Available online at www.sciencedirect.com
ScienceDirect
Journal of Functional Analysis 250 (2007) 400–425
JOURNAL OF Functional Analysis
www.elsevier.com/locate/jfa

Characterization of Talagrand's like transportation-cost inequalities on the real line

Nathael Gozlan

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

Received 18 October 2006; accepted 24 May 2007

Available online 23 July 2007

Communicated by C. Villani

Abstract

In this paper, we give necessary and sufficient conditions for Talagrand's like transportation cost inequalities on the real line. This brings a new wide class of examples of probability measures enjoying a dimension-free concentration of measure property. Another byproduct is the characterization of modified Log-Sobolev inequalities for log-concave probability measures on $\mathbb{R}$.

Keywords: Transportation cost inequalities; Concentration of measure; Logarithmic Sobolev inequalities

1. Introduction

1.1. Transportation-cost inequalities

This article is devoted to the study of probability measures on the real axis satisfying some kind of transportation-cost inequalities. These inequalities relate two quantities: on the one hand, an optimal transportation cost in the sense of Kantorovich and on the other hand, the relative entropy (also called Kullback–Leibler distance). Let us recall that if $\alpha : \mathbb{R} \to \mathbb{R}^+$ is a continuous

E-mail address: nathael.gozlan@univ-mlv.fr.

0022-1236/$ – see front matter
doi:10.1016/j.jfa.2007.05.025

even function, the optimal transportation-cost to transport $\nu \in \mathcal{P}(\mathbb{R})$ on $\mu \in \mathcal{P}(\mathbb{R})$ (the set of all probability measures on $\mathbb{R}$) is defined by

$$\mathcal{T}_\alpha(\nu, \mu) = \inf_{\pi \in P(\nu,\mu)} \iint_{\mathbb{R}\times\mathbb{R}} \alpha(x-y)\,\pi(dx\,dy), \tag{1}$$

where $P(\nu, \mu)$ is the set of all the probability measures on $\mathbb{R} \times \mathbb{R}$ such that $\pi(dx \times \mathbb{R}) = \nu$ and $\pi(\mathbb{R} \times dy) = \mu$. The relative entropy of ν with respect to μ is defined by

$$\mathrm{H}(\nu \mid \mu) = \begin{cases} \int \log \frac{d\nu}{d\mu}\, d\nu & \text{if } \nu \ll \mu, \\ +\infty & \text{otherwise.} \end{cases} \tag{2}$$

One will say that $\mu \in \mathcal{P}(\mathbb{R})$ satisfies the *Transportation Cost Inequality* (*TCI*) *with the cost function* $(x, y) \mapsto \alpha(x - y)$ if

$$\forall \nu \in \mathcal{P}(\mathbb{R}), \quad \mathcal{T}_\alpha(\nu, \mu) \leqslant \mathrm{H}(\nu \mid \mu). \tag{3}$$

Transportation-cost inequalities of the form (3) were introduced by K. Marton in [15,16] and M. Talagrand in [21]. After them, several authors studied inequality (3), possibly in a multidimensional setting, for particular choices of the cost function α (see for example [3,4,7,9,19,23]). The best known example of transportation-cost inequality is the so-called $\mathbb{T}_2$-inequality (also called Talagrand's inequality). It corresponds to the choice $\alpha(x) = \frac{1}{a}x^2$. One says that μ satisfies $\mathbb{T}_2$ with the constant a if

$$\forall \nu \in \mathcal{P}(\mathbb{R}), \quad \mathcal{T}_2(\nu, \mu) \leqslant a\mathrm{H}(\nu \mid \mu), \tag{4}$$

writing $\mathcal{T}_2(\nu, \mu)$ instead of $\mathcal{T}_{x^2}(\nu, \mu)$.

It will be convenient to introduce a class of slightly stronger transportation-cost inequalities. A probability measure will be said to satisfy the *strong TCI with the cost function* $(x, y) \mapsto \alpha(x - y)$ if

$$\forall \nu, \beta \in \mathcal{P}(\mathbb{R}), \quad \mathcal{T}_\alpha(\nu, \beta) \leqslant \mathrm{H}(\nu \mid \mu) + \mathrm{H}(\beta \mid \mu). \tag{5}$$

Note that this inequality is a sort of symmetrized version of the usual TCI (3). Of course, since $\mathrm{H}(\mu \mid \mu) = 0$, (5) implies (3). In fact, if α is convex, these two inequalities are equivalent up to constant factors (see Proposition 7). Strong TCIs are not new. The strong TCI (5) is in fact equivalent to an infimal-convolution inequality. Infimal-convolution inequalities were introduced by B. Maurey in [17]. The translation of (5) in terms of infimal-convolution inequalities will be stated in Theorem 11.

One of the reasons of the increasing interest to transportation-cost inequalities is their links with the concentration of measure phenomenon. Roughly speaking, a probability measure which satisfies a TCI of the form (3) or (5) also satisfies a *dimension-free* concentration of measure property. This link was first pointed out by K. Marton in [15] and M. Talagrand in [21]. If μ verifies the strong TCI (5), then

$$\forall n \in \mathbb{N}^*,\ \forall A \subset \mathbb{R}^n \text{ measurable},\ \forall r \geqslant 0, \quad \mu^n\left(A_\alpha^r\right) \geqslant 1 - \frac{1}{\mu^n(A)}e^{-r}, \tag{6}$$

where

$$A_\alpha^r = \left\{ x \in \mathbb{R}^n \colon \exists y \in A \text{ such that } \sum_{i=1}^n \alpha\big(|x_i - y_i|\big) \leqslant r \right\}.$$

For example, the $\mathbb{T}_2$-inequality is related to Gaussian dimension-free concentration: if μ satisfies $\mathbb{T}_2$ with constant $a > 0$, then (6) can be restated in the following way:

$$\forall n \in \mathbb{N}^*,\ \forall A \subset \mathbb{R}^n \text{ measurable},\ \forall r \geqslant 0, \quad \mu^n\big(d_2(\cdot, A) \geqslant r\big) \geqslant 1 - \frac{1}{\mu^n(A)} e^{-r^2/a},$$

where d_2 denotes the usual Euclidean metric on $\mathbb{R}^n$.

1.2. Presentation of the results

In this paper, we will give necessary and sufficient conditions under which a probability measure μ on $\mathbb{R}$ satisfies a strong TCI. We will always assume that μ has no atom ($\mu\{x\} = 0$ for all $x \in \mathbb{R}$) and full support ($\mu(A) > 0$ for all open set $A \subset \mathbb{R}$).

First let us define the set of admissible cost functions.

Definition 1 *(Admissible cost functions).* The class $\mathcal{A}$ will be the set of all the functions $\alpha : \mathbb{R} \to \mathbb{R}^+$ such that:

- α is even,
- α is a continuous function, non-decreasing on $\mathbb{R}^+$ with $\alpha(0) = 0$,
- α is super-additive on $\mathbb{R}^+$: $\alpha(x + y) \geqslant \alpha(x) + \alpha(y)$, $\forall x, y \geqslant 0$,
- α is quadratic near 0: $\alpha(t) = |t|^2$, $\forall t \in [-1, 1]$.

One will write $\mu \in \mathbb{T}_\alpha(a)$ (respectively $\mu \in \mathbb{ST}_\alpha(a)$) if μ satisfies the TCI (respectively the strong TCI) with the cost function $(x, y) \mapsto \alpha(a(x - y))$.

1.2.1. The main result

Our main result (Theorem 2) characterizes the strong TCIs on a large class $\mathcal{L}ip_\sharp\mu_1 \subset \mathcal{P}(\mathbb{R})$. Roughly speaking this set is the class of all probability measures which are Lipschitz deformation of the exponential probability measure $d\mu_1(x) = \frac{1}{2}e^{-|x|}\,dx$.

To properly define $\mathcal{L}ip_\sharp\mu_1$ let $T : \mathbb{R} \to \mathbb{R}$: $x \mapsto F^{-1} \circ F_1(x)$, where F (respectively F_1) is the cumulative distribution function of μ (respectively μ_1). The function T is called the *monotone rearrangement map*. It transports μ_1 on μ: $\mu = T\sharp\mu_1$, where $T\sharp\mu_1$ denotes the image of μ_1 under T, i.e.

$$\forall A \subset \mathbb{R} \text{ measurable}, \quad T\sharp\mu_1(A) = \mu_1\big(T^{-1}(A)\big).$$

By definition, the class $\mathcal{L}ip_\sharp\mu_1$ is the class of probability measures μ such that T is Lipschitz.

Theorem 2. *Let $\mu \in \mathcal{L}ip_\sharp \mu_1$ and $\alpha \in \mathcal{A}$. The probability measure μ belongs to $\mathbb{ST}_\alpha(a)$ for some $a > 0$ if and only if there is some $b > 0$ such that*

$$K^+(b) := \sup_{x \geqslant m} \int e^{\alpha(bu)}\, d\mu_x^+(u) < +\infty \quad \text{and} \quad K^-(b) := \sup_{x \leqslant m} \int e^{\alpha(bu)}\, d\mu_x^-(u) < +\infty,$$

where m is the median of μ and where μ_x^+ and μ_x^- are probability measures on $\mathbb{R}^+$ defined as follows:

$$\mu_x^+ = \mathcal{L}(X - x | X \geqslant x) \quad \text{and} \quad \mu_x^- = \mathcal{L}(x - X | X \leqslant x),$$

with X a random variable of law μ.

Remark 3. In a more explicit way, for all bounded and measurable f defined on $\mathbb{R}^+$,

$$\int_0^{+\infty} f(u)\, d\mu_x^+(u) = \frac{1}{\mu[x, +\infty)} \int_x^{+\infty} f(u - x)\, d\mu(u),$$

$$\int_0^{+\infty} f(u)\, d\mu_x^-(u) = \frac{1}{\mu(-\infty, x]} \int_{-\infty}^{x} f(x - u)\, d\mu(u).$$

In particular, the integrability conditions appearing in the preceding theorem read

$$K^+(b) = \sup_{x \geqslant m} \frac{\int_x^{+\infty} e^{\alpha(b(u-x))}\, d\mu(u)}{\mu[x + \infty)} < +\infty \quad \text{and}$$

$$K^-(b) = \sup_{x \leqslant m} \frac{\int_{-\infty}^{x} e^{\alpha(b(x-u))}\, d\mu(u)}{\mu(-\infty, x]} < +\infty.$$

The result furnished by Theorem 2 is quite satisfactory. Though partial, this result covers all the ‘regular’ cases. Namely, it can be shown that whenever μ satisfies $\mathbb{ST}_\alpha(a)$ then it satisfies a Poincaré inequality. But if μ has a positive density of the form $d\mu = e^{-V}\, dx$, where V is a good potential (see the definition bellow), then μ satisfies Poincaré if and only if it belongs to $\mathcal{L}ip_\sharp \mu_1$. We are unable to construct a potential V such that the probability measure $d\mu = e^{-V}\, dx$ satisfies Poincaré without being in $\mathcal{L}ip_\sharp \mu_1$. However, if the density is allowed to vanish, then the probability measures $d\nu_r(x) = \frac{1}{Z_r}|x|^r e^{-|x|}\, dx$, with $r \in (0, 1)$ do the job (see Section 4.1).

In the following theorem we derive from the above result an explicit condition on V ensuring that μ belongs to $\mathbb{ST}_\alpha(a)$ for some $a > 0$. This condition requires that V is regular enough:

Definition 4 *(Good potentials).* The class $\mathcal{V}$ will be the set of all the functions $f : \mathbb{R} \to \mathbb{R}$ of class $\mathcal{C}^2$ such that:

- there is $x_o > 0$ such that $xf'(x) > 0$ for all $x \in (-\infty, -x_o] \cup [x_o, +\infty)$,
- $\frac{f''(x)}{f'^2(x)} \underset{x \to \pm\infty}{\longrightarrow} 0$.

Theorem 5. *Let $d\mu = e^{-V}\,dx$ with $V \in \mathcal{V}$ and $\alpha \in \mathcal{A} \cap \mathcal{V}$, then μ satisfies $\mathbb{ST}_\alpha(a)$ for some constant $a > 0$ as soon as the following conditions hold*:

$$\liminf_{x\to\pm\infty}\left|V'(x)\right| > 0 \quad \textit{and} \quad \exists\lambda > 0 \textit{ such that } \limsup_{x\to\pm\infty}\frac{\alpha'(\lambda x)}{V'(x+m)} < +\infty.$$

The first condition guaranties that μ belongs to $\mathcal{L}ip_\sharp\mu_1$ and the second one that $K^+(b) < +\infty$ and $K^-(b) < +\infty$ for some positive b. In the particular case of $\mathbb{T}_2$-inequality, we recover a sufficient condition of P. Cattiaux and A. Guillin (see [7, Proposition 5.5]).

1.2.2. The particular case of Log-concave distributions

A particularly nice case is when μ is Log-concave. Recall that μ is said to be Log-concave if $\log(1-F)$ is concave, F being the cumulative distribution function of μ. If μ is Log-concave then it belongs to $\mathcal{L}ip_\sharp\mu_1$ (see Proposition 44). The following theorem gives a complete characterization of strong TCIs for Log-concave measures.

Theorem 6. *Let $\alpha \in \mathcal{A}$ and $\mu \in \mathcal{P}(\mathbb{R})$ a Log-concave distribution. The following propositions are equivalent*:

(i) *There is some constant $a > 0$ such that $\mu \in \mathbb{ST}_\alpha(a)$.*
(ii) *There is some constant $b > 0$ such that $\int e^{\alpha(bx)}\,d\mu(x) < +\infty$.*

If $\alpha \in \mathcal{A}$ is convex then the same is true for the TCI.

Using well known techniques, we derive from this result sufficient conditions for the modified Logarithmic Sobolev inequalities introduced by I. Gentil, A. Guillin and L. Miclo in [9]. We prove that under the moment condition $\int e^{\alpha(bx)}\,d\mu(x) < +\infty$, the Log-concave distribution μ satisfies the following inequality

$$\mathrm{Ent}_\mu\left(f^2\right) \leqslant C\int \alpha^*\left(t\frac{f'}{f}\right)f^2\,d\mu, \quad \forall f, \tag{7}$$

for some $c, t > 0$ (see Theorem 47 and Corollary 49). This extends and completes the results of Gentil, Guillin and Miclo (see [9,10]). The interested reader will find other results concerning modified Logarithmic Sobolev inequalities in the recent papers [2] and [13]. Using Hardy type inequalities, F. Barthe and C. Roberto give in [2] fine estimations of optimal constants for the one-dimensional modified Logarithmic Sobolev inequalities (7). In [13], A.V. Kolesnikov gives necessary and sufficient conditions for modified Logarithmic Sobolev inequalities for Log-concave distributions on $\mathbb{R}^d$. His results are expressed in terms of integrability conditions of a certain isoperimetric function.

1.2.3. A word on the method

The classical approach to study TCIs is to relate them to other functional inequalities such as Logarithmic Sobolev inequalities. The main work on the subject is the article by F. Otto and C. Villani on Talagrand's inequality (see [19]). They proved that if $\mu \in \mathcal{P}(\mathbb{R})$ satisfies the Logarithmic Sobolev inequality

$$\mathrm{Ent}_\mu\big(f^2\big) \leqslant C \int f'^2\, d\mu, \quad \forall f$$

then it satisfies Talagrand's inequality (4) with the same constant C. In fact, this result is true in a multidimensional setting. Soon after Otto and Villani, S.G. Bobkov, I. Gentil and M. Ledoux provided an other proof of this result (see [3]). Different authors have tried to generalize this approach to study TCIs associated to other cost functions (see [7,9,23]). This approach has its limits. In [7], P. Cattiaux and A. Guillin were able to construct a probability measure satisfying $\mathbb{T}_2$ but not the Logarithmic Sobolev inequality.

The originality of the present paper is that transportation cost inequalities are studied without the help of Logarithmic Sobolev inequalities. Our results rely on a simple but powerful perturbation method which is explained in Section 3. Roughly speaking, we show that if μ satisfies some (strong) TCI then $T_\sharp\mu$ satisfies a (strong) TCI with a skewed cost function. This principle enables us to derive new (strong) TCIs from old ones. More precisely if μ_{ref} is a known probability measure satisfying some (strong) TCI and if one is able to construct a map T transporting μ_{ref} on an other probability measure μ, then μ will satisfy a (strong) TCI too. This principle is true in any dimension. The reason why this paper deals with dimension one only is that transportation of measures is extremely simple in this framework.

This paper is organized as follows. In Section 1, we recall some known results about transportation-cost inequalities. Since in these results the dimension plays no role, we will place ourselves in an abstract Polish setting. In Section 2, we explain the above mentioned perturbation method. In Section 3, we apply the perturbation method to characterize strong TCIs for probability measures belonging to the class $\mathcal{L}ip_\sharp\mu_1$. In Section 4, we use the preceding results to give a complete characterization of strong TCIs for Log-concave probability measures.

2. Preliminary results

In this section, we recall some well-known results on TCIs, namely their dual representation, their tensorization and their links with the concentration of measure phenomenon.

General Framework. Most of the forthcoming results are available in more general framework which we shall now describe.

Let $\mathcal{X}$ be a Polish space and let $c : \mathcal{X} \times \mathcal{X} \to \mathbb{R}^+$ be a lower semi-continuous function, called the *cost function*. The set of all the probability measures on $\mathcal{X}$ will be denoted by $\mathcal{P}(\mathcal{X})$. The optimal transportation cost between $\nu \in \mathcal{P}(\mathcal{X})$ and $\mu \in \mathcal{P}(\mathcal{X})$ is defined by

$$\mathcal{T}_c(\nu, \mu) = \inf_{\pi \in P(\nu,\mu)} \iint_{\mathcal{X}\times\mathcal{X}} c(x, y)\, \pi(dx\, dy),$$

where $P(\nu, \mu)$ is the set of all the probability measures on $\mathcal{X} \times \mathcal{X}$ such that $\pi(dx \times \mathcal{X}) = \nu$ and $\pi(\mathcal{X} \times dy) = \mu$.

A probability measure μ is said to satisfy the TCI with the cost function c if

$$\forall \nu \in \mathcal{P}(\mathcal{X}), \quad \mathcal{T}_c(\nu, \mu) \leqslant \mathrm{H}(\nu \mid \mu). \tag{8}$$

A probability measure μ is said to satisfy the strong TCI with the cost function c if

$$\forall \nu, \beta \in \mathcal{P}(\mathcal{X}), \quad \mathcal{T}_c(\nu, \beta) \leqslant \mathrm{H}(\nu \mid \mu) + \mathrm{H}(\beta \mid \mu). \tag{9}$$

2.1. TCIs vs strong TCIs

We begin with elementary facts concerning TCIs and strong TCIs. In the sequel, $\theta : \mathbb{R}^p \to \mathbb{R}^+$ will be a symmetric function, i.e. $\theta(x) = \theta(-x)$, for all $x \in \mathbb{R}^p$. If $\mu \in \mathcal{P}(\mathbb{R}^p)$ satisfies the strong TCI with the cost function $c(x, y) = \theta(a(x - y))$, one will write $\mu \in \mathbb{ST}_\theta(a)$.

Proposition 7. *Suppose that θ is convex. If $\mu \in \mathbb{T}_\theta(a)$ for some $a > 0$, then $\mu \in \mathbb{ST}_{\tilde{\theta}}(a)$, where $\tilde{\theta}(x) = 2\theta(x/2)$, for all $x \in \mathbb{R}^p$.*

Proof. (See also the proof of [21, Corollary 1.3].) Let $\pi_1 \in P(\nu, \mu)$ and $\pi_2 \in P(\mu, \beta)$. One can construct X, Y, Z, three random variables such that $\mathcal{L}(X, Y) = \pi_1$ and $\mathcal{L}(Y, Z) = \pi_2$ (see for instance [22, Gluing Lemma, p. 208]). Let $c(x, y) = \theta(a(x - y))$ and $\tilde{c}(x, y) = \tilde{\theta}(a(x - y))$. Using the convexity of θ, one has

$$\begin{aligned}\mathcal{T}_{\tilde{c}}(\nu, \beta) &\leqslant \mathbb{E}\big[2\theta\big(a(X - Z)/2\big)\big] \leqslant \mathbb{E}\big[\theta\big(a(X - Y)\big)\big] + \mathbb{E}\big[\theta\big(a(Y - Z)\big)\big] \\ &= \int c(x, y)\,\pi_1(dx\,dy) + \int c(y, z)\,\pi_2(dy\,dz).\end{aligned}$$

Optimizing in π_1 and π_2 yields

$$\mathcal{T}_{\tilde{c}}(\nu, \beta) \leqslant \mathcal{T}_c(\nu, \mu) + \mathcal{T}_c(\beta, \mu), \quad \forall \nu, \beta \in \mathcal{P}\big(\mathbb{R}^p\big).$$

Consequently, if $\mu \in \mathbb{T}_\theta(a)$, then $\mu \in \mathbb{ST}_{\tilde{\theta}}(a)$. □

Lemma 8. *Suppose that $\theta(kx) \geqslant k\theta(x)$, $\forall k \in \mathbb{N}$, $\forall x \in \mathbb{R}^p$. Let $b_1, b_2 > 0$ and define $\tilde{\theta}(x) = b_1\theta(b_2 x)$, $\forall x \in \mathbb{R}^p$. Then, $\mu \in \mathbb{ST}_\theta(a)$ for some $a > 0$ if and only if $\mu \in \mathbb{ST}_{\tilde{\theta}}(\tilde{a})$ for some $\tilde{a} > 0$.*

Proof. Suppose that $\mu \in \mathbb{ST}_\theta(a)$ for some $a > 0$. Let $k \in \mathbb{N}$ such that $k \geqslant b_1$, then $\theta(ax) \geqslant k\theta(ax/k) \geqslant \tilde{\theta}(ax/(b_2 k))$. Hence, $\mu \in \mathbb{ST}_{\tilde{\theta}}(\frac{a}{b_2 k})$. The other way is identical. □

Remark 9. Suppose that θ is a convex symmetric function such that $\theta(0) = 0$, then it is easy to check that $\theta(kx) \geqslant k\theta(x)$ for all x. According to what precedes, one thus has

$$\exists a > 0 \quad \text{such that } \mu \in \mathbb{ST}_\alpha(a) \quad \Leftrightarrow \quad \exists \tilde{a} > 0 \quad \text{such that } \mu \in \mathbb{T}_\alpha(\tilde{a}).$$

In other words, as far as convex cost functions are considered, TCIs and strong TCIs are qualitatively equivalent.

2.2. Links with the concentration of measure phenomenon

The following theorem explains how to deduce concentration of measure estimates from a strong TCI. The argument used in the proof is due to K. Marton and M. Talagrand (see [15] and [21, the proof of Corollary 1.3]).

Theorem 10. *Let* (X, d) *be a Polish space and* $c : \mathcal{X} \times \mathcal{X} \to \mathbb{R}^+$ *be a continuous cost function. Suppose that* μ *satisfies the strong TCI* (9) *with the cost function c, then*

$$\forall A \subset \mathcal{X} \text{ measurable}, \forall r \geqslant 0, \quad \mu\big(A_c^r\big) \geqslant 1 - \frac{1}{\mu(A)} e^{-r}, \tag{10}$$

where $A_c^r = \{y \in \mathcal{X}\colon \exists x \in A \text{ such that } c(x, y) \leqslant r\}$.

Proof. Let $A, B \in \mathcal{X}$ and define $\mu_A(\cdot) = \frac{\mu(\cdot \cap A)}{\mu(A)}$ and $\mu_B(\cdot) = \frac{\mu(\cdot \cap B)}{\mu(B)}$. Since μ satisfies the strong TCI, one has:

$$c(A, B) \leqslant \mathcal{T}_c(\mu_A, \mu_B) \leqslant \mathrm{H}(\mu_A \mid \mu) + \mathrm{H}(\mu_B \mid \mu) = -\log \mu(A) - \log \mu(B), \tag{11}$$

with $c(A, B) = \inf\{c(x, y)\colon x \in A,\ y \in B\}$. Now taking $B = \mathcal{X} \setminus A_c^r$ in (11) yields the desired result. $\square$

2.3. Dual representation of transportation-cost inequalities

2.3.1. Kantorovich–Rubinstein theorem and its consequences

According to the celebrated Kantorovich–Rubinstein theorem, optimal transportation costs admit a dual representation which is the following:

$$\forall \nu, \mu \in \mathcal{P}(\mathcal{X}), \quad \mathcal{T}_c(\nu, \mu) = \sup_{(\psi, \varphi) \in \Phi_c} \left\{ \int \psi \, d\nu - \int \varphi \, d\mu \right\}, \tag{12}$$

where $\Phi_c = \{(\psi, \varphi) \in B(\mathcal{X}) \times B(\mathcal{X})\colon \psi(x) - \varphi(y) \leqslant c(x, y),\ \forall x, y \in \mathcal{X}\}$ and $B(X)$ is the set of bounded measurable functions on $\mathcal{X}$ (see [12]). The dual representation (12) is in particular true if c is lower semi-continuous function defined on a Polish space $\mathcal{X}$ (see for instance [22, Theorem 1.3]).

The infimal-convolution operator Q_c is defined by

$$Q_c\varphi(x) = \inf_y \big\{\varphi(y) + c(x, y)\big\},$$

for all $\varphi \in B(\mathcal{X})$. If c is continuous, $x \mapsto Q_c\varphi(x)$ is measurable (in fact upper semi-continuous) and it is easy to see that

$$\forall \nu, \mu \in \mathcal{P}(\mathcal{X}), \quad \mathcal{T}_c(\nu, \mu) = \sup_{\varphi \in B(\mathcal{X})} \left\{ \int Q_c\varphi \, d\nu - \int \varphi \, d\mu \right\}.$$

Since optimal transportation costs admit a dual representation, it is natural to ask if TCIs and strong TCIs admit a dual translation too. The answer is given in the following theorem.

Theorem 11. *Suppose that c is a continuous cost-function on the Polish space $\mathcal{X}$.*

(i) *μ satisfies the TCI* (8) *if and only if*

$$\forall \varphi \in B(\mathcal{X}), \quad \int e^{Q_c \varphi}\, d\mu \cdot e^{-\int \varphi\, d\mu} \leqslant 1. \tag{13}$$

(ii) *μ satisfies the strong TCI* (9) *if and only if*

$$\forall \varphi \in B(\mathcal{X}), \quad \int e^{Q_c \varphi}\, d\mu \cdot \int e^{-\varphi}\, d\mu \leqslant 1. \tag{14}$$

Proof. The first point is due to S.G. Bobkov and F. Götze (see [4, the proofs of Theorem 1.3 and (1.7)]). The interested reader can also find an alternative proof of this result in [11, Corollary 1]. In this latter proof, Large Deviations Theory techniques are used. One can easily adapt the one or the other approach to derive the dual version of strong TCIs (14). This is left to the reader. □

Remark 12. As mentioned in the introduction, inequalities of the form (14) are called infimal-convolution inequalities. These inequalities were introduced by B. Maurey in [17]. A good introduction on infimal-convolution inequalities can be found in [14]. In this article, we have chosen to privilege the strong TCI (9) form, which is the primal form of (14). The reason is that we find (9) more intuitive.

2.3.2. Application: Strong TCIs and integrability

Let us detail an important application of the infimal-convolution formulation of strong TCIs.

Proposition 13. *Let c be a continuous cost function on the Polish space $\mathcal{X}$. Suppose that $\mu \in \mathcal{P}(\mathcal{X})$ satisfies the strong TCI with the cost function c. Let $A \subset \mathcal{X}$ be a measurable set and define $c(x, A) = \inf_{y \in A} c(x, y)$. One has*

$$\int e^{c(x,A)}\, d\mu(x) \cdot \mu(A) \leqslant 1. \tag{15}$$

Remark 14. This integrability property was first noticed by B. Maurey in [17]. Note that the inequality (15) implies the concentration estimate (10).

Proof. Define, for all $p \in \mathbb{N}$,

$$\varphi_A^p(x) = \begin{cases} 0 & \text{if } x \in A, \\ p & \text{if } x \in A^c. \end{cases}$$

As φ_A^p is bounded, one can apply (14), this yields

$$\int e^{Q_c \varphi_A^p}\, d\mu \cdot \int e^{-\varphi_A^p}\, d\mu \leqslant 1.$$

An easy computation shows that

$$Q_c \varphi_A^p(x) = \min\bigl(c(x, A), p\bigr) \underset{p \to +\infty}{\longrightarrow} c(x, A) \quad \text{and} \quad e^{-\varphi_A^p} \underset{p \to +\infty}{\longrightarrow} \mathbb{1}_A.$$

Using the monotone convergence theorem, one gets the desired inequality. □

The following corollary will be very useful in the sequel.

Corollary 15. *Let μ be a probability measure on $\mathbb{R}$ satisfying the strong TCI with the cost function $c(x,y)=\alpha(x-y)$, with α a continuous symmetric non-decreasing function. For all $x\in\mathbb{R}$, define*

$$\mu_x^+=\mathcal{L}(X-x\,|\,X\geqslant x)\quad\text{and}\quad \mu_x^-=\mathcal{L}(x-X\,|\,X\leqslant x),$$

where X is a random variable with law μ. Then,

$$\int_0^{+\infty} e^{\alpha}\,d\mu_x^+\leqslant \frac{1}{\mu(-\infty,x]}+1,\quad \forall x\in\mathbb{R},$$

$$\int_0^{+\infty} e^{\alpha}\,d\mu_x^-\leqslant \frac{1}{\mu[x,+\infty)}+1,\quad \forall x\in\mathbb{R}.$$

In particular,

$$\int e^{\alpha}\,d\mu\leqslant \frac{1}{\mu(\mathbb{R}^+)\mu(\mathbb{R}^-)}-1.$$

Proof. Let $A=(-\infty,x]$. It is easy to show that $c(y,A)=\alpha(y-x)$ if $y\geqslant x$ and 0 else. Applying (15) with this A yields

$$\left(\mu(-\infty,x]+\int_x^{+\infty} e^{\alpha(y-x)}\,d\mu(y)\right)\cdot\mu(-\infty,x]\leqslant 1.$$

Rearranging the terms, one gets

$$\int_x^{+\infty} e^{\alpha(y-x)}\,d\mu(y)\leqslant \frac{1-\mu(-\infty,x]^2}{\mu(-\infty,x]}.$$

Dividing both sides by $\mu[x,+\infty)$ gives the result. Working with $A=[x,+\infty)$ gives the integrability property for μ_x^-. Now,

$$\begin{aligned}\int e^{\alpha}\,d\mu&=\mu\big(\mathbb{R}^+\big)\int_0^{+\infty} e^{\alpha}\,d\mu_0^+ +\mu(\mathbb{R}^-)\int_0^{+\infty} e^{\alpha}\,d\mu_0^-\\&\leqslant 1+\frac{\mu(\mathbb{R}^+)}{\mu(\mathbb{R}^-)}+\frac{\mu(\mathbb{R}^-)}{\mu(\mathbb{R}^+)}\\&=\frac{1}{\mu(\mathbb{R}^+)\mu(\mathbb{R}^-)}-1.\qquad\square\end{aligned}$$

2.4. Tensorization property of (strong) TCIs

If μ_1 and μ_2 satisfy a (strong) TCI, does $\mu_1 \otimes \mu_2$ satisfy a (strong) TCI? The following theorem gives an answer to this question.

Theorem 16. *Let $(\mathcal{X}_i)_{i=1\ldots n}$ be a family of Polish spaces. Suppose that μ_i is a probability measure on $\mathcal{X}_i$ satisfying a (strong) TCI on $\mathcal{X}_i$ with a continuous cost function c_i such that $c_i(x,x)=0$, $\forall x \in \mathcal{X}_i$. Then the probability measure $\mu_1 \otimes \cdots \otimes \mu_n$ satisfies a (strong) TCI on $\mathcal{X}_1 \times \cdots \times \mathcal{X}_n$ with the cost function $c_1 \oplus \cdots \oplus c_n$ defined as follows:*

$$\forall x, y \in \mathcal{X}_1 \times \cdots \times \mathcal{X}_n, \quad c_1 \oplus \cdots \oplus c_n(x,y) = \sum_{i=1}^{n} c_i(x_i, y_i).$$

Proof. There are two methods to prove this tensorization property. The first one is due to K. Marton and makes use of a coupling argument (the so called *Marton's coupling argument*). It is explained in several places: in Marton's original paper [15], in Talagrand's paper on $\mathbb{T}_2$ [21] or in M. Ledoux book [14, Chapter 6]. The second method uses the dual forms (13) and (14). This approach was originally developed by B. Maurey in [17] for infimal-convolution inequalities (see [17, Lemma 1]). In the case of TCIs, the proof is given in great details in [11, Theorem 5]. $\square$

Remark 17. Several authors have obtained non-independent tensorization results for transportation cost inequalities and related inequalities (see [8,16,20]).

Applying Theorem 16 together with Theorem 10, one obtains the following corollary.

Corollary 18. *Let c be a continuous cost function on the Polish space $\mathcal{X}$ such that $c(x,x)=0$, $\forall x \in \mathcal{X}$. Suppose that $\mu \in \mathcal{P}(\mathcal{X})$ satisfies the strong TCI with the cost function c. Then,*

$$\forall n \in \mathbb{N}^*, \ \forall A \text{ measurable}, \ \forall r \geqslant 0, \quad \mu^n\big(A_c^r\big) \geqslant 1 - \frac{1}{\mu^n(A)} e^{-r},$$

where $A_c^r = \{x \in \mathcal{X}^n \colon \exists y \in A$ such that $\sum_{i=1}^n c(x_i, y_i) \leqslant r\}$.

3. The perturbation method for (strong) TCIs

3.1. The contraction principle in an abstract setting

In the sequel, $\mathcal{X}$ and $\mathcal{Y}$ will be Polish spaces. If μ is a probability measure on $\mathcal{X}$ and $T: \mathcal{X} \to \mathcal{Y}$ is a measurable map, the image of μ under T will be denoted by $T_\sharp\mu$, it is the probability measure on $\mathcal{Y}$ defined by

$$\forall A \subset \mathcal{Y} \text{ measurable}, \quad T_\sharp\mu(A) = \mu\big(T^{-1}(A)\big).$$

In this section, we will explain how a (strong) TCI is modified when the reference probability measure μ is replaced by the image $T_\sharp\mu$ of μ under some map T.

Theorem 19. *Let $T:\mathcal{X}\to\mathcal{Y}$ be a measurable bijection. If μ_{ref} satisfies the (strong) TCI with a cost function c_{ref} on $\mathcal{X}$, then $T_\sharp\mu_{\mathrm{ref}}$ satisfies the (strong) TCI with the cost function c_{ref}^T defined on $\mathcal{Y}$ by*

$$c_{\mathrm{ref}}^T(y_1,y_2)=c_{\mathrm{ref}}\big(T^{-1}y_1,T^{-1}y_2\big),\quad \forall y_1,y_2\in\mathcal{Y}.$$

In other word, $T_\sharp\mu_{\mathrm{ref}}$ satisfies the (strong) TCI with a *skewed cost function*.

Proof. Let us define $Q(y_1,y_2)=(T^{-1}y_1,T^{-1}y_2)$, $\forall y_1,y_2\in\mathcal{Y}$. Let $\nu,\beta\in\mathcal{P}(\mathcal{Y})$ and take $\pi\in P(\nu,\beta)$, then

$$\int c_{\mathrm{ref}}^T(y_1,y_2)\,d\pi=\int c(x,y)\,dQ_\sharp\pi,$$

so

$$\mathcal{T}_{c_{\mathrm{ref}}^T}(\nu,\beta)=\inf_{\pi\in Q_\sharp P(\nu,\beta)}\int c(x,y)\,d\pi.$$

But it is easily seen that $Q_\sharp P(\nu,\beta)=P(T_\sharp^{-1}\nu,T_\sharp^{-1}\beta)$. Consequently

$$\mathcal{T}_{c_{\mathrm{ref}}^T}(\nu,\beta)=\mathcal{T}_{c_{\mathrm{ref}}}\big(T_\sharp^{-1}\nu,T_\sharp^{-1}\beta\big).$$

If μ_{ref} satisfies the strong TCI with the cost function c_{ref}, then

$$\mathcal{T}_{c_{\mathrm{ref}}}\big(T_\sharp^{-1}\nu,T^{-1}\beta\big)\leqslant \mathrm{H}\big(T_\sharp^{-1}\nu\mid\mu_{\mathrm{ref}}\big)+\mathrm{H}\big(T_\sharp^{-1}\beta\mid\mu_{\mathrm{ref}}\big).$$

But

$$\mathrm{H}\big(T_\sharp^{-1}\nu\mid\mu_{\mathrm{ref}}\big)=\mathrm{H}\big(T_\sharp^{-1}\nu\mid T_\sharp^{-1}T_\sharp\mu_{\mathrm{ref}}\big)=\mathrm{H}(\nu\mid T_\sharp\mu_{\mathrm{ref}}),$$

where the last equality comes from the following classical invariance property of relative entropy: $\mathrm{H}(S_\sharp\nu_1\mid S_\sharp\nu_2)=\mathrm{H}(\nu_1\mid\nu_2)$. Hence

$$\forall\nu,\beta\in\mathcal{P}(\mathcal{Y}),\quad \mathcal{T}_{c_{\mathrm{ref}}^T}(\nu,\beta)\leqslant\mathrm{H}(\nu\mid T_\sharp\mu_{\mathrm{ref}})+\mathrm{H}(\beta\mid T_\sharp\mu_{\mathrm{ref}}).$$

The proof works in the same way for TCI. □

The corollary below explains the method we will use in the sequel to derive new (strong) TCIs from known ones.

Corollary 20 *(Contraction principle). Let μ_{ref} be a probability measure on $\mathcal{X}$ satisfying a (strong) TCI with a continuous cost function c_{ref}. In order to prove that a probability measure μ on $\mathcal{Y}$ satisfies the (strong) TCI with a continuous cost function c, it is enough to build an application $T:\mathcal{X}\to\mathcal{Y}$ such that $\mu=T_\sharp\mu_{\mathrm{ref}}$ and*

$$c(Tx_1,Tx_2)\leqslant c_{\mathrm{ref}}(x_1,x_2),\quad\forall x_1,x_2\in\mathcal{X}.$$

This contraction property of strong TCIs (written in their infimal-convolution form) was first observed by B. Maurey (see [17, Lemma 2]).

Proof. We assume that μ_{ref} satisfies the strong TCI with the cost function c_{ref}. Let $\varphi : \mathcal{Y} \to \mathbb{R}$ be a bounded map. Then, for all $x_1 \in \mathcal{X}$

$$\begin{aligned} Q_c\varphi(Tx_1) &= \inf_{y\in\mathcal{Y}}\left\{\varphi(y) + c(Tx_1, y)\right\} \leqslant \inf_{x_2\in\mathcal{X}}\left\{\varphi(Tx_2) + c(Tx_1, Tx_2)\right\} \\ &\leqslant \inf_{x_2\in\mathcal{X}}\left\{\varphi\circ T(x_2) + c_{\mathrm{ref}}(x_1, x_2)\right\} = Q_{c_{\mathrm{ref}}}(\varphi\circ T). \end{aligned}$$

Thus,

$$\begin{aligned} \int e^{Q_c\varphi}\,d\mu \cdot \int e^{-\varphi}\,d\mu &= \int e^{Q_c\varphi}\circ T\,d\mu_{\mathrm{ref}} \cdot \int e^{-\varphi\circ T}\,d\mu_{\mathrm{ref}} \\ &\leqslant \int e^{Q_{c_{\mathrm{ref}}}(\varphi\circ T)}\,d\mu_{\mathrm{ref}} \cdot \int e^{-\varphi\circ T}\,d\mu_{\mathrm{ref}} \\ &\leqslant 1, \end{aligned}$$

where the last inequality follows from (14). $\square$

Remark 21. If T is invertible, the proof above can be simplified using Theorem 19. Namely, according to Theorem 19, μ satisfies the (strong) TCI with the cost function c^T_{ref}. But, by hypothesis, $c \leqslant c^T_{\mathrm{ref}}$, so μ satisfies the (strong) TCI with the cost function c.

3.2. The contraction principle on the real line

3.2.1. Monotone rearrangement

We are going to apply the contraction principle to probability measures on the real line. The reason why dimension one is so easy to handle is the existence of an explicit and computable map T which pushes forward μ_{ref} on μ: the monotone rearrangement.

Theorem 22 *(Monotone rearrangement). Let μ_{ref} and μ be probability measures on $\mathbb{R}$ and let F_{ref} and F denote their cumulative distribution functions*:

$$F_{\mathrm{ref}}(t) = \mu_{\mathrm{ref}}(-\infty, t], \quad \forall t \in \mathbb{R}, \quad \textit{and} \quad F(t) = \mu(-\infty, t], \quad \forall t \in \mathbb{R}.$$

If F_{ref} and F are continuous and increasing (*equivalently μ_{ref} and μ have no atom and full support*), *then the map $T = F^{-1} \circ F_{\mathrm{ref}}$ transports μ_{ref} on μ, that is $T_\sharp\mu_{\mathrm{ref}} = \mu$.*

The proof of this theorem is elementary. From now on, T will always be the map defined in the preceding theorem.

3.2.2. About the exponential distribution

The reference probability measure μ_{ref} will be the symmetric exponential distribution μ_1 on $\mathbb{R}$:

$$d\mu_{\mathrm{ref}}(x) = d\mu_1(x) := \frac{1}{2}e^{-|x|}\,dx.$$

Theorem 23 *(Maurey, Talagrand). The exponential measure μ_1 satisfies the (strong) TCI with the cost function $\frac{1}{\kappa}c_1$, for some constant $\kappa > 0$, with c_1 defined by*

$$c_1(x,y) := \alpha_1(x-y), \quad \forall x,y \in \mathbb{R}, \quad \textit{where} \quad \alpha_1(t) = \min\bigl(|t|, t^2\bigr), \quad \forall t \in \mathbb{R}.$$

Remark 24.

(1) One can take $\kappa = 36$.
(2) B. Maurey proved the strong TCI with the sharper cost functions $c(x,y) = \tilde{\alpha}_1(x-y)$, where $\tilde{\alpha}_1(x) = 1/36x^2$ if $|x| \leqslant 4$ and $2/9(|x|-2)$ otherwise (see [17, Proposition 1]). One can show that $\tilde{\alpha}_1 \geqslant 1/36\alpha_1$.
(3) M. Talagrand proved independently that μ_1 satisfies the TCI with the cost functions $c_\lambda(x,y) = \gamma_\lambda(x-y)$ where $\gamma_\lambda(x) = (\frac{1}{\lambda} - 1)(e^{-\lambda|x|} - 1 + \lambda|x|)$ for all $\lambda \in (0,1)$ (see [21, Theorem 1.2]).

Transportation-cost inequalities associated to the cost function c_1 were fully characterized by S. Bobkov, I. Gentil and M. Ledoux in [3] in terms of Poincaré inequalities:

Theorem 25 *(Bobkov–Gentil–Ledoux). A probability measure μ on $\mathbb{R}^p$ satisfies the TCI with the cost function $(x,y) \mapsto \lambda \min(|x-y|_2, |x-y|_2^2)$, for some $\lambda > 0$ (where $|\cdot|_2$ is the Euclidean norm on $\mathbb{R}^p$) if and only if it satisfies a Poincaré inequality, that is if there is some constant $C > 0$ such that*

$$\mathrm{Var}_\mu(f) \leqslant C \int_{\mathbb{R}^p} |\nabla f|_2^2\,d\mu, \quad \forall f. \tag{16}$$

3.2.3. Application of the contraction principle on the real line

A nice feature of the exponential distribution is that its cumulative distribution function can be explicitly computed:

$$F_1(x) = \begin{cases} 1 - \frac{1}{2}e^{-|x|} & \text{if } x \geqslant 0, \\ \frac{1}{2}e^{-|x|} & \text{if } x \leqslant 0 \end{cases} \quad \text{and} \quad F_1^{-1}(t) = \begin{cases} -\log(2(1-t)) & \text{if } t \geqslant \frac{1}{2}, \\ \log(2t) & \text{if } t \leqslant \frac{1}{2}. \end{cases} \tag{17}$$

Suppose that μ is a probability measure on $\mathbb{R}$ having no atom and full support, then its cumulative distribution function F is invertible, and the map T transporting μ_1 on μ can be expressed as follows:

$$T(x) = \begin{cases} F^{-1}(1 - \frac{1}{2}e^{-|x|}) & \text{if } x \geqslant 0, \\ F^{-1}(\frac{1}{2}e^{-|x|}) & \text{if } x \leqslant 0 \end{cases} \quad \text{and}$$

$$T^{-1}(x) = \begin{cases} -\log(2(1-F(x))) & \text{if } x \geqslant m, \\ \log(2F(x)) & \text{if } x \leqslant m, \end{cases} \tag{18}$$

where m denotes the median of μ.

Let us introduce the following quantity:

$$\omega_\mu(h) = \inf\{ |T^{-1}x - T^{-1}y| \colon |x-y| \geqslant h \}, \quad \forall h \geqslant 0.$$

Proposition 26. *If $\mu \in \mathcal{P}(\mathbb{R})$ is a probability measure with no atom and full support, then μ satisfies the strong TCI with the cost function $c_\mu(x,y) = \frac{1}{\kappa}\alpha_1 \circ \omega_\mu(|x-y|)$, where $\alpha_1(t) = \min(t, t^2)$, $\forall t \geqslant 0$.*

Proof. By definition of ω_μ, $|T^{-1}x - T^{-1}y| \geqslant \omega_\mu(|x-y|)$, for all $x, y \in \mathbb{R}$. Thus, $c_1^T(x,y) \geqslant \alpha_1(\omega_\mu(|x-y|))$, for all $x, y \in \mathbb{R}$. This achieves the proof by Corollary 20 and Theorem 23. □

To better understand ω_μ it is good to relate it to the continuity modulus of T.

Definition 27 *(The class $\mathcal{UC}_\sharp\mu_1$).* The set of all probability measures μ on $\mathbb{R}$, with no atom and full support, such that the monotone rearrangement map transporting the exponential measure $d\mu_1(x) = \frac{1}{2}e^{-|x|}\,dx$ on μ is uniformly continuous is denoted by $\mathcal{UC}_\sharp\mu_1$.

The proof of the following proposition is left to the reader.

Proposition 28. *Suppose $\mu \in \mathcal{UC}_\sharp\mu_1$, then the continuity modulus Δ_μ of T is defined by $\Delta_\mu(h) = \sup\{|Tx - Ty| \colon |x-y| \leqslant h\}$, $\forall h \geqslant 0$. It is a continuous increasing function and*

$$\omega_\mu = \Delta_\mu^{-1}.$$

Remark 29.

(1) All the elements of $\mathcal{UC}_\sharp\mu_1$ enjoy a dimension-free concentration of measure property. Namely, if $\mu \in \mathcal{UC}_\sharp\mu_1$, then μ satisfies the strong TCI with the cost function $c_\mu(x,y) = \alpha_\mu(x-y)$, where $\alpha_\mu(x) = \frac{1}{\kappa}\alpha_1 \circ \omega_\mu(|x|)$. Thus according to Corollary 18, one has

$$\forall n \in \mathbb{N}^*,\ \forall A \subset \mathbb{R}^n,\ \forall r \geqslant 0, \quad \mu^n(A^r_{\alpha_\mu}) \geqslant 1 - \frac{1}{\mu^n(A)}e^{-r},$$

with $A^r_{\alpha_\mu} = \{x \in \mathbb{R}^n \colon \exists y \in A \text{ such that } \sum_{i=1}^n \alpha_\mu(x_i - y_i) \leqslant r\}$.

(2) The class of all the probability measures on $\mathbb{R}$ satisfying a dimension-free concentration of measure property is not yet identified. In [6], S.G. Bobkov and C. Houdré studied probability measures enjoying a weak dimension-free concentration property (roughly speaking one can estimate $\mu^n(A^r_\infty)$ independently of the dimension, where A^r_∞ denotes the blow-up of A with respect to the norm $|x|_\infty = \max_i |x_i|$). They proved that a probability measure has this weak property if and only if the map T generate a finite modulus, which means that $\Delta_\mu(h) < +\infty$ for some (equivalently for all) $h \in \mathbb{R}$.

In order to obtain explicit concentration properties, one has to estimate ω_μ.

Proposition 30. *Define*

$$\omega_\mu^+(h) = \inf\{|T^{-1}x - T^{-1}y|:\ |x-y| \geqslant h,\ x, y \geqslant m\},$$
$$\omega_\mu^-(h) = \inf\{|T^{-1}x - T^{-1}y|:\ |x-y| \geqslant h,\ x, y \leqslant m\}$$

then

$$\omega_\mu(h) \geqslant \min\left(\omega_\mu^+\left(\frac{h}{2}\right), \omega_\mu^-\left(\frac{h}{2}\right)\right).$$

Proof. Let $x, y \in \mathbb{R}$ with $x \leqslant m \leqslant y$ and $y - x \geqslant h \geqslant 0$. One has

$$\begin{aligned}|T^{-1}y - T^{-1}x| &= T^{-1}y - T^{-1}x = (T^{-1}y - T^{-1}m) + (T^{-1}m - T^{-1}x)\\ &\geqslant \omega_\mu^+(y-m) + \omega_\mu^-(m-x).\end{aligned}$$

Since $y - x \geqslant h$ and $m \in [x, y]$, one has $y - m \geqslant \frac{h}{2}$ or $m - x \geqslant \frac{h}{2}$, thus

$$|T^{-1}y - T^{-1}x| \geqslant \min\left(\omega_\mu^+\left(\frac{h}{2}\right), \omega_\mu^-\left(\frac{h}{2}\right)\right). \qquad \square$$

Let X be a random variable with law μ and define

$$\mu_x^+ = \mathcal{L}(X - x \mid X \geqslant x) \in \mathcal{P}(\mathbb{R}^+), \quad \forall x \geqslant m, \tag{19}$$
$$\mu_x^- = \mathcal{L}(x - X \mid X \leqslant x) \in \mathcal{P}(\mathbb{R}^+), \quad \forall x \leqslant m. \tag{20}$$

In the following proposition, the quantities ω_μ^+ and ω_μ^- are expressed in terms of the cumulative distribution functions of the probability measures μ_x^+ and μ_x^-.

Proposition 31.

$$\omega_\mu^+(h) = \inf\{-\log \mu_x^+[h, +\infty):\ x \geqslant m\},$$
$$\omega_\mu^-(h) = \inf\{-\log \mu_x^-[h, +\infty):\ x \leqslant m\}, \quad \forall h \geqslant 0.$$

Proof. It is easy to see that $\omega_\mu^+(h) = \inf\{T^{-1}(x+h) - T^{-1}(x):\ x \geqslant m\}$. Using (18) one sees that

$$T^{-1}(x+h) - T^{-1}(x) = -\log\left(\frac{1 - F(x+h)}{1 - F(x)}\right) = -\log \mu_x^+[h, +\infty),$$

which gives the result for $\omega_\mu^+(h)$. The proof is identical for $\omega_\mu^-(h)$. $\square$

The proof of the following corollary is immediate.

Corollary 32. *Let $\omega : \mathbb{R}^+ \to \mathbb{R}^+$ a continuous non-decreasing function with $\omega(0) = 0$. In order to show that $c_\mu(x, y) \geqslant \frac{1}{\kappa}\alpha_1 \circ \omega(\frac{|x-y|}{2})$, it is enough to show that*

$$\sup_{x \geqslant m} \mu_x^+[h, +\infty) \leqslant e^{-\omega(h)}, \quad \forall h \geqslant 0, \tag{21}$$

$$\sup_{x \leqslant m} \mu_x^-[h, +\infty) \leqslant e^{-\omega(h)}, \quad \forall h \geqslant 0. \tag{22}$$

4. Characterization of strong TCIs on $\mathcal{L}ip_\sharp\mu_1$

In this section, we give a characterization of strong TCIs for probability measures belonging to a certain class $\mathcal{L}ip_\sharp\mu_1$ which we shall now define.

4.1. The Lipschitz images of the exponential measure

4.1.1. Description of $\mathcal{L}ip_\sharp\mu_1$

Definition 33 *(The class $\mathcal{L}ip_\sharp\mu_1$).* The set of all probability measures μ on $\mathbb{R}$, with no atom and full support, such that the monotone rearrangement map transporting the exponential measure $d\mu_1(x) = \frac{1}{2}e^{-|x|}\,dx$ on μ is Lipschitz is denoted by $\mathcal{L}ip_\sharp\mu_1$.

The following proposition describes the elements of $\mathcal{L}ip_\sharp\mu_1$.

Proposition 34. *Let $\mu \in \mathcal{P}(\mathbb{R})$ with no atom and full support and let T be the monotone rearrangement transporting μ_1 on μ. The map T is $1/a$-Lipschitz if and only if one has*

$$\sup_{x \geqslant m} \mu_x^+[h, +\infty) \leqslant e^{-ah}, \quad \forall h \geqslant 0, \tag{23}$$

$$\sup_{x \leqslant m} \mu_x^-[h, +\infty) \leqslant e^{-ah}, \quad \forall h \geqslant 0. \tag{24}$$

If μ is of the form $d\mu(z) = e^{-V(z)}\,dz$ where V is a continuous function, then T is $1/a$-Lipschitz if and only if

$$A^+ := \sup_{x \geqslant m} \bigl(1 - F(x)\bigr)e^{V(x)} \leqslant \frac{1}{a} \quad \text{and} \quad A^- := \sup_{x \leqslant m} F(x)e^{V(x)} \leqslant \frac{1}{a}. \tag{25}$$

Furthermore, if V is of class $\mathcal{C}^1$, a sufficient condition for A^+ and A^- to be finite is that

$$\liminf_{x \to +\infty} V' > 0 \quad \text{and} \quad \limsup_{x \to -\infty} V' < 0. \tag{26}$$

Proof. It is easy to see that the map T is $1/a$-Lipschitz if and only if

$$T^{-1}z - T^{-1}y \geqslant a(z - y), \quad \forall z \geqslant y. \tag{27}$$

This is equivalent to

$$T^{-1}(x+h) - T^{-1}x \geqslant ah, \quad \forall x \geqslant m,\ \forall h \geqslant 0 \quad \text{and}$$
$$T^{-1}x - T^{-1}(x-h) \geqslant ah, \quad \forall x \leqslant m,\ \forall h \geqslant 0.$$

Using the fact that

$$T^{-1}(z) = \begin{cases} -\log(2(1-F(z))) & \text{if } z \geqslant m, \\ \log(2F(z)) & \text{if } z \leqslant m, \end{cases}$$

one sees immediately that these conditions are equivalent to (23) and (24).

If $d\mu(z) = e^{-V(z)}\,dz$ with a continuous V, T^{-1} is differentiable. Observe that (27) means that $z \mapsto T^{-1}z - az$ is non-decreasing and this is equivalent to $\sup_{z\in\mathbb{R}} dT^{-1}/dz(z) \leqslant a$. Computing dT^{-1}/dz, one obtains immediately (25).

Finally, let us show that the condition $\liminf_{x\to+\infty} V' > 0$ implies that A^+ is finite. Under this assumption, there is $v_0 > 0$ and $z_0 > m$ such that for all $z \geqslant z_0$, one has $V'(z) \geqslant v_0$. If $z \geqslant z_0$, one thus has

$$e^{-V(z)} = \int_z^{+\infty} V'(y)e^{-V(y)}\,dy \geqslant v_0 \int_z^{+\infty} e^{-V(y)}\,dy = v_0\big(1-F(y)\big).$$

So, $\sup_{z\geqslant z_0}(1-F(z))e^{V(z)} \leqslant \frac{1}{v_0}$. Since $\sup_{m\leqslant z\leqslant z_0}(1-F(z))e^{V(z)} < +\infty$, one concludes that $A^+ < +\infty$. The same reasoning shows that the condition $\limsup_{z\to-\infty} V' < 0$ implies $A^- < +\infty$. $\square$

4.1.2. The class $\mathcal{L}ip_\sharp\mu_1$ and the Poincaré inequality

According to Corollary 20, one concludes that a sufficient condition for a probability measure to satisfy the inequality $\mathbb{ST}_{\alpha_1}(a)$ for some constant a is that μ belongs to $\mathcal{L}ip_\sharp\mu_1$. But, according to Theorem 25, one knows that $\mu \in \mathbb{ST}_{\alpha_1}(a)$ for some $a > 0$ if and only if μ satisfies the Poincaré inequality (16) for some constant $C > 0$. For a large class of probability measures, μ satisfies Poincaré if and only if $\mu \in \mathcal{L}ip_\sharp\mu_1$. This is explained in the next proposition.

We refer to the introduction for the definition of the class $\mathcal{V}$ of 'good' potentials.

Proposition 35. *Let $d\mu = e^{-V}\,dx$ with $V \in \mathcal{V}$, then*

$$\mu \text{ satisfies Poincaré} \quad \Leftrightarrow \quad \liminf_{x\to+\infty} V'(x) > 0 \quad \text{and} \quad \limsup_{x\to-\infty} V'(x) < 0 \quad \Leftrightarrow \quad \mu \in \mathcal{L}ip_\sharp\mu_1.$$

Proof. According to the celebrated Muckenhoupt criterion (see [18] and [1, Theorems 6.2.1 and 6.2.2]), a probability measure $d\mu = h\,dx$ with a continuous h satisfies (16) for some constant $C > 0$ if and only if

$$D^+ := \sup_{x\geqslant m}\big(1-F(x)\big)\cdot\int_m^x \frac{1}{h(y)}\,dy < +\infty \quad \text{and} \quad D^- := \sup_{x\leqslant m} F(x)\cdot\int_x^m \frac{1}{h(y)}\,dy < +\infty,$$

m being the median of μ. Applying Proposition 40, one shows that

$$\bigl(1 - F(x)\bigr) \cdot \int_m^x e^V(y)\,dy \sim_{x\to+\infty} \frac{1}{V'^2(x)} \quad \text{and} \quad \bigl(1 - F(x)\bigr) \cdot e^{V(x)} \sim_{x\to+\infty} \frac{1}{V'(x)}.$$

From this one easily concludes that A^+ and D^+ are finite if and only if

$$\liminf_{x\to+\infty} V'(x) > 0. \qquad \square$$

In fact, there are probability measures satisfying Poincaré inequality which do not belong to $\mathcal{L}ip_\sharp\mu_1$. For example, it is not difficult to see that the probability measure $d\nu_r(x) = \frac{1}{Z_r}|x|^r e^{-|x|}\,dx$ with $r \in (0,1)$ verifies Muckenhoupt criterion. But since the density vanishes at zero, one gets $T'(0) = +\infty$ and so T is not Lipschitz.

We end this section with a result of S.G. Bobkov and C. Houdré which makes the link between $\mathcal{L}ip_\sharp\mu_1$ and L_1-Poincaré type inequalities (see [5, Theorem 1.2]).

Theorem 36 *(Bobkov–Houdré). A probability measure $\mu \in \mathcal{P}(\mathbb{R})$ belongs to $\mathcal{L}ip_\sharp\mu_1$ if and only if there is $C > 0$ such that*

$$\forall f, \quad \int \bigl|f(x) - m(f)\bigr|\,d\mu(x) \leqslant C \int \bigl|f'(x)\bigr|\,d\mu(x), \tag{28}$$

where $m(f)$ is the median of f under μ.

The L_1-Poincaré type inequality (28) has a dimension-free tensorization property and is equivalent to Cheeger type isoperimetric inequality (see [5]).

4.2. *Characterization of strong TCI on $\mathcal{L}ip_\sharp\mu_1$*

Proof of Theorem 2. Assume first that μ satisfies $\mathbb{ST}_\alpha(a)$ for some $a > 0$. Then according to Corollary 15, one has

$$\int e^{\alpha(az)}\,d\mu_x^+(z) \leqslant \frac{1}{\mu(-\infty, x]} + 1 \leqslant 3, \quad \forall x \geqslant m,$$
$$\int e^{\alpha(az)}\,d\mu_x^-(z) \leqslant \frac{1}{\mu[x, +\infty)} + 1 \leqslant 3, \quad \forall x \leqslant m.$$

Thus $K^+(a) < +\infty$ and $K^-(a) < +\infty$.

Now let us assume that $K^+(b_0) < +\infty$ and $K^-(b_0) < +\infty$ for some $b_0 > 0$ and let us prove that μ satisfies $\mathbb{ST}_\alpha(a)$ for some $a > 0$. According to Corollary 32, if there is some $a > 0$ such that

$$\sup_{x\geqslant m} \mu_x^+[h, +\infty) \leqslant e^{-\alpha_1^{-1}\circ\alpha(ah)}, \quad \forall h \geqslant 0, \tag{29}$$

$$\sup_{x\leqslant m} \mu_x^-[h, +\infty) \leqslant e^{-\alpha_1^{-1}\circ\alpha(ah)}, \quad \forall h \geqslant 0, \tag{30}$$

then μ satisfies the strong TCI with the cost function $\frac{1}{\kappa}\alpha(a|x-y|/2) \geqslant \alpha(a|x-y|/(2\kappa))$. Hence it is enough to prove (29) and (30).

Let us prove (29) (the proof of (30) will be the same). To prove (29), it is enough to find $a > 0$ such that

$$\begin{cases} \mu_x^+[h,+\infty) \leqslant e^{-ah}, & \forall h \leqslant \frac{1}{a}, \\ \mu_x^+[h,+\infty) \leqslant e^{-\alpha(ah)}, & \forall h \geqslant \frac{1}{a}, \end{cases} \tag{31}$$

holds for all $x \geqslant m$. Since $\mu \in \mathcal{L}ip_\sharp\mu_1$, there is $a_0 > 0$ such that

$$\sup_{x \geqslant m} \mu_x^+[h,+\infty) \leqslant e^{-a_0 h}, \quad \forall h \geqslant 0.$$

By hypothesis $K^+(b_0) := \sup_{x\geqslant m} \int e^{\alpha(b_0 z)}\, d\mu_x^+(z) < +\infty$. Using Markov's inequality, one gets

$$K^+(b_0) \geqslant e^{\alpha(b_0 h)} \mu_x^+[h,+\infty), \quad \forall h \geqslant 0.$$

Thus, using the super-additivity of α, one has

$$\mu_x^+[h,+\infty) \leqslant K^+(b_0)e^{-\alpha(b_0 h)} \leqslant \left[K^+(b_0)e^{-\alpha(b_0 h/2)}\right]e^{-\alpha(b_0 h/2)} \leqslant e^{-\alpha(b_0 h/2)},$$

as soon as $h \geqslant \frac{2}{b_0}\alpha^{-1}(\log K^+(b_0))$. It is now easy to check that (31) holds with

$$a = \min\left(a_0, b_0/2, \left[\frac{2}{b_0}\alpha^{-1}\left(\log K^+(b_0)\right)\right]^{-1}\right). \qquad \square$$

In the following proposition, we derive from Theorem 2 a perturbation result.

Proposition 37. *Let $\mu = e^{-V}\,dx$, with V continuous, be in $\mathcal{L}ip_\sharp\mu_1$. If μ satisfies $\mathbb{ST}_\alpha(a)$ for some $a > 0$, then for every bounded and continuous function $h : \mathbb{R} \to \mathbb{R}$, the probability measure $\tilde{\mu} = \frac{1}{Z}e^h \cdot \mu$ satisfies $\mathbb{ST}_\alpha(\tilde{a})$ for some $\tilde{a} > 0$.*

Proof. Using obvious notations, one observes that

$$\sup_{x\geqslant m} (1-\widetilde{F})Ze^{V-h} \leqslant e^{\operatorname{Osc}(h)} \sup_{x\geqslant m}(1-F)e^V$$

and that

$$\sup_{x\geqslant m} \int e^{\alpha(bu)}\, d\tilde{\mu}_x^+(du) \leqslant e^{\operatorname{Osc}(h)} \sup_{x\geqslant m} \int e^{\alpha(bu)}\, d\mu_x^+(du),$$

where $\operatorname{Osc}(h) = \sup(h) - \inf(h)$ is the *oscillation* of h. Applying Proposition 34 and Theorem 2, one concludes that $\tilde{\mu}$ satisfies $\mathbb{ST}_\alpha(\tilde{a})$, for some $\tilde{a} > 0$. $\square$

Remark 38. We do not know how to prove this result directly. The same perturbation property holds for Logarithmic Sobolev or Poincaré inequalities without restriction, and in this case, the proof is completely immediate.

4.3. Tractable sufficient condition for good potentials

Theorem 39. *Let $d\mu = e^{-V}\,dx$ with $V \in \mathcal{V}$ and $\alpha \in \mathcal{A} \cap \mathcal{V}$. If $\mu \in \mathcal{L}ip_\sharp \mu_1$ and if*

$$\exists \lambda > 0 \text{ such that } \limsup_{u \to \pm\infty} \frac{\alpha'(\lambda u)}{V'(u+m)} < +\infty, \tag{32}$$

where m is the median of μ, then μ satisfies the inequality $\mathbb{ST}_\alpha(a)$ for some $a > 0$.

To prove this theorem, we will use the following lemma.

Lemma 40. *Let $\Phi \in \mathcal{V}$, then*

$$\int_x^{+\infty} e^{-\Phi(t)}\,dt \sim \frac{e^{-\Phi(x)}}{\Phi'(x)} \quad \text{and} \quad \int_0^x e^{-\Phi(t)}\,dt \sim \frac{e^{\Phi(x)}}{\Phi'(x)}, \quad \text{as } x \text{ goes to } +\infty.$$

Proof. See [1, Corollary 6.4.2]. □

Proof of Theorem 39. Let $\tilde{\mu} = \mathcal{L}(X - m)$, where X is a random variable with law μ. The density of $\tilde{\mu}$ with respect to Lebesgues measure is $e^{-\widetilde{V}}$, with $\widetilde{V}(x) = V(x+m), \forall x \in \mathbb{R}$. As $x \mapsto x + m$ is 1-Lipschitz, it follows from Corollary 20 that μ satisfies $\mathbb{ST}_\alpha(a)$ if and only if $\tilde{\mu}$ satisfies $\mathbb{ST}_\alpha(a)$. Observe that $\tilde{\mu} \in \mathcal{L}ip_\sharp \mu_1$. According to Theorem 2, to prove that $\tilde{\mu}$ satisfies $\mathbb{ST}_\alpha(a)$ for some $a > 0$, it suffices to prove that there is $b > 0$ such that

$$K^+(b) = \sup_{x \geqslant 0} \int e^{\alpha(bu)}\,d\tilde{\mu}_x^+(u) < +\infty \quad \text{and} \quad K^-(b) = \sup_{x \leqslant 0} \int e^{\alpha(bu)}\,d\tilde{\mu}_x^-(u) < +\infty,$$

where $\tilde{\mu}_x^+ = \mathcal{L}(\widetilde{X} - x \mid \widetilde{X} \geqslant x)$ and $\tilde{\mu}_x^- = \mathcal{L}(x - \widetilde{X} \mid \widetilde{X} \leqslant x)$ with $\widetilde{X}$ of law $\tilde{\mu}$.

The proof of $K^-(b) < +\infty$ being similar, we will only prove that $K^+(b) < +\infty$ for some $b > 0$. One can suppose without restriction that $\lambda = 1$ in (32). Define

$$\begin{aligned} K(b,x) &= \int e^{\alpha(bt)}\,d\tilde{\mu}_x^+(t) \\ &= \frac{\int_x^{+\infty} e^{\alpha(b(u-x))} e^{-\widetilde{V}(u)}\,du}{\int_x^{+\infty} e^{-\widetilde{V}(u)}\,du}, \quad \forall x \geqslant 0,\ \forall b \geqslant 0. \end{aligned}$$

Let us show that there is $k \in \mathbb{N}^*$ such that $K(k^{-1}, x) < +\infty$ for all $x \geqslant 0$. Since α is super-additive and non-decreasing, one gets

$$K\left(k^{-1}, x\right) \leqslant \frac{\int_0^{+\infty} e^{k^{-1}\alpha(u)} e^{-\widetilde{V}(u)}\,du}{\int_x^{+\infty} e^{-\widetilde{V}(u)}\,du}.$$

Since $\limsup_{u\to+\infty}\frac{\alpha'(u)}{\widetilde{V}'(u)}<+\infty$, there are $M>0$ and $u_*>0$ such that $\widetilde{V}'(u)\geqslant 2M\alpha'(u)$, for all $u\geqslant u_*$. Integrating yields

$$\widetilde{V}(u)\geqslant 2M\alpha(u)+C,\quad \forall u\geqslant u_*,$$

where C is a constant. Let k_* be a positive integer such that $k_*\geqslant\frac{1}{M}$. Then one has

$$e^{k_*^{-1}\alpha(u)-\widetilde{V}(u)}\leqslant e^{-M\alpha(u)-C}\leqslant e^{-M(u-1)-C},\quad \forall u\geqslant u_*,$$

where the last inequality follows from the inequality $\alpha(u)\geqslant\alpha(1)(u-1)=u-1, \forall u\geqslant 0$ which is easy to establish using the fact that $\alpha\in\mathcal{A}$. From this follows easily that $K(k_*^{-1},x)<+\infty$ for all $x\geqslant 0$.

Now, let us show that $\sup_{x\geqslant 0}K(k_*^{-1},x)<+\infty$. Since the map $x\mapsto K(k_*^{-1},x)$ is continuous, it suffices to check that $\limsup_{x\to+\infty}K(k_*^{-1},x)<+\infty$. Using the super-additivity of α, one gets

$$\alpha(u-x)\leqslant\alpha(u)-\alpha(x),\quad \forall u\geqslant x\geqslant 0.$$

So

$$K\left(k_*^{-1},x\right)\leqslant e^{-k_*^{-1}\alpha(x)}\frac{\int_x^{+\infty}e^{k_*^{-1}\alpha(u)-\widetilde{V}(u)}\,du}{\int_x^{+\infty}e^{-\widetilde{V}(u)}\,du}.$$

Applying Lemma 40, with $\Phi=\widetilde{V}-k_*^{-1}\alpha$, and then with $\Phi=\widetilde{V}$, one gets

$$e^{-k_*^{-1}\alpha(x)}\frac{\int_x^{+\infty}e^{k_*^{-1}\alpha(u)-\widetilde{V}(u)}\,du}{\int_x^{+\infty}e^{-\widetilde{V}(u)}\,du}\sim e^{-k_*^{-1}\alpha(x)}\frac{e^{k_*^{-1}\alpha(x)-\widetilde{V}(x)}}{\widetilde{V}'(x)-k_*^{-1}\alpha'(x)}e^{\widetilde{V}(x)}\widetilde{V}'(x)=\frac{1}{1-k_*^{-1}\frac{\alpha'(x)}{\widetilde{V}'(x)}}.$$

Since

$$\limsup_{x\to+\infty}\frac{1}{1-k_*^{-1}\frac{\alpha'(x)}{\widetilde{V}'(x)}}<+\infty,$$

one deduces that $\limsup_{x\to+\infty}K(k_*^{-1},x)<+\infty$, which ends the proof. □

5. The particular case of Log-concave distributions

5.1. Proof of Theorem 6

To prove Theorem 6, one needs to recall the notion of stochastic domination.

Definition 41 *(Stochastic domination).* Let $\nu_1,\nu_2\in\mathcal{P}(\mathbb{R}^+)$; one says that ν_1 is stochastically dominated by ν_2, and one writes $\nu_1\prec_{\mathrm{st}}\nu_2$ if $\nu_1(h,+\infty)\leqslant\nu_2(h,+\infty)$, for all $h\geqslant 0$.

The following proposition is well known.

Proposition 42. *Let $\nu_1, \nu_2 \in \mathcal{P}(\mathbb{R}^+)$. Then $\nu_1 \prec_{\mathrm{st}} \nu_2$ if and only if $\int f\, d\nu_1 \leqslant \int f\, d\nu_2$, for all non-decreasing $f : \mathbb{R}^+ \to \mathbb{R}^+$.*

Proof. For all $f : \mathbb{R}^+ \to \mathbb{R}^+$, one has $\int f\, d\nu_1 = \int_0^{+\infty} \nu_1(f > t)\, dt$. If f is non-decreasing and $\nu_1 \prec_{\mathrm{st}} \nu_2$, one thus has $\nu_1(f > t) \leqslant \nu_2(f > t)$ for almost all t and so $\int f\, d\nu_1 \leqslant \int f\, d\nu_2$. □

Log-concave distributions are examples of NBU distributions ("New Better than Used"), this is explained bellow.

Proposition 43. *If $\mu \in \mathcal{P}(\mathbb{R})$ is a Log-concave distribution, then*

$$\mu_x^+ \prec_{\mathrm{st}} \mu_m^+, \quad \forall x \geqslant m \quad \textit{and} \quad \mu_x^- \prec_{\mathrm{st}} \mu_m^-, \quad \forall x \leqslant m.$$

Proof. Let us show that $\mu_x^+ \prec_{\mathrm{st}} \mu_m^+$ for all $x \geqslant m$. By definition, this means that $\mu_x^+(h, +\infty) \leqslant \mu_m^+(h, +\infty)$, $\forall h \geqslant 0$, $\forall x \geqslant m$ and this is equivalent to

$$\frac{1 - F(x+h)}{1 - F(x)} \leqslant \frac{1 - F(m+h)}{1 - F(m)}, \quad \forall h \geqslant 0,\ \forall x \geqslant m.$$

Defining $\overline{F}_m^+(h) = \frac{1-F(m+h)}{1-F(m)}$, $\forall h \geqslant 0$, the preceding inequality is equivalent to

$$\overline{F}_m^+(x - m + h) \leqslant \overline{F}_m^+(x - m) \times \overline{F}_m^+(h), \quad \forall h \geqslant 0,\ \forall x \geqslant m.$$

In other word, $\mu_x^+ \prec_{\mathrm{st}} \mu_m^+$ if and only if the function $\log \overline{F}_m^+$ is sub-additive. Since μ is Log-concave, the function $\log \overline{F}_m^+$ is concave. It is easy to check that every concave function defined on $\mathbb{R}^+$ and vanishing at 0 is sub-additive. This achieves the proof. □

Now let us show that Log-concave distributions belong to $\mathcal{L}ip_\sharp \mu_1$.

Proposition 44. *Log-concave distributions on $\mathbb{R}$ belong to $\mathcal{L}ip_\sharp \mu_1$.*

Proof. According to Proposition 34, it is enough to show that (23) and (24) hold. But using Proposition 43, one just has to check that there is some $a > 0$ such that $\mu_m^+(h, +\infty) \leqslant e^{-ah}$ and $\mu_m^-(h, +\infty) \leqslant e^{-ah}$, for all $h \geqslant 0$. Let us prove this for μ_m^+. Let $\varphi = \log(1 - F)$. The function φ is concave, so

$$\varphi(m + h) \leqslant \varphi(m) + \varphi_r'(m)h, \quad \forall h \geqslant 0,$$

where $\varphi_r'(m)$ is the right derivative of φ at point m. If $\varphi_r'(m) < 0$, then one can take $a = -\varphi_r'(m)$. Now, let us prove that $\varphi_r'(m) < 0$. Since φ is non-increasing, $\varphi_r'(m) \leqslant 0$. The function φ being concave, φ_r' is non-increasing. Consequently, if $\varphi_r'(m) = 0$, then $\varphi_r'(x) = 0$, for all $x \leqslant m$. This would imply that $1 - F$ is constant on $(-\infty, m]$, which is absurd. □

Proof of Theorem 6. (i) $\Rightarrow$ (ii). If μ satisfies $\mathbb{ST}_\alpha(a)$, then according to Corollary 15, one has

$$\int e^{\alpha(az)}\, d\mu(z) < +\infty.$$

Hence, (ii) holds with $b = a$.

(ii) $\Rightarrow$ (i). According to Propositions 42 and 43, $\int e^{\alpha(bu)}\,d\mu_x^+(u) \leqslant \int e^{\alpha(bu)}\,d\mu_m^+(u)$. Thus the constants $K^+(b)$ and $K^-(b)$ of Theorem 2 are finite. Since μ belongs to $\mathcal{L}ip_\sharp\mu_1$, one concludes that μ satisfies $\mathbb{ST}_\alpha(a)$ for some $a > 0$. $\square$

5.2. Links with modified Log-Sobolev inequalities

Recall the definition of the entropy functional:

$$\mathrm{Ent}_\mu(f) := \int f \log f\,d\mu - \int f\,d\mu \log \int f\,d\mu.$$

Definition 45. Let $\beta : \mathbb{R} \to \mathbb{R}^+$ be an even convex function with $\beta(0) = 0$. One says that $\mu \in \mathcal{P}(\mathbb{R})$ satisfies the modified Logarithmic Sobolev inequality $LSI_\beta(C, t)$ if

$$\mathrm{Ent}_\mu\big(f^2\big) \leqslant C \int \beta\left(t\frac{f'}{f}\right) f^2\,d\mu,$$

for all smooth $f > 0$.

Note that if $\beta(x) = x^2$, one recovers the classical Logarithmic Sobolev inequality. The links between transportation cost inequalities and Logarithmic Sobolev inequalities have been studied by several authors (see the works by Otto and Villani [19], Bobkov, Gentil and Ledoux [3] and more recently Gentil, Guillin and Miclo [9]). The usual point of view is to prove TCI using Log-Sobolev type inequalities. Here we will do the opposite and derive Log-Sobolev inequalities from TCIs. To this end we will use the following result.

Theorem 46. *Let $\alpha \in \mathcal{A}$ be a convex function. If $\mu = e^{-V}\,dx \in \mathcal{P}(\mathbb{R})$ with $V : \mathbb{R} \to \mathbb{R}$ a convex function satisfies the inequality $\mathbb{T}_\alpha(a)$, then it satisfies $LSI_{\alpha^*}(\frac{\lambda}{1-\lambda}, \frac{1}{a\lambda})$, for all $\lambda \in (0, 1)$, where α^* is the convex conjugate of α:*

$$\alpha^*(s) = \sup_{t \in \mathbb{R}}\big\{st - \alpha(t)\big\}, \quad \forall s \in \mathbb{R}.$$

Proof. The proof of Theorem 46 can be easily adapted from the one of [9, Theorem 2.9]. The regularity issue mentioned by the authors during the proof, is irrelevant in our framework. Namely, in dimension one, the Brenier map is simply the monotone rearrangement map, and the regularity of this latter can be easily checked by hand. $\square$

The following result follows immediately from Theorems 6 and 46.

Theorem 47. *Let $\alpha \in \mathcal{A}$ be a convex function and $\mu = e^{-V}\,dx \in \mathcal{P}(\mathbb{R})$ with V convex. If $\int e^{\alpha(b|x|)}\,d\mu(x) < +\infty$, for some $b > 0$, then μ satisfies the inequality $LSI_{\alpha^*}(C, t)$ for some $C, t > 0$.*

Remark 48. Let θ_p be defined by

$$\theta_p(x) = \begin{cases} x^2 & \text{if } |x| \leqslant 1, \\ \frac{2}{p}|x|^p + 1 - \frac{2}{p} & \text{if } |x| \geqslant 1, \end{cases} \quad \forall p \in [1,2].$$

In [9], I. Gentil, A. Guillin and L. Miclo proved that the measure $d\mu_p(x) = \frac{1}{Z_p} e^{-|x|^p}\, dx$ with $p \in [1,2]$ satisfies the inequality $LSI_{\theta_p^*}(C,t)$ for some $C, t > 0$. Using classical tools, one can show that

$$\exists C, t > 0 \quad \text{s.t. } \mu \text{ satisfies } LSI_{\theta_p^*}(C,t) \quad \Rightarrow \quad \exists b > 0 \quad \text{s.t.} \int e^{\theta_p(ax)}\, d\mu(x) < +\infty.$$

Consequently, a Log-concave measure μ satisfies the inequality $LSI_{\theta_p^*}(C,t)$ if and only if there is some $b > 0$ such that $\int e^{\theta_p(ax)}\, d\mu(x) < +\infty$.

Suppose that $d\mu = e^{-V}\, dx$ with V a convex and symmetric function. It is tempting to take $\alpha = V$ in the above theorem. To do this one has to modify the potential V near 0. Define

$$\widetilde{V}(x) = \begin{cases} x^2 & \text{if } |x| \leqslant 1, \\ V(a_0 x) + 1 - V(a_0) & \text{if } |x| \geqslant 1. \end{cases}$$

Choosing $a_0 > 0$ such that $a_0 V'(a_0) = 2$ (which is always possible), one obtains a convex function. Furthermore, it is clear that one can find some $b > 0$ such that $\int e^{\widetilde{V}(bx)}\, d\mu(x) < +\infty$. Applying the above theorem, one obtains the following result.

Corollary 49. *With the above notations, μ satisfies the inequality $LSI_{\widetilde{V}^*}(C,t)$ for some $C, t > 0$.*

Remark 50. In [10], Gentil, Guillin and Miclo have obtained the preceding corollary under the following additional assumption on V:

$$\exists \varepsilon \in [0, 1/2],\ \exists M > 0,\ \forall x \geqslant M, \quad (1+\varepsilon)V(x) \leqslant xV'(x) \leqslant (2-\varepsilon)V(x).$$

This hypothesis seems to be useless.

Acknowledgments

I want to warmly acknowledge Patrick Cattiaux, Arnaud Guillin and Christian Léonard for so many interesting conversations on functional inequalities and other topics.

References

[1] C. Ané, S. Blachère, D. Chafaï, P. Fougères, I. Gentil, F. Malrieu, C. Roberto, G. Scheffer, Sur les inégalités de Sobolev logarithmiques, in: Panor. Synthèses, vol. 10, Soc. Math. de France, Paris, 2000.

[2] F. Barthe, C. Roberto, Modified logarithmic-Sobolev inequalities, preprint, 2006.

[3] S.G. Bobkov, I. Gentil, M. Ledoux, Hypercontractivity of Hamilton–Jacobi equations, J. Math. Pures Appl. (9) 80 (7) (2001) 669–696.

[4] S.G. Bobkov, F. Götze, Exponential integrability and transportation cost related to logarithmic Sobolev inequalities, J. Funct. Anal. 163 (1) (1999) 1–28.

[5] S.G. Bobkov, C. Houdré, Isoperimetric constants for product probability measures, Ann. Probab. 25 (1) (1997) 184–205.
[6] S.G. Bobkov, C. Houdré, Weak dimension-free concentration of measure, Bernoulli 6 (4) (2000) 621–632.
[7] P. Cattiaux, A. Guillin, Talagrand's like quadratic transportation cost inequalities, J. Math. Pures Appl. 86 (9) (2004) 341–361.
[8] 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.
[9] I. Gentil, A. Guillin, L. Miclo, Modified logarithmic Sobolev inequalities and transportation inequalities, Probab. Theory Related Fields 133 (3) (2005) 409–436.
[10] I. Gentil, A. Guillin, L. Miclo, Modified logarithmic Sobolev inequalities in null curvature, Rev. Math. Iberoamericana 23 (2007) 235–258.
[11] N. Gozlan, C. Léonard, A large deviation approach to some transportation cost inequalities, Probab. Theory Related Fields 139 (2) (2007) 235–283.
[12] L.V. Kantorovich, On an effective method of solving certain classes of extremal problems, Dokl. Akad. Nauk USSR 28 (1940) 212–215.
[13] A.V. Kolesnikov, Modified log-Sobolev inequalities and isoperimetry, preprint, 2006.
[14] M. Ledoux, The Concentration of Measure Phenomenon, Math. Surveys Monogr., vol. 89, Amer. Math. Soc., Providence, RI, 2001.
[15] K. Marton, A simple proof of the blowing-up lemma, IEEE Trans. Inform. Theory 32 (3) (1986) 445–446.
[16] K. Marton, Bounding $\bar{d}$-distance by informational divergence: A method to prove measure concentration, Ann. Probab. 24 (2) (1996) 857–866.
[17] B. Maurey, Some deviation inequalities, Geom. Funct. Anal. 1 (2) (1991) 188–197.
[18] B. Muckenhoupt, Hardy's inequality with weights, Studia Math. 44 (1972) 31–38.
[19] 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.
[20] P.M. Samson, Concentration of measure inequalities for Markov chains and Φ-mixing processes, Ann. Probab. 28 (1) (2000) 416–461.
[21] M. Talagrand, Transportation cost for Gaussian and other product measures, Geom. Funct. Anal. 6 (3) (1996) 587–600.
[22] C. Villani, Topics in Optimal Transportation, Grad. Stud. Math., vol. 58, Amer. Math. Soc., Providence, RI, 2003.
[23] F.-Y. Wang, Probability distance inequalities on Riemannian manifolds and path spaces, J. Funct. Anal. 206 (1) (2004) 167–190.