with a universal link between the constants a, M, c_1 and c_2 (see Corollary 2.20). When plugging (1.17) into (1.16), we naturally arrive to a defective logarithmic Sobolev inequality of the form

$$H(v|\mu) \leq d_1 I_{\mu,U}(v) + d_2, \tag{1.18}$$

for all v with a Lipschitz density with respect to μ . The rest of the proof consists in tightening this inequality. For that purpose, we use a result by Wang showing that, when d_2 is small enough, (1.18) implies $\mathbf{LSI}(C)$, for some C depending only on d_1 and d_2 .

Non-tight transport inequalities like (1.17) have their own interest. In Section 2, we establish a general link between concentration inequalities and non-tight transport inequalities involving functionals U_μ in the right-hand side. We also give a dual formulation of them, extending a celebrated theorem by Bobkov and Götze [12]. Moreover, this analysis enables us to recover, in a completely analytic way, the fact that dimension free Gaussian concentration is equivalent to \mathbf{T}_2 , a result obtained by the first named author in [18] using large deviation techniques.

To conclude this introduction let us mention two closely related papers.

In [7], Barthe and Kolesnikov have followed a similar scheme of proof to go from concentration to functional inequalities. They have shown, that in a very general framework, integrability conditions like (1.7) were sufficient to obtain Sobolev type and isoperimetric inequalities. So the main difference between their paper and ours, is that our results do not involve dimensional quantity like (1.7).

In [36], E. Milman has proved that concentration inequalities can be encoded by transport inequalities involving the relative entropy and the L_1 -Wasserstein distance:

$$\mathcal{W}_1(v, \mu) = \inf_{\pi} \iint d(x, y) d\pi,$$

where the infimum runs over all the couplings of v and μ . The main difference with our work is that we characterize concentration in terms of the quadratic transport cost \mathcal{T}_2 . Replacing \mathcal{W}_1 by \mathcal{T}_2 requires rather subtle techniques developed in Section 2.

2. Concentration, inf-convolutions and non-tight transport inequalities

In this section we deal with the links between concentration and non-tight transport inequalities. Our first task is to establish a dual version of the latter, in the spirit of Bobkov and Götze dual theorem [12]. Then we express the concentration property in term of inf-convolution operator. Finally, we use this approach to recover the characterization of the dimension free Gaussian concentration of [18].

Let us introduce some general notation. In all the section, (\mathcal{X}, d) is a polish space. A probability measure μ is said to satisfy a concentration inequality if there is a measurable function $\alpha : [0, \infty) \rightarrow [0, \infty)$ such that for all $A \subset \mathcal{X}$ with $\mu(A) \geq 1/2$,

$$\mu(A^r) \geq 1 - \alpha(r), \quad \forall r > 0, \tag{2.1}$$

where $A^r = \{x \in X; d(x, A) \leq r\}$.

2.1. Transport inequalities and their dual forms

In all what follows, we let $U : [0, \infty) \rightarrow \mathbb{R}$ be a lower semicontinuous strictly convex function which we moreover assume to be superlinear (i.e. $U(x)/x \rightarrow \infty$ when x goes to ∞). We also impose that $U(1) \geq 0$ so that, by Jensen’s inequality, $U_\mu(v) \geq 0$ for all $v \in \mathcal{P}(\mathcal{X})$. We will mainly be concerned with the two following particular cases: $U(x) = x \log(x)$, for which $U_\mu(v) = H(v|\mu)$ is the relative entropy of v with respect to μ , and $U(x) = x \log^2(e + x)$.

One of our main object of interest is the following *non-tight* transport inequality.

Definition 2.2 (*Non-tight transport inequality*). One says that μ verifies the transport inequality $\mathbf{T}_2U(c_1, c_2)$, $c_1, c_2 \geq 0$, if

$$\mathcal{T}_2(v, \mu) \leq c_1U_\mu(v) + c_2, \quad \forall v \in \mathcal{P}(\mathcal{X}).$$

When $U(x) = x \log(x)$, we denote this inequality by $\mathbf{T}_2(c_1, c_2)$, and when $c_2 = 0$, by $\mathbf{T}_2(c_1)$.

It is well known that $\mathbf{T}_2(c_1)$ implies the Gaussian concentration property applying the Marton’s argument. Proposition 2.3 below shows that more generally $\mathbf{T}_2U(c_1, c_2)$ also implies concentration properties. Moreover we will show in Section 2.2 that for some specific choices of the function U , $\mathbf{T}_2U(c_1, c_2)$ is actually equivalent to the Gaussian or the exponential concentration property.

Proposition 2.3. *If μ verifies the inequality $\mathbf{T}_2U(c_1, c_2)$, then μ verifies the following concentration inequality: for all $A \subset \mathcal{X}$ with $\mu(A) \geq 1/2$,*

$$\mu(A^c) \geq 1 - \varphi_U^{-1}\left(\frac{1}{c_1}(r - r_o)^2\right), \quad \forall r \geq r_o + \sqrt{c_1[U'_+(0) + U(0)]_+},$$

where $\varphi_U(t) = tU(1/t) + (1 - t)U(0)$, $t > 0$, and $r_o = 2\sqrt{c_2} + c_1\sqrt{(U(0) + U(2))/2}$.

Remark 2.4. Observe that the function $\varphi_U(t) = tU(1/t) + (1 - t)U(0)$, $t > 0$, is strictly decreasing on $(0, \infty)$ with values in $(U'_+(0) + U(0), \infty)$. Its inverse function φ_U^{-1} is thus strictly decreasing (and well defined) on $(U'_+(0) + U(0), \infty)$ and $\varphi_U^{-1}(r) \rightarrow 0$ when $r \rightarrow \infty$ (all these facts are an immediate consequence of the strict convexity and the superlinearity of U).

For $U(x) = x \log(x)$, one has $\varphi_U^{-1}(t) = e^{-t}$, $t \in \mathbb{R}$. So we conclude that $\mathbf{T}_2(c_1, c_2)$ implies Gaussian concentration.

For $U(x) = x \log^2(e + x)$, one has $\varphi_U^{-1}(t) = \frac{e^{-\sqrt{t}}}{1 - e^{-\sqrt{t}}}$, $t \geq 1$. In this case, $\mathbf{T}_2U(c_1, c_2)$ implies exponential concentration.

Proof of Proposition 2.3. What follows is a straightforward adaptation of Marton’s argument. If $\mathbf{T}_2U(c_1, c_2)$ holds then the sub-additivity property of \mathcal{W}_2 implies that for all $v_1, v_2 \in \mathcal{P}_2(\mathcal{X})$,

$$\mathcal{W}_2(v_1, v_2) \leq \mathcal{W}_2(v_1, \mu) + \mathcal{W}_2(\mu, v_2) \leq \sqrt{c_1U_\mu(v_1)} + \sqrt{c_1U_\mu(v_2)} + 2\sqrt{c_2}.$$

Let A be a subset of \mathcal{X} . Choosing v_1 with density $\mathbb{1}_A/\mu(A)$ and v_2 with density $(1 - \mathbb{1}_A)/$

$(1 - \mu(A^r))$ with respect to μ , one has $r \leq \mathcal{W}_2(v_1, v_2)$, $U_\mu(v_1) = \varphi_U(\mu(A)) \leq \varphi_U(1/2) = (U(0) + U(2))/2$ (since φ_U is decreasing) and $U_\mu(v_2) = \varphi_U(\mu(B))$. Note that $(U(0) + U(2))/2 \geq U(1) \geq 0$. So, $r - r_o \leq \sqrt{c_1 \varphi_U(\mu(B))}$ and if r is large enough, we get the desired concentration inequality. \square

Our next aim is to give a dual formulation of the non-tight transport inequalities introduced above. Let \mathcal{B}_b be the space of all bounded measurable functions on \mathcal{X} and \mathcal{C}_b the space of bounded continuous functions on \mathcal{X} . For all functions $U : [0, \infty) \rightarrow \mathbb{R}$, define

$$\Lambda_\mu(h) = \sup_{\nu \in \mathcal{P}(\mathcal{X})} \left\{ \int h \, d\nu - U_\mu(\nu) \right\}, \quad \forall h \in \mathcal{B}_b. \tag{2.5}$$

If $f : \mathcal{X} \rightarrow \mathbb{R} \cup \{+\infty\}$ is a measurable function bounded from below, let us define the inf-convolution operators $(Q_\lambda)_{\lambda>0}$ as follows

$$Q_\lambda f(x) = \inf_{y \in \mathcal{X}} \left\{ f(y) + \frac{1}{\lambda} d^2(x, y) \right\}, \quad x \in \mathcal{X}.$$

For $\lambda = 1$ we denote $Q_1 f$ by Qf . The following result is a straightforward extension of Bobkov–Götze theorem [12], providing a dual formulation of transport inequalities involving the relative entropy.

Theorem 2.6. *A probability μ verifies $\mathbf{T}_2 U(c_1, c_2)$ if and only if*

$$\Lambda_\mu(Q_{c_1} f) \leq \int f \, d\mu + c_2/c_1, \quad \forall f \in \mathcal{C}_b. \tag{2.7}$$

Proof. According to Kantorovich’s theorem

$$\mathcal{T}_2(\nu, \mu) = \sup_{h \in \mathcal{C}_b} \left\{ \int Qh \, d\nu - \int h \, d\mu \right\}.$$

So, μ verifies the transport inequality if and only if for all $h \in \mathcal{C}_b$, it holds

$$\int (Qh)/c_1 \, d\nu - U_\mu(\nu) \leq \int h/c_1 \, d\mu + c_2/c_1, \quad \forall \nu \in \mathcal{P}(\mathcal{X}).$$

Optimizing over ν and according to (2.5), we arrive at the equivalent condition

$$\Lambda_\mu((Qh)/c_1) \leq \int h/c_1 \, d\mu + c_2/c_1, \quad \forall h \in \mathcal{C}_b.$$

Letting $f = h/c_1$ gives (2.7). \square

Define $U^*(t) = \sup_{s \geq 0} \{st - U(s)\}$, $t \in \mathbb{R}$, then

$$\Lambda_\mu(h) \leq \inf_{t \in \mathbb{R}} \int U^*(h + t) - t \, d\mu, \quad \forall h \in \mathcal{C}_b. \tag{2.8}$$

This easily follows from the fact that $\Lambda_\mu(h) = \Lambda_\mu(h + t) - t$ for all $t \in \mathbb{R}$, and from Young’s inequality: $xy \leq U(x) + U^*(y)$, $x \geq 0, y \in \mathbb{R}$. A direct consequence of (2.8) is that $\mathbf{T}_2U(c_1, c_2)$ holds as soon as for all $f \in \mathcal{C}_b$, there exists some $t_f \in \mathbb{R}$ such that

$$\int U^*(Q_{c_1} f + t_f) - t_f d\mu \leq \int f d\mu + c_2/c_1.$$

This will be a key point to show that concentration implies $\mathbf{T}_2U(c_1, c_2)$ (see Section 2.2 below).

For the sake of completeness, let us show that (2.8) is actually an equality. It will follow that U_μ is the dual function of Λ_μ .

Proposition 2.9. *The following duality formulas hold*

$$\Lambda_\mu(h) = \inf_{t \in \mathbb{R}} \int U^*(h + t) - t d\mu, \quad \forall h \in \mathcal{C}_b, \tag{2.10}$$

and for all $\nu \in \mathcal{P}(\mathcal{X})$,

$$\begin{aligned} U_\mu(\nu) &= \sup_{h \in \mathcal{C}_b} \left\{ \int h d\nu - \int U^*(h) d\mu \right\} \\ &= \sup_{h \in \mathcal{C}_b} \left\{ \int h d\nu - \Lambda_\mu(h) \right\}. \end{aligned} \tag{2.11}$$

For example, when $U(x) = x \log(x)$, $x \geq 0$, then $U^*(y) = e^{y-1}$ and we recover the well-known identity

$$\Lambda_\mu(h) = \inf_{t \in \mathbb{R}} \left\{ e^{t-1} \int e^h d\mu - t \right\} = \log \left(\int e^h d\mu \right).$$

Proof of Proposition 2.9. For the proof of (2.10), since (2.8) holds, it remains to show the reverse inequality. We can restrict to the case where ν is absolutely continuous with respect to μ . Let $g = d\nu/d\mu$. Fix $h \in \mathcal{C}_b$ and set $F(t) = \int U^*(h + t) - t d\mu, t \in \mathbb{R}$. Since U is superlinear, U^* is finite everywhere, and so is F . Furthermore, as U is strictly convex and lower semicontinuous, U^* is differentiable (see [40, Chapter 26]). Obviously $F(t) \geq -U(0) - t$ and therefore $F(t) \rightarrow +\infty$ as $t \rightarrow -\infty$. Since for all $t \geq 0, U(t) < \infty$ and $U^*(x)/x \geq t - U(t)/x$ it is easy to conclude that $U^*(x)/x \rightarrow +\infty$ as $x \rightarrow +\infty$. As a consequence we also have $F(t) \rightarrow +\infty$ as $t \rightarrow +\infty$. Therefore the infimum $\inf_{t \in \mathbb{R}} \int U^*(h + t) - t d\mu$ is reached at some point t_0 such that $F'(t_0) = 0$ or equivalently

$$\int (U^*)'(h + t_0) d\mu = 1.$$

Being a supremum of increasing functions, U^* is increasing and so $(U^*)'(h + t_0) \geq 0$. Let ν_0 denote the probability measure with density $g_0 = (U^*)'(h + t_0)$ with respect to μ : one has

$$\begin{aligned} \sup_{\nu \in \mathcal{P}(\mathcal{X})} \left\{ \int h \, d\nu - U_{\mu}(\nu) \right\} &\geq \int (hg_0 - U(g_0)) \, d\mu \\ &= \int U^*(h + t_0) - t_0 \, d\mu \\ &= \inf_{t \in \mathbb{R}} \int U^*(h + t) - t \, d\mu, \end{aligned}$$

where we used the fact that for all $x \in \mathbb{R}$

$$x(U^*)'(x) - U((U^*)'(x)) = U^*(x). \tag{2.12}$$

Let us sketch the proof of this identity in the case when U is of class \mathcal{C}^1 (but (2.12) is true without this assumption). For $x < U'(0^+)$, this follows from $U^*(x) = -U(0)$ and $(U^*)'(x) = 0$. For $x \geq U'(0^+)$, since U is superlinear, the supremum $U^*(x)$ is reached at some point $s \geq 0$ such that $U'(s) = x$; the equality (2.12) then follows since $(U^*)'(x) = U'^{-1}(x)$. The proof of (2.10) is complete.

The first equality in (2.11) is proved for instance in [32, Theorem B.2]. The second equality follows from (2.10) and an immediate optimization noticing that \mathcal{C}_b is stable under the translations $h \mapsto h + t$. \square

2.2. From concentration to non-tight transport inequalities

The purpose of the next proposition is to give a new formulation of the concentration inequality (2.1) in terms of deviation inequalities of inf-convolution operators.

Proposition 2.13. *A probability μ verifies the concentration property (2.1), if and only if, for all $\lambda > 0$, and all measurable functions $f : \mathcal{X} \rightarrow \mathbb{R} \cup \{+\infty\}$ bounded from below, it holds*

$$\mu(Q_{\lambda}f > \text{med}_{\mu}(f) + r) \leq \alpha(\sqrt{\lambda r}), \quad r \geq 0. \tag{2.14}$$

Proof. We first show that (2.14) implies (2.1). Take $A \subset \mathcal{X}$, with $\mu(A) \geq 1/2$, and consider the function f_A which is 0 on A and $+\infty$ otherwise. It is clear that 0 is a median for f_A , and that $Q_{\lambda}f_A(x) = d^2(x, A)/\lambda$. So applying (2.14) yields

$$\mu(\{x; d(x, A) > \sqrt{\lambda r}\}) \leq \alpha(\sqrt{\lambda r}), \quad r \geq 0,$$

which gives back (2.1).

Now we prove that (2.1) implies (2.14). Let $f : \mathcal{X} \rightarrow \mathbb{R} \cup \{+\infty\}$ be a measurable function bounded from below and define $A = \{x \in \mathcal{X}; f(x) \leq \text{med}_{\mu}(f)\}$. If $x \in A^c$, then there is some $y \in A$ such that $d(x, y) \leq r$. Consequently,

$$Q_{\lambda}f(x) \leq f(y) + d^2(x, y)/\lambda \leq \text{med}_{\mu}(f) + r^2/\lambda.$$

So $A^c \subset \{x \in \mathcal{X}; Q_{\lambda}f(x) \leq \text{med}_{\mu}(f) + r^2/\lambda\}$ and therefore

$$\mu(Q_\lambda f(x) > \text{med}_\mu(f) + r^2/\lambda) \leq \alpha(r), \quad r \geq 0,$$

which is (2.14). \square

Now our objective is to deduce transport inequalities from concentration. We need some preparation.

Lemma 2.15. *If $c_1, \lambda_1, \lambda_2$ are positive and such that $c_1 = \lambda_1 + \lambda_2$, then for any $f \in \mathcal{C}_b$,*

$$\Lambda_\mu(Q_{c_1} f) \leq \int f \, d\mu + \sup_{g \in \mathcal{B}_b} \Lambda_\mu(Q_{\lambda_1} g - \text{med}_\mu(g)) + \sup_{g \in \mathcal{B}_b} \left\{ \text{med}_\mu(Q_{\lambda_2} g) - \int g \, d\mu \right\},$$

where \mathcal{B}_b denotes the space of bounded and measurable functions on \mathcal{X} .

Proof. First, the following inequality holds $Q_{c_1} f \leq Q_{\lambda_1}(Q_{\lambda_2} f)$, for all $f \in \mathcal{C}_b$ (see e.g. [4, proof of Theorem 2.5(ii)]). Furthermore, it follows easily from its definition that Λ_μ is order preserving: $f_1 \leq f_2 \Rightarrow \Lambda_\mu(f_1) \leq \Lambda_\mu(f_2)$. So,

$$\begin{aligned} \Lambda_\mu(Q_{c_1} f) &\leq \Lambda_\mu(Q_{\lambda_1}(Q_{\lambda_2} f)) \\ &= \int f \, d\mu + \Lambda_\mu(Q_{\lambda_1}(Q_{\lambda_2} f) - \text{med}_\mu(Q_{\lambda_2} f)) + \text{med}_\mu(Q_{\lambda_2} f) - \int f \, d\mu, \end{aligned}$$

where the last equality follows from the fact that $\Lambda_\mu(h + r) = \Lambda_\mu(h) + r$, for all $h \in \mathcal{B}_b$ and $r \in \mathbb{R}$. Since f and $Q_{\lambda_2} f$ are bounded, the claim follows by taking supremums over $g \in \mathcal{B}_b$. \square

Now, we will use Proposition 2.13 to bound the two supremums in Lemma 2.15.

Lemma 2.16. *Assume that (2.1) holds for some function α . Then,*

$$\sup_{g \in \mathcal{B}_b} \left\{ \text{med}_\mu(Q_\lambda g) - \int g \, d\mu \right\} \leq \int_0^\infty \alpha(\sqrt{\lambda t}) \, dt, \quad \forall \lambda > 0.$$

Lemma 2.17. *Assume that (2.1) holds for some function α . Then, if $W : \mathbb{R} \rightarrow \mathbb{R}$ is an increasing function of class \mathcal{C}^1 bounded from below, it holds*

$$\sup_{g \in \mathcal{B}_b} \int W(Q_\lambda g - \text{med}_\mu(g)) \, d\mu \leq W(0) + \int_0^\infty W'(t)\alpha(\sqrt{\lambda t}) \, dt, \quad \forall \lambda > 0.$$

Proof of Lemma 2.16. Fix a bounded function $g : \mathcal{X} \rightarrow \mathbb{R}$ and assume without loss of generality that $\int_0^\infty \alpha(\sqrt{t}) \, dt < \infty$. Since $Q_\lambda g(x) \leq g(y) + \frac{1}{\lambda} d^2(x, y)$ for any $y \in \mathcal{X}$, it holds that

$$-g(y) \leq \inf_{x \in \mathcal{X}} \left\{ -Q_\lambda g(x) + \frac{1}{\lambda} d^2(x, y) \right\} = Q_\lambda(-Q_\lambda g)(y).$$

In particular,

$$-\int g \, d\mu \leq \int Q_\lambda(-Q_\lambda g) \, d\mu. \tag{2.18}$$

Set $h = -Q_\lambda g$. Then, integrating by part and using Proposition 2.13, one has

$$\begin{aligned} \int (Q_\lambda h - \text{med}_\mu(h)) \, d\mu &\leq \int (Q_\lambda h - \text{med}_\mu(h))_+ \, d\mu \\ &= \int_0^\infty \mu((Q_\lambda h - \text{med}_\mu(h))_+ > t) \, dt \\ &\leq \int_0^\infty \mu(Q_\lambda h - \text{med}_\mu(h) > t) \, dt \\ &\leq \int_0^\infty \alpha(\sqrt{\lambda t}) \, dt. \end{aligned}$$

The expected result follows from (2.18) and the fact that $-\text{med}_\mu(h) = \text{med}_\mu(Q_\lambda g)$. \square

Proof of Lemma 2.17. Integrating by part and using Proposition 2.13 yields

$$\begin{aligned} \int W(Q_\lambda g - \text{med}_\mu(g)) \, d\mu &= W(-\infty) + \int_{-\infty}^{+\infty} W'(s)\mu(Q_\lambda g - \text{med}_\mu(g) \geq s) \, ds \\ &\leq W(-\infty) + \int_{-\infty}^0 W'(s) \, ds + \int_0^{+\infty} W'(s)\alpha(\sqrt{\lambda s}) \, ds, \end{aligned}$$

which proves the claim. \square

We are now in position to prove that concentration implies non-tight transport inequalities.

Theorem 2.19. Assume in addition that U is of class \mathcal{C}^1 on $(0, \infty)$. If μ verifies (2.1) for some function α and $\int_0^{+\infty} (U^*)'(t)\alpha(\sqrt{\lambda_1 t}) \, dt < \infty$ for some $0 < \lambda_1$, then μ verifies the inequality $\mathbf{T}_2 U(c_1, c_2)$ for some constants $c_1, c_2 > 0$. More precisely, for all $\lambda_2 > 0$, one can take

$$c_1 = \lambda_1 + \lambda_2$$

and

$$c_2 = (\lambda_1 + \lambda_2) \left(U^*(0) + \int_0^{+\infty} (U^*)'(t)\alpha(\sqrt{\lambda_1 t}) \, dt + \int_0^{+\infty} \alpha(\sqrt{\lambda_2 t}) \, dt \right).$$

Proof. The inequality (2.8) provides $\Lambda_\mu(h) \leq \int U^*(h) d\mu$ for all bounded functions h . Applying Lemma 2.17 with $W = U^*$ (which is in this case of class \mathcal{C}^1) yields

$$\sup_{g \in \mathcal{B}_b} \int U^*(Q_{\lambda_1} g - \text{med}_\mu(g)) d\mu \leq U^*(0) + \int_0^{+\infty} (U^*)'(t) \alpha(\sqrt{\lambda_1 t}) dt.$$

The rest of the proof follows from Lemmas 2.15 and 2.16 and Theorem 2.6. \square

Let us emphasize two important particular cases corresponding to Gaussian and exponential concentrations.

Corollary 2.20 (*Gaussian concentration*). *Let μ be a probability measure on \mathcal{X} . If μ verifies the concentration inequality (2.1) with $\alpha(r) = M e^{-ar^2}$, $r \geq 0$, for some $a, M \geq 0$, then μ verifies $\mathbf{T}_2(u/a, c(u)M/a)$, for all $u > 1$, with $c(u) = 4u/(u - 1)$.*

Proof. In this case, $U(x) = x \log(x) + 1 - x$, $x \geq 0$ and $U^*(x) = e^x - 1$, $x \in \mathbb{R}$. For all $\lambda_1 > 1/a$,

$$\int_0^{+\infty} (U^*)'(t) \alpha(\sqrt{\lambda_1 t}) dt = M \int_0^{+\infty} e^{-(a\lambda_1 - 1)t} dt = \frac{M}{a\lambda_1 - 1}.$$

On the other hand, for all $\lambda_2 > 0$, $\int_0^{+\infty} \alpha(\sqrt{\lambda_2 t}) dt = M \int_0^{+\infty} e^{-a\lambda_2 t} dt = M/(a\lambda_2)$. According to Theorem 2.19, we conclude that μ verifies the inequality $\mathbf{T}_2(c_1, c_2)$ with

$$c_1 = \lambda_1 + \lambda_2 \quad \text{and} \quad c_2 = (\lambda_1 + \lambda_2) \left(\frac{M}{a\lambda_1 - 1} + \frac{M}{a\lambda_2} \right), \quad \forall \lambda_1 > 1/a, \forall \lambda_2 > 0.$$

Equivalently, for all $u > 1$, μ verifies $\mathbf{T}_2(u/a, c_2)$, with

$$c_2 = \frac{Mu}{a} \inf_{u > a\lambda_1 > 1} \left(\frac{1}{a\lambda_1 - 1} + \frac{1}{u - a\lambda_1} \right) = \frac{4Mu}{a(u - 1)},$$

since the infimum is attained at $\lambda_1 = (u + 1)/(2a)$. This completes the proof. \square

Corollary 2.21 (*Exponential concentration*). *Let μ be a probability measure on \mathcal{X} . If μ verifies the concentration inequality (2.1) with $\alpha(r) = M e^{-ar}$, $r \geq 0$, for some $a, M \geq 0$, then μ verifies $\mathbf{T}_2 U(u/a^2, c(u)M/a^2)$, for all $u > 1$, with $U(x) = x \log^2(e + x)$ and $c(u) = 8u/(\sqrt{2u - 1} - 1)$.*

Proof. According to Lemma A.1(iii), for all $\lambda_1 > 1/a^2$

$$\int_0^{+\infty} (U^*)'(t) \alpha(\sqrt{\lambda_1 t}) dt \leq M \int_0^{+\infty} e^{-(a\sqrt{\lambda_1} - 1)\sqrt{t}} dt < +\infty.$$

Hence, applying Theorem 2.19, μ verifies $\mathbf{T}_2 U(c_1, c_2)$, with $c_1 = \lambda_1 + \lambda_2$, $\lambda_1 > 1/a^2$ and

$$\begin{aligned}
 c_2 &= (\lambda_1 + \lambda_2)M \left(\int_0^{+\infty} e^{-(a\sqrt{\lambda_1}-1)\sqrt{t}} dt + \int_0^{+\infty} e^{-a\sqrt{\lambda_2}t} dt \right) \\
 &= 2(\lambda_1 + \lambda_2)M \left(\frac{1}{(\sqrt{a^2\lambda_1} - 1)^2} + \frac{1}{a^2\lambda_2} \right),
 \end{aligned}$$

since $\int_0^{+\infty} e^{-\sqrt{t}} dt = 2$. Equivalently, μ verifies $\mathbf{T}_2U(c_1, c_2)$, with $c_1 = u/a^2$, $u > 1$ and

$$c_2 = \frac{2Mu}{a^2} \inf_{1 < s < u} \left(\frac{1}{(\sqrt{s} - 1)^2} + \frac{1}{u - s} \right) \leq \frac{8Mu}{a^2(\sqrt{2u - 1} - 1)},$$

taking $s = \frac{1}{2}(1 + \sqrt{2u - 1}) \in (1, u)$ for which $(u - s) = (\sqrt{s} - 1)^2$. \square

Remark 2.22 (*Integrability conditions for transport inequalities*). Let us mention that in the literature, many papers have adopted another point of view to relate transport inequalities with tails estimates of μ . It was first observed by Djellout, Guillin and Wu [16] that the integrability condition

$$I = \int e^{ad^2(x,x_0)} d\mu(x) < \infty,$$

for some $a > 0$ and $x_0 \in \mathcal{X}$, implies Talagrand’s \mathbf{T}_1 transport inequality:

$$\mathcal{W}_1(v, \mu) \leq \sqrt{CH(v|\mu)}, \quad \forall v \in \mathcal{P}(\mathcal{X}),$$

where the constant C depends on I . After that, many variants have been proposed to handle different transport costs with different tails behaviours [13,17,19]. All these results are dimensional since the quantity I depends on the dimension of \mathcal{X} . E. Milman [36] has obtained a universal translation of concentration inequalities in terms of transport inequalities of the form

$$\mathcal{W}_1(v, \mu) \leq C\Psi(H(v|\mu)), \quad \forall v \in \mathcal{P}(\mathcal{X}),$$

where Ψ is some concave function related to the concentration function, and C is a constant independent of the dimension. For our purpose, Milman’s results are not adapted since we need to control \mathcal{W}_2 .

2.3. A characterization of dimension free Gaussian concentration

A consequence of the preceding section is that it enables us to give a completely analytic proof of a recent result by the first named author about the equivalence between dimension free Gaussian concentration and Talagrand’s inequality [18]. This was pointed out to us by M. Ledoux.

Corollary 2.23. *A probability measure μ on \mathcal{X} verifies $\mathbf{T}_2(C)$ if and only if for all $a < 1/C$, there is some positive $M(a)$ such that for all positive integers n , the product probability measure μ^n verifies the concentration inequality*

$$\mu^n(A^r) \geq 1 - M(a)e^{-ar^2}, \quad \forall r \geq 0,$$

for all $A \subset \mathcal{X}^n$ with $\mu^n(A) \geq 1/2$, where the enlargement A^r is defined by

$$A^r = \left\{ x \in \mathcal{X}^n; \inf_{y \in A} \sum_{i=1}^n d^2(x_i, y_i) \leq r^2 \right\}.$$

Proof. The fact that \mathbf{T}_2 implies dimension free Gaussian concentration is well known, so we will only prove the converse. According to Corollary 2.20, the assumed concentration property implies that for all $a < 1/C$, for all positive integers n and all $u > 1$,

$$\mathcal{T}_2(\beta, \mu^n) \leq \frac{u}{a} H(\beta|\mu^n) + \frac{4M(a)u}{a(u-1)},$$

for all probability measure β on \mathcal{X}^n (here the transport cost is defined with respect to the metric $d_2(x, y) = \sqrt{\sum_{i=1}^n d^2(x_i, y_i)}$ on \mathcal{X}^n). In particular, taking $\beta = \nu^n$ and using the following easy to check relations: $\mathcal{T}_2(\nu^n, \mu^n) = n\mathcal{T}_2(\nu, \mu)$ and $H(\nu^n|\mu^n) = nH(\nu|\mu)$, we obtain

$$\mathcal{T}_2(\nu, \mu) \leq \frac{u}{a} H(\nu|\mu) + \frac{1}{n} \cdot \frac{4M(a)u}{a(u-1)}.$$

Letting $n \rightarrow \infty$ and then $u \rightarrow 1$ and $a \rightarrow 1/C$, we arrive at $\mathbf{T}_2(C)$, which completes the proof. \square

3. Log-Sobolev inequality: proof of Theorem 1.13

This section is devoted to the proof of the following quantitative version of Theorem 1.13.

Theorem 3.1. Define $c(u) = 4u/(u - 1)$, $u > 1$, and for $M > 0$

$$\tau(M) = \sup \left\{ \frac{r}{Mc(u) + ru}; u > 1, 0 \leq r < \log(2) \right\} = \frac{\log(2)}{(2\sqrt{M} + \sqrt{\log(2)})^2}.$$

Suppose that (\mathcal{X}, d, μ) has ∞ -Ricci curvature bounded below by $K \leq 0$ and assume that μ verifies the Gaussian concentration property (1.1) with positive constants a and M .

If the constants a, M and K satisfy the relation $|K|/(2a) < \tau(M)$, then μ verifies the logarithmic Sobolev inequality **LSI**(C) for some C depending only on K, a and M .

More precisely, if $|K|/(2a) < r/(Mc(u) + ru)$ for some $u > 1$ and $r \in (0, \log(2))$, then μ verifies for all $f : \mathcal{X} \rightarrow \mathbb{R}$ smooth enough

$$\text{Ent}_\mu(f^2) \leq \frac{1}{a} B(u, r, M, K) \int |\nabla^- f|^2 d\mu,$$

with $B(u, r, M, K) = \frac{(Mc(u)+ru)^2}{Mc(u)(2r-|K|(Mc(u)+ru))} (1 + 2\frac{2+r}{2-e^r})$.

In particular, when $K = 0$, the following logarithmic Sobolev inequality holds

$$\text{Ent}_\mu(f^2) \leq \frac{DM}{a} \int |\nabla^- f|^2 d\mu,$$

where D is some absolute numerical constant.

The sketch of the proof of this theorem is the following. Corollary 2.20 ensures that the Gaussian concentration hypothesis implies $\mathbf{T}_2(c_1, c_2)$. Proposition 3.2 indicates that $\mathbf{T}_2(c_1, c_2)$ and the curvature condition imply a non-tight logarithmic-Sobolev inequality. The proof is completed by tightening this Sobolev inequality thanks to a Poincaré inequality.

Proposition 3.2. *Suppose that (\mathcal{X}, d, μ) has ∞ -Ricci curvature bounded below by $K \leq 0$. If μ verifies the transport inequality $\mathbf{T}_2(c_1, c_2)$, then it verifies the following non-tight logarithmic Sobolev inequality: for all bounded Lipschitz functions f*

$$\text{Ent}_\mu(f^2) \leq \frac{2(c_2 + c_1 r)^2}{c_2(2r - |K|(c_2 + c_1 r))} \int |\nabla^- f|^2 d\mu + r \int f^2 d\mu,$$

for all $r > |K|c_2/(2 - |K|c_1)$.

Proof. Under the curvature condition, by Proposition 1.11 applied with the function $U(x) = x \log(x)$, the following HWI inequality holds

$$\begin{aligned} H(v|\mu) &\leq \sqrt{\mathcal{T}_2(v, \mu)} \sqrt{I_{\mu,U}(v)} - \frac{K}{2} \mathcal{T}_2(v, \mu), \quad v \in \mathcal{P}(\mathcal{X}) \\ &\leq \frac{\lambda + |K|}{2} \mathcal{T}_2(v, \mu) + \frac{1}{2\lambda} I_{\mu,U}(v), \end{aligned}$$

for all $\lambda > 0$, where $I_{\mu,U}(v)$ is the usual Fisher information (associated to $U(x) = x \log x$). Consequently, the transport inequality $\mathbf{T}_2(c_1, c_2)$ yields

$$H(v|\mu) \leq \frac{(\lambda + |K|)c_1}{2} H(v|\mu) + \frac{1}{2\lambda} I_{\mu,U}(v) + \frac{(\lambda + |K|)c_2}{2}.$$

So, if $\lambda + |K| < 2$, it holds

$$H(v|\mu) \leq \frac{1}{\lambda(2 - (\lambda + |K|)c_1)} I_{\mu,U}(v) + \frac{(\lambda + |K|)c_2}{(2 - (\lambda + |K|)c_1)}.$$

Applying this inequality to $dv = \frac{f^2}{\int f^2 d\mu} d\mu$, with a Lipschitz bounded function f , we get

$$\text{Ent}_\mu(f^2) \leq \frac{4}{\lambda(2 - (\lambda + |K|)c_1)} \int |\nabla^- f|^2 d\mu + \frac{(\lambda + |K|)c_2}{(2 - (\lambda + |K|)c_1)} \int f^2 d\mu.$$

Letting $r = \frac{(\lambda + |K|)c_2}{(2 - (\lambda + |K|)c_1)}$ gives the result. \square

To obtain a logarithmic Sobolev inequality we will take advantage of a self-tightening phenomenon first observed by Wang [48] and described in the proposition below.

Proposition 3.3 (*Self-tightening phenomenon*). *If a probability measure μ verifies the non-tight logarithmic-Sobolev inequality*

$$\text{Ent}_\mu(f^2) \leq b_1 \int |\nabla^- f|^2 d\mu + b_2 \int f^2 d\mu,$$

with $b_2 < \log(2)$, then it satisfies the logarithmic Sobolev inequality **LSI**(C) with

$$C = b_1 \left(1 + 2 \frac{2 + b_2}{2 - e^{b_2}} \right).$$

Remark 3.4. The condition $b_2 < \log(2)$ is sharp. Wang has obtained a counterexample in [48].

The proof of Proposition 3.3 is based on the two following lemmas due to Wang [47,48].

Lemma 3.5 (*Non-tight Poincaré inequality*). *If a probability measure μ verifies the non-tight logarithmic-Sobolev inequality*

$$\text{Ent}_\mu(f^2) \leq b_1 \int |\nabla^- f|^2 d\mu + b_2 \int f^2 d\mu, \tag{3.6}$$

for all bounded Lipschitz functions f , then it verifies the following Poincaré type inequality

$$\int f^2 d\mu \leq b_1 \int |\nabla^- f|^2 d\mu + e^{b_2} \left(\int |f| d\mu \right). \tag{3.7}$$

Proof. Below we improve the constants in a proof of Wang [47]. Take f such that $\int |f| d\mu = 1$ and denote $\alpha = \int f^2 d\mu$ and $\beta = \int |\nabla^- f|^2 d\mu$. Then, we use the formula

$$x \log(x^2/\alpha) = \sup_{s \in \mathbb{R}} \{sx - 2\sqrt{\alpha}e^{\frac{s}{2}-1}\}.$$

So, for all $s \in \mathbb{R}$,

$$s\alpha - 2\sqrt{\alpha}e^{\frac{s}{2}-1} \leq \int |f| \log\left(\frac{f^2}{\alpha}\right) |f| d\mu \leq b_1\beta + b_2\alpha.$$

So, in particular, $(s - b_2)\alpha - 2\sqrt{\alpha}e^{\frac{s}{2}-1} - b_1\beta \leq 0$, for all $s > b_2$. One conclude from this that

$$\sqrt{\alpha} \leq \frac{e^{\frac{s}{2}-1} + \sqrt{e^{s-2} + (s - b_2)b_1\beta}}{s - b_2},$$

and this implies that

$$\alpha \leq \frac{2b_1}{s - b_2} \beta + \frac{4e^{s-2}}{(s - b_2)^2}.$$

In other words, for all f it holds

$$\int f^2 d\mu \leq \frac{2b_1}{s - b_2} \int |\nabla^- f|^2 d\mu + \frac{4e^{s-2}}{(s - b_2)^2} \left(\int |f| d\mu \right)^2, \quad s > b_2,$$

and therefore, since $\inf_{s > b_2} 4e^{s-2}/(s - b_2)^2$ is reached for $s = b_2 + 2$, one gets (3.7). \square

The next lemma states that if the second constant in the Poincaré type inequality (3.7) is sufficiently small then the Poincaré inequality holds.

Lemma 3.8. *If a probability measure μ verifies the inequality*

$$\int f^2 d\mu \leq d_1 \int |\nabla^- f|^2 d\mu + d_2 \left(\int |f| d\mu \right)^2, \tag{3.9}$$

for all bounded Lipschitz functions f with a constant $d_2 < 2$, then it verifies the following Poincaré inequality

$$\text{Var}_\mu(f) \leq \frac{2d_1}{2 - d_2} \int |\nabla^- f|^2 d\mu.$$

Remark 3.10. Wang [48] has shown that the condition $d_2 < 2$ is optimal.

Proof of Lemma 3.8. Take f a bounded Lipschitz function and consider the bounded Lipschitz functions

$$f_+ = \max(f - m, 0) \quad \text{and} \quad f_- = \min(f - m, 0),$$

where m is the median of f . It is not difficult to check that

$$|\nabla^- f_+| = |\nabla^- f| \mathbb{1}_{\{f > m\}} \quad \text{and} \quad |\nabla^- f_-| = |\nabla^- f| \mathbb{1}_{\{f \leq m\}}. \tag{3.11}$$

Apply (3.9) to the function f_+ ; then Cauchy–Schwarz inequality yields

$$\int_{f > m} (f - m)^2 d\mu \leq d_1 \int_{f > m} |\nabla^- f|^2 d\mu + \frac{d_2}{2} \int_{f > m} (f - m)^2 d\mu.$$

Doing the same with f_- and summing the two inequalities yields

$$\int (f - m)^2 d\mu \leq d_1 \int |\nabla^- f|^2 d\mu + \frac{d_2}{2} \int (f - m)^2 d\mu,$$

and so

$$\int (f - m)^2 d\mu \leq \frac{2d_1}{2 - d_2} \int |\nabla^- f|^2 d\mu.$$

Since, $\text{Var}_\mu(f) \leq \int (f - m)^2 d\mu$, this ends the proof. \square

Proof of Proposition 3.3. Lemma 3.5 and (3.6) imply that

$$\int f^2 d\mu \leq b_1 \int |\nabla^- f|^2 d\mu + e^{b_2} \left(\int |f| d\mu \right)^2.$$

According to Lemma 3.8, we conclude that if $b_2 < \log 2$, then μ verifies the following Poincaré inequality:

$$\text{Var}_\mu(f) \leq \frac{2b_1}{2 - e^{b_2}} \int |\nabla^- f|^2 d\mu. \tag{3.12}$$

We can now tighten the inequality (3.6). Namely, according to Rothaus’ lemma [41] (see also [2, Lemma 4.3.8]), it holds

$$\text{Ent}_\mu(f^2) \leq \text{Ent}_\mu(\bar{f}^2) + 2 \text{Var}_\mu(f),$$

with $\bar{f} = f - \int f d\mu$. So applying (3.6) together with (3.12) gives the result. \square

Proof of Theorem 3.1. According to Corollary 2.20, μ verifies $\mathbf{T}_2(u/a, c(u)M/a)$, for all $u > 1$, with $c(u) = 4u/(u - 1)$. Thus it follows from Proposition 3.2 that

$$\text{Ent}_\mu(f^2) \leq \frac{1}{a} b(u, r, M, K) \int |\nabla^- f|^2 d\mu + r \int f^2 d\mu, \tag{3.13}$$

with

$$b(u, r, M, K) = \frac{(Mc(u) + ru)^2}{Mc(u)(2r - |K|(Mc(u) + ru))},$$

for all $r > \frac{|K|Mc(u)}{2a - |K|u}$ or equivalently when $\frac{|K|}{a} < \frac{2r}{Mc(u) + ru}$. According to Proposition 3.3, we conclude that if $\frac{|K|}{a} < \frac{2r}{Mc(u) + ru}$ for some $u > 1$ and $r < \log 2$, then

$$\text{Ent}_\mu(f^2) \leq \frac{1}{a} B(u, r, M, K) \int |\nabla^- f|^2 d\mu,$$

with $B(u, r, M, K) = b(u, r, M, K)(1 + 2\frac{2+r}{2-e^r})$. The proof of Theorem 3.1 is complete. \square

4. Poincaré inequality: proof of Theorem 1.14

In this section we deal with probability measures verifying an exponential concentration inequality as follows

$$\mu(A^r) \geq 1 - Me^{-ar}, \quad \forall r \geq 0, \tag{4.1}$$

for all $A \subset \mathcal{X}$ such that $\mu(A) \geq 1/2$. In all this part, the measured length space (\mathcal{X}, d, μ) is assumed to have ∞ -Ricci curvature bounded below by 0. The basic reason is that we want to use the HWI inequality with $U(x) = x \log^2(e + x)$, $x \geq 0$, for which $\kappa(U) = -\infty$ if $K < 0$ according to (1.8).

The proof of Theorem 1.14 is similar to the proof of Theorem 3.1. We first establish a non-tight “ U -Sobolev” inequality (see Proposition 4.2 below) that provides a non-tight Poincaré inequality (Proposition 4.4). Under null curvature condition, we may also obtain a weak Poincaré inequality (see Proposition 4.7). And it is known in the literature that a non-tight Poincaré inequality together with a weak Poincaré inequality implies a Poincaré inequality (see Proposition 4.7 below). For completeness, this is recalled in Proposition A.2 in Appendix A. Our strategy can be summarized as follows

$$\left. \begin{array}{l} \text{Concentration} \Rightarrow \mathcal{T}_2 U(c_1, c_2) \\ \text{Bounded curv.} \Rightarrow \left. \begin{array}{l} + \\ \text{HWI} \\ \text{HWI} \end{array} \right\} \Rightarrow \left. \begin{array}{l} \text{Non-tight “}U\text{-Sobolev”} \\ + \\ \text{Weak Poincaré} \end{array} \right\} \Rightarrow \text{Poincaré.} \end{array} \right\}$$

Proposition 4.2. *Suppose that (\mathcal{X}, d, μ) has ∞ -Ricci curvature bounded below by 0. If μ verifies (4.1), then, it holds*

$$\int U(g) d\mu \leq b_1 \int \frac{\log^2(e + g)}{e + g} |\nabla^- g|^2 d\mu + b_2, \tag{4.3}$$

for all Lipschitz functions g , where $b_1 = 16u/a^2$ and $b_2 = 4 + \frac{8M}{\sqrt{2u-1}}$.

Proof. The function $U(x) = x \log^2(e + x)$, $x \in [0, \infty)$ is in the class $\mathcal{DC}_\infty \cap \mathcal{C}^2$. Under the non-negative curvature condition, Proposition 1.11 ensures that for every probability measures $d\nu = g d\mu$ with a positive Lipschitz function g such that $\int U(g) d\mu < \infty$,

$$\int U(g) d\mu \leq U(1) + \sqrt{\int g U''(g)^2 |\nabla^- g|^2 d\mu} \sqrt{\mathcal{T}_2(v, \mu)}.$$

Using the inequalities $\sqrt{ab} \leq \varepsilon b + a/(4\varepsilon)$ for all $\varepsilon > 0$, $U(1) = \log^2(e + 1) \leq 2$ and $g/(e + g) \leq 1$, and Lemma A.1(ii), we get for every $\varepsilon > 0$,

$$\int U(g) d\mu \leq 2 + \varepsilon \mathcal{T}_2(v, \mu) + \frac{4}{\varepsilon} \int \frac{\log^2(e + g)}{e + g} |\nabla^- g|^2 d\mu.$$

Under the concentration property (4.1), Corollary 2.21 ensures that

$$\mathcal{T}_2(v, \mu) \leq c_1 \int U(g) d\mu + c_2,$$

with $c_1 = u/a^2$, $c_2 = c(u)M/a^2$ for all $u > 1$ and $c(u) = \frac{8u}{\sqrt{2u-1}-1}$. It follows that for every $\varepsilon < 1/c_1$,

$$\int U(g) d\mu \leq b_1 \int \frac{\log^2(e+g)}{e+g} |\nabla^- g|^2 d\mu + b_2,$$

with $b_1 = 4/(\varepsilon(1 - \varepsilon c_1))$ and $b_2 = (2 + \varepsilon c_2)/(1 - \varepsilon c_1)$. Taking $\varepsilon = 1/(2c_1)$ completes the proof. \square

Proposition 4.4 (*Non-tight Poincaré inequality*). *Assume that there exist some non-negative constants b_1 and b_2 such that for any positive Lipschitz function $g : \mathcal{X} \rightarrow \mathbb{R}^+$ with $\int g d\mu = 1$ and $\int U(g) d\mu < \infty$, it holds*

$$\int g \log^2(e+g) d\mu \leq b_1 \int \frac{\log^2(e+g)}{e+g} |\nabla^- g|^2 d\mu + b_2. \tag{4.5}$$

Then, one has the following non-tight Poincaré inequality: for any bounded Lipschitz function $h : \mathcal{X} \rightarrow \mathbb{R}$,

$$\int h^2 d\mu \leq 16b_1 \int |\nabla^- h|^2 d\mu + 4b_2 e^{2b_2} \left(\int |h| d\mu \right)^2. \tag{4.6}$$

Proposition 4.7 (*Weak Poincaré inequality*). *Suppose that (\mathcal{X}, d, μ) has ∞ -Ricci curvature bounded below by 0. If μ verifies*

$$C = \sup_{g \in \mathcal{B}_b} \int \left[Qg - \int g d\mu \right]_+ d\mu < +\infty, \tag{4.8}$$

then for any bounded Lipschitz function $h : \mathcal{X} \rightarrow \mathbb{R}$, for all $s > 0$, one has

$$\text{Var}_\mu(h) \leq \frac{C}{s} \int |\nabla^- h|^2 d\mu + s \text{Osc}(h)^2, \tag{4.9}$$

where $\text{Osc}(h) = \sup(h) - \inf(h)$.

We postpone the proof of these two propositions in order to prove Theorem 1.14.

Proof of Theorem 1.14. First let us show that under the concentration property (4.1), it holds

$$\sup_{g \in \mathcal{B}_b} \int \left[Qg - \int g d\mu \right]_+ d\mu \leq \frac{8M}{a^2}.$$

Namely, for every bounded function g , it holds

$$\begin{aligned} \int \left[Qg - \int g d\mu \right]_+ d\mu &\leq \int \left[Q_{1/2}(Q_{1/2}g) - \int g d\mu \right]_+ d\mu \\ &\leq \int \left[Q_{1/2}(Q_{1/2}g) - \text{med}_\mu(Q_{1/2}g) \right]_+ d\mu \\ &\quad + \left[\text{med}_\mu(Q_{1/2}g) - \int g d\mu \right]_+, \end{aligned}$$

where the first inequality comes from the inequality $Qg \leq Q_{1/2}(Q_{1/2}g)$. According to Lemmas 2.16 and 2.17,

$$\begin{aligned} &\sup_{g \in \mathcal{B}_b} \int \left[Q_{1/2}g - \text{med}_\mu(g) \right]_+ d\mu + \sup_{g \in \mathcal{B}_b} \left[\text{med}_\mu(Q_{1/2}g) - \int g d\mu \right]_+ \\ &\leq 2M \int_0^{+\infty} e^{-a\sqrt{t/2}} dt = \frac{8M}{a^2}, \end{aligned}$$

which gives the result.

According to Propositions 4.2, 4.4, 4.7 and A.2 we conclude that μ verifies a Poincaré inequality of the form

$$\text{Var}_\mu(f) \leq \frac{c(M)}{a^2} \int |\nabla^- f|^2 d\mu,$$

for all bounded Lipschitz functions f . \square

The two following subsections are devoted to the proofs of Proposition 4.4 and Proposition 4.7.

4.1. Non-tight Poincaré inequality

Eq. (4.5) is close to (but yet different from) an inequality called $I(\tau)$ introduced by Kolesnikov [24] and further studied in [7]. Using some techniques from [7] we deduce from (4.5) a non-tight Poincaré inequality.

Proof of Proposition 4.4. Let $\psi : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ be the inverse function of U , $\psi = U^{-1}$. The function ψ is increasing and concave with $\psi(0) = 0$ since U is increasing and convex. Fix a bounded positive Lipschitz function f on \mathcal{X} with $\int f d\mu = 1$ and consider its Luxembourg-norm like

$$L = \inf \left\{ \lambda : \int \psi \left(\frac{f}{\lambda} \right) d\mu \leq 1 \right\}.$$

Set $g = \psi(f/L)$. By construction, $\int g d\mu = 1$. Since ψ is increasing, one has $|\nabla^- g| = \psi'(f/L)|\nabla^-(f/L)|$. Hence, applying (4.5) to g leads to

$$\int f d\mu \leq b_1 L \int \frac{\log^2(e + \psi(f/L))}{e + \psi(f/L)} (\psi'(f/L))^2 |\nabla^-(f/L)|^2 d\mu + b_2 L. \tag{4.10}$$

Lemma A.1(i) implies that

$$(\psi'(u))^2 = \frac{1}{U'(\psi(u))^2} \leq \frac{1}{\log^4(e + \psi(u))} = \frac{\psi(u)}{u \log^2(e + \psi(u))}.$$

Therefore

$$\int f \, d\mu \leq b_1 \int \frac{|\nabla^- f|^2}{f} \, d\mu + b_2 L.$$

From the concavity of ψ , $1 = \int \psi(f/L) \, d\mu \leq \psi(\int f \, d\mu/L) = \psi(1/L)$. Consequently $L \leq 1/U(1) < 1$ and

$$L = L \int \psi(f/L) \, d\mu \leq \int \psi(f) \, d\mu.$$

Hence

$$\int f \, d\mu \leq b_1 \int \frac{|\nabla^- f|^2}{f} \, d\mu + b_2 \int \psi(f) \, d\mu.$$

The latter applied to $f = h^2/\mu(h^2)$ leads to

$$\int h^2 \, d\mu \leq 4b_1 \int |\nabla^- h|^2 \, d\mu + b_2 \sqrt{\mu(h^2)} \int |h| F\left(\frac{|h|}{\sqrt{\mu(h^2)}}\right) \, d\mu, \tag{4.11}$$

where $\mu(h^2) = \int h^2 \, d\mu$ and $F(y) = \psi(y^2)/y, y > 0$.

The next step is to bound the function F by affine functions. Since $x/s \leq U(x) + U^*(1/s)$ for every $s > 0, x \geq 0$, then one has for every real numbers y ,

$$\psi(y^2) \leq y^2 s + s U^*(1/s).$$

It follows that

$$F(y) = \frac{\psi(y^2)}{\sqrt{\psi(y^2) \log^2(e + \psi(y^2))}} \leq \sqrt{\psi(y^2)} \leq |y| \sqrt{s} + \sqrt{s U^*(1/s)}.$$

Fix a bounded Lipschitz function h (not necessarily positive) with $\int |h| \, d\mu = 1$ and set $\alpha = \sqrt{\mu(h^2)}$. It follows from (4.11) and the previous computations that for $0 < s \leq 1/b_2^2$,

$$\alpha^2(1 - b_2 \sqrt{s}) - \alpha b_2 \sqrt{s U^*(1/s)} - \beta \leq 0,$$

where $\beta = 4b_1 \int |\nabla^- h|^2 \, d\mu$. This implies that

$$\alpha \leq \frac{b_2 \sqrt{s U^*(1/s)} + \sqrt{b_2^2 s U^*(1/s) + 4\beta(1 - b_2 \sqrt{s})}}{2(1 - b_2 \sqrt{s})}, \quad \forall s \in (0, 1/b_2^2),$$

and therefore

$$\int h^2 d\mu \leq \frac{8b_1}{1 - b_2\sqrt{s}} \int |\nabla^- h|^2 d\mu + \frac{b_2^2 s U^*(1/s)}{(1 - b_2\sqrt{s})^2}.$$

Choosing $s = 1/(4b_2^2)$ and using Lemma A.1(iv) leads to the desired inequality (4.6). \square

4.2. Weak Poincaré inequality

In this section, we prove Proposition 4.7.

Proof of Proposition 4.7. The weak Poincaré inequality (4.9) is a simple consequence of the usual HWI inequality (1.12) for $U(x) = x \log(x)$ that holds when the ∞ -Ricci curvature is bounded below by 0. Namely, for any bounded Lipschitz function $f > 0$ with $\int f d\mu = 1$,

$$\text{Ent}_\mu(f) \leq \sqrt{\mathcal{T}_2(f\mu, \mu) \int \frac{|\nabla^- f|^2}{f} d\mu},$$

and therefore for all $s > 0$,

$$\text{Ent}_\mu(f) \leq s\mathcal{T}_2(f\mu, \mu) + \frac{1}{4s} \int \frac{|\nabla^- f|^2}{f} d\mu. \tag{4.12}$$

The first step is to bound $\mathcal{T}_2(f\mu, \mu)$. By the Kantorovich’s dual characterization of \mathcal{T}_2 ,

$$\mathcal{T}_2(f\mu, \mu) = \sup_{g \in \mathcal{C}_b} \int \left(Qg - \int g d\mu \right) f d\mu.$$

Since $Qg \leq g$, it holds

$$\begin{aligned} \int \left(Qg - \int g d\mu \right) f d\mu &= \int \left(Qg - \int g d\mu \right) (f - \inf f) d\mu \\ &\quad + \inf(f) \left(\int Qg d\mu - \int g d\mu \right) \\ &\leq \text{Osc}(f) \int \left[Qg - \int g d\mu \right]_+ d\mu. \end{aligned}$$

Consequently one has

$$\mathcal{T}_2(f\mu, \mu) \leq C \text{Osc}(f).$$

By homogeneity, applying (4.12) to $f = (h - \inf h)^2 / \int (h - \inf h)^2 d\mu$, it follows that

$$\begin{aligned} \text{Ent}_\mu((h - \inf h)^2) &\leq sC \text{Osc}((h - \inf h)^2) + \frac{1}{s} \int |\nabla^-(h - \inf h)|^2 d\mu \\ &= sC \text{Osc}^2(h) + \frac{1}{s} \int |\nabla^-h|^2 d\mu. \end{aligned}$$

The standard inequality $\text{Var}_\mu(f) \leq \text{Ent}_\mu(f^2)$ for $f \geq 0$ (see e.g. [2, inequality (1.9)]) ends the proof of (4.9) since

$$\text{Var}_\mu(h) = \text{Var}_\mu(h - \inf h) \leq \text{Ent}_\mu((h - \inf h)^2). \quad \square$$

5. Extensions

As mentioned in the introduction, our approach generalizes to other types of concentration (different from Gaussian and exponential). We present in this section some results in this direction, without details.

For example, one could consider concentration between exponential and Gaussian of the type: for all $A \subset \mathcal{X}$ with $\mu(A) \geq 1/2$,

$$\mu(A^r) \geq 1 - Me^{-ar^{2/\gamma}}, \quad \forall r \geq 0, \tag{5.1}$$

where $\gamma \in [1, 2)$. In this case, one has to apply Proposition 1.11 with the function $U(x) = x \log^\gamma(e + x)$, $x \geq 0$, which belongs to the class $\mathcal{DC}_\infty \cap \mathcal{C}^2$. This, together with Theorem 2.19 and few rearrangements, lead to a non-tight inequality of the type (for any positive Lipschitz function f with $\int f d\mu = 1$)

$$\int U(f) d\mu \leq \frac{b_1}{a^\gamma} \int \frac{\log^{2(\gamma-1)}(e + f)}{e + f} |\nabla^- f|^2 d\mu + b_2$$

for some positive constants b_1, b_2 depending only on M .

Now, following the same lines as in the proof of Proposition 4.4 (see Section 4), it is possible to derive the following non-tight F -Sobolev inequality

$$\int f^2 \log^{2-\gamma}(e + f^2) d\mu \leq \frac{c_1}{a^\gamma} \int |\nabla^- f|^2 d\mu + c_2, \tag{5.2}$$

for all bounded Lipschitz functions f with $\int f^2 d\mu = 1$ (where c_1, c_2 are positive constants depending only on M), and also a non-tight Poincaré inequality.

If (\mathcal{X}, d, μ) has ∞ -Ricci curvature bounded below by 0, Proposition 4.7 applies and leads to a weak Poincaré inequality. This inequality together with the non-tight Poincaré inequality previously obtained imply a Poincaré inequality, as explained in Proposition A.2.

Finally, one applies the analogue of Rothaus' lemma for F -Sobolev inequalities (see [6,39,7]) in order to tighten inequality (5.2) and end up with

$$\int f^2(\log^{2-\gamma}(e + f^2) - \log^{2-\gamma}(e + 1)) d\mu \leq \frac{D}{a^\gamma} \int |\nabla^- f|^2 d\mu,$$

for all bounded Lipschitz functions f with $\int f^2 d\mu = 1$, where D is a positive constant that depends only on M . Such an inequality does not enjoy the tensorization property as the Poincaré inequality or the logarithmic Sobolev inequality. However, it is known [50] to be equivalent to Beckner–Latała–Oleszkiewicz inequalities (i.e. inequality (5.5) below) that do tensorize [8,25].

Adjusting all the previous computations would lead to the following theorem.

Theorem 5.3. *Suppose that (\mathcal{X}, d, μ) has ∞ -Ricci curvature bounded below by 0 and fix $\gamma \in [1, 2)$. If μ verifies the concentration property (5.1) for some $M, a > 0$, then there exists a constant C that depends only on M such that for any bounded Lipschitz function $f : \mathcal{X} \rightarrow \mathbb{R}$ with $\int f^2 d\mu = 1$, it holds*

$$\int f^2 (\log^{2-\gamma}(e + f^2) - \log^{2-\gamma}(e + 1)) d\mu \leq \frac{C}{a^\gamma} \int |\nabla^- f|^2 d\mu, \tag{5.4}$$

and

$$\sup_{p \in (1,2)} \frac{\int f^2 d\mu - (\int |f|^p d\mu)^{2/p}}{(2 - p)^{2-\gamma}} \leq \frac{C}{a^\gamma} \int |\nabla^- f|^2 d\mu. \tag{5.5}$$

As the reader might have noticed, inequality (5.4) has the flavour of the logarithmic Sobolev inequality and the Poincaré inequality, respectively, when $\gamma = 1$ and $\gamma \rightarrow 2$, respectively.

Acknowledgments

The authors warmly thank Christian Léonard and Michel Ledoux for useful discussions on the topics of this paper.

Appendix A. Technical results

In this appendix we collect some technical facts about the function $U(x) = x \log^2(e + x)$. Also, we recall the known result that a non-tight Poincaré inequality together with any weak Poincaré inequality imply a (tight) Poincaré inequality.

Lemma A.1. *Let $U(x) = x \log^2(e + x)$ for $x \geq 0$. Then,*

- (i) $\log^2(e + x) \leq U'(x), x \geq 0$.
- (ii) $U''(x) \leq 4 \frac{\log(e+x)}{e+x}, x \geq 0$.
- (iii) $U^*(x) \leq -e + \exp\{\sqrt{x}\}, x \geq 1$.
- (iv) $U^*(x) \leq 2\sqrt{x}e^{\sqrt{x}}, x \geq 0$.

Proof. Point (i) follows from

$$U'(x) = \log(e + x)^2 + \frac{2x}{e + x} \log(e + x).$$

Point (ii) is a consequence of the fact that

$$U''(x) = \frac{2 \log(e+x)}{e+x} \left(\frac{2e+x}{e+x} + \frac{x}{(e+x) \log(e+x)} \right)$$

and that $0 \leq \frac{2e+x}{e+x} + \frac{x}{(e+x) \log(e+x)} \leq 2$. We omit details. Point (iii) follows from point (i) and $U^{*'} = U'^{-1}$. Using point (iii), $U^*(0) = -\inf_y \{U(y)\} = 0$ and an integration by parts, we get point (iv):

$$\begin{aligned} U^*(x) &= \int_0^x U^{*'}(y) dy \leq -ex + \int_0^x e^{\sqrt{y}} dy = -ex + 2(\sqrt{x}e^{\sqrt{x}} - e^{\sqrt{x}} + 1) \\ &\leq 2\sqrt{x}e^{\sqrt{x}}. \quad \square \end{aligned}$$

The next proposition shows that the Poincaré inequality is a consequence of both non-tight Poincaré inequality and weak Poincaré inequality. This result is well known, see e.g. [47,49,7,51]. We write here the version by Wang [49, Corollary 4.1.2].

Proposition A.2 (Wang). *Assume that there exist two constants $d_1, d_2 > 0$ and a non-increasing positive function β , on $(0, 1/2)$ such that for any bounded Lipschitz function $f : \mathcal{X} \rightarrow \mathbb{R}$, it holds*

$$\int f^2 d\mu \leq d_1 \int |\nabla^- f|^2 d\mu + d_2 \left(\int |f| d\mu \right)^2 \quad (\text{Non-tight Poincaré})$$

and

$$\text{Var}_\mu(f) \leq \beta(s) \int |\nabla^- f|^2 d\mu + s \text{Osc}(f)^2 \quad (\text{Weak Poincaré}).$$

Then, every bounded Lipschitz functions $f : \mathcal{X} \rightarrow \mathbb{R}$ verifies

$$\text{Var}_\mu(f) \leq \sup_{s \in (0, d_2^{-1})} \frac{d_1 + \beta(s)}{1 - \sqrt{d_2 s}} \int |\nabla^- f|^2 d\mu \quad (\text{Poincaré}).$$

References

- [1] M. Agueh, N. Ghoussoub, X. Kang, Geometric inequalities via a general comparison principle for interacting gases, *Geom. Funct. Anal.* 14 (1) (2004) 215–244.
- [2] C. Ané, S. Blachère, D. Chafaï, P. Fougères, I. Gentil, F. Malrieu, C. Roberto, G. Scheffer, *Sur les inégalités de Sobolev logarithmiques*, Panor. Synthèses, vol. 10, Société Mathématique de France, Paris, 2000, with a preface by Dominique Bakry and Michel Ledoux.
- [3] D. Bakry, M. Ledoux, Lévy–Gromov’s isoperimetric inequality for an infinite-dimensional diffusion generator, *Invent. Math.* 123 (2) (1996) 259–281.
- [4] Z. Balogh, A. Engoulatov, L. Hunziker, O.E. Maasalo, Functional inequalities and Hamilton–Jacobi equations in geodesic spaces, preprint, available on <http://arxiv.org/abs/0906.0476>, 2009.
- [5] F. Barthe, Isoperimetric inequalities, probability measures and convex geometry, in: *European Congress of Mathematics*, Eur. Math. Soc., Zürich, 2005, pp. 811–826.
- [6] F. Barthe, P. Cattiaux, C. Roberto, Interpolated inequalities between exponential and Gaussian, Orlicz hypercontractivity and isoperimetry, *Rev. Mat. Iberoamericana* 22 (3) (2006) 993–1067.
- [7] F. Barthe, A.V. Kolesnikov, Mass transport and variants of the logarithmic Sobolev inequality, *J. Geom. Anal.* 18 (4) (2008) 921–979.

- [8] W. Beckner, A generalized Poincaré inequality for Gaussian measures, *Proc. Amer. Math. Soc.* 105 (2) (1989) 397–400.
- [9] S.G. Bobkov, Isoperimetric and analytic inequalities for log-concave probability measures, *Ann. Probab.* 27 (4) (1999) 1903–1921.
- [10] S.G. Bobkov, On isoperimetric constants for log-concave probability distributions, in: *Geometric Aspects of Functional Analysis*, in: *Lecture Notes in Math.*, vol. 1910, Springer, Berlin, 2007, pp. 81–88.
- [11] S.G. Bobkov, I. Gentil, M. Ledoux, Hypercontractivity of Hamilton–Jacobi equations, *J. Math. Pures Appl.* (9) 80 (7) (2001) 669–696.
- [12] S.G. Bobkov, F. Götze, Exponential integrability and transportation cost related to logarithmic Sobolev inequalities, *J. Funct. Anal.* 163 (1) (1999) 1–28.
- [13] F. Bolley, C. Villani, Weighted Csiszár–Kullback–Pinsker inequalities and applications to transportation inequalities, *Ann. Fac. Sci. Toulouse Math.* (6) 14 (3) (2005) 331–352.
- [14] D. Cordero-Erausquin, Some applications of mass transport to Gaussian-type inequalities, *Arch. Ration. Mech. Anal.* 161 (3) (2002) 257–269.
- [15] D. Cordero-Erausquin, W. Gangbo, C. Houdré, Inequalities for generalized entropy and optimal transportation, in: *Recent Advances in the Theory and Applications of Mass Transport*, in: *Contemp. Math.*, vol. 353, Amer. Math. Soc., Providence, RI, 2004, pp. 73–94.
- [16] H. Djellout, A. Guillin, L. Wu, Transportation cost-information inequalities and applications to random dynamical systems and diffusions, *Ann. Probab.* 32 (3B) (2004) 2702–2732.
- [17] N. Gozlan, Integral criteria for transportation-cost inequalities, *Electron. Comm. Probab.* 11 (2006) 64–77 (electronic).
- [18] N. Gozlan, A characterization of dimension free concentration in terms of transport inequalities, *Ann. Probab.* 37 (6) (2009) 2480–2498.
- [19] N. Gozlan, C. Léonard, A large deviation approach to some transportation cost inequalities, *Probab. Theory Related Fields* 139 (1–2) (2007) 235–283.
- [20] N. Gozlan, C. Léonard, Transport inequalities. A survey, *Markov Process. Related Fields*, in press, available on <http://arxiv.org/abs/1003.3852>, 2010.
- [21] N. Gozlan, C. Roberto, P.M. Samson, A new characterization of Talagrand’s transport-entropy inequalities and applications, *Ann. Probab.* (2010), doi:10.1214/10-AOP570, in press.
- [22] L. Gross, Logarithmic Sobolev inequalities, *Amer. J. Math.* 97 (4) (1975) 1061–1083.
- [23] L. Gross, Logarithmic Sobolev inequalities and contractivity properties of semi-groups, in: G. Dell’Antonio, U. Mosco (Eds.), *Dirichlet Forms*, in: *Lecture Notes in Math.*, vol. 1563, Springer, Berlin, 1993, pp. 54–88.
- [24] A.V. Kolesnikov, Modified log-Sobolev inequalities and isoperimetry, *Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. Rend. Lincei (9) Mat. Appl.* 18 (2) (2007) 179–208.
- [25] R. Latała, K. Oleszkiewicz, Between Sobolev and Poincaré, in: *Geometric Aspects of Functional Analysis*, in: *Lecture Notes in Math.*, vol. 1745, Springer, Berlin, 2000, pp. 147–168.
- [26] M. Ledoux, A simple analytic proof of an inequality by P. Buser, *Proc. Amer. Math. Soc.* 121 (3) (1994) 951–959.
- [27] M. Ledoux, Isoperimetry and Gaussian analysis, in: *Lectures on Probability Theory and Statistics*, Saint-Flour, 1994, in: *Lecture Notes in Math.*, vol. 1648, Springer, Berlin, 1996, pp. 165–294.
- [28] M. Ledoux, The geometry of Markov diffusion generators, *Ann. Fac. Sci. Toulouse Math.* (6) 9 (2) (2000) 305–366, Probability theory.
- [29] M. Ledoux, *The Concentration of Measure Phenomenon*, *Math. Surveys Monogr.*, vol. 89, Amer. Math. Soc., Providence, RI, 2001.
- [30] M. Ledoux, From concentration to isoperimetry: semigroup proofs, preprint, 2009.
- [31] J. Lott, C. Villani, Hamilton–Jacobi semigroup on length spaces and applications, *J. Math. Pures Appl.* (9) 88 (3) (2007) 219–229.
- [32] J. Lott, C. Villani, Ricci curvature for metric-measure spaces via optimal transport, *Ann. of Math.* (2) 169 (3) (2009) 903–991.
- [33] K. Marton, A simple proof of the blowing-up lemma, *IEEE Trans. Inform. Theory* 32 (3) (1986) 445–446.
- [34] E. Milman, Isoperimetric and concentration inequalities. Equivalence under curvature lower bound, *Duke Math. J.* 154 (2) (2010) 207–239.
- [35] E. Milman, On the role of convexity in functional and isoperimetric inequalities, *Proc. Lond. Math. Soc.* (3) 99 (1) (2009) 32–66.
- [36] E. Milman, Properties of isoperimetric, functional and transport-entropy inequalities via concentration, preprint, available on <http://arxiv.org/abs/0909.0207>, 2010.
- [37] F. Otto, C. Villani, Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality, *J. Funct. Anal.* 173 (2) (2000) 361–400.

- [38] C. Roberto, Isoperimetry for product of probability measures: recent results, *Markov Process. Related Fields*, in press.
- [39] C. Roberto, B. Zegarliński, Orlicz–Sobolev inequalities for sub-Gaussian measures and ergodicity of Markov semi-groups, *J. Funct. Anal.* 243 (1) (2007) 28–66.
- [40] R. Tyrrell Rockafellar, *Convex Analysis*, Princeton Landmarks in Mathematics, Princeton University Press, Princeton, NJ, 1997, reprint of the 1970 original, Princeton Paperbacks.
- [41] O.S. Rothaus, Analytic inequalities, isoperimetric inequalities and logarithmic Sobolev inequalities, *J. Funct. Anal.* 64 (1985) 296–313.
- [42] A.J. Stam, Some inequalities satisfied by the quantities of information of Fisher and Shannon, *Inform. Control* 2 (1959) 101–112.
- [43] K.T. Sturm, On the geometry of metric measure spaces. I, *Acta Math.* 196 (1) (2006) 65–131.
- [44] K.T. Sturm, On the geometry of metric measure spaces. II, *Acta Math.* 196 (1) (2006) 133–177.
- [45] C. Villani, *Optimal Transport*, Grundlehren Math. Wiss., vol. 338, Springer-Verlag, Berlin, 2009, old and new.
- [46] F.Y. Wang, Logarithmic Sobolev inequalities on noncompact Riemannian manifolds, *Probab. Theory Related Fields* 109 (3) (1997) 417–424.
- [47] F.Y. Wang, Functional inequalities for empty essential spectrum, *J. Funct. Anal.* 170 (1) (2000) 219–245.
- [48] F.Y. Wang, Spectral gap for hyperbounded operators, *Proc. Amer. Math. Soc.* 132 (9) (2004) 2629–2638 (electronic).
- [49] F.Y. Wang, *Functional Inequalities, Markov Processes and Spectral Theory*, Science Press, Beijing, 2005.
- [50] F.Y. Wang, A generalization of Poincaré and log-Sobolev inequalities, *Potential Anal.* 22 (1) (2005) 1–15.
- [51] P. Zitt, Super Poincaré inequalities, Orlicz norms and essential spectrum, preprint, available at hal-00426367.