Annales de l'Institut Henri Poincaré - Probabilités et Statistiques
2018, Vol. 54, No. 4, 2075–2091
https://doi.org/10.1214/17-AIHP865

ANNALES DE L'INSTITUT HENRI POINCARÉ
PROBABILITÉS ET STATISTIQUES
www.imstat.org/aihp

A characterization of a class of convex log-Sobolev inequalities on the real line

Yan Shu[a] and Michał Strzelecki[b,1]

[a] *Université Paris Ouest Nanterre La Défense – Modal'X, 200 avenue de la République, 92000 Nanterre, France. E-mail: yshu@u-paris10.fr*
[b] *Institute of Mathematics, University of Warsaw, Banacha 2, 02-097 Warsaw, Poland. E-mail: M.Strzelecki@mimuw.edu.pl*

Received 16 February 2017; revised 23 May 2017; accepted 20 September 2017

Abstract. We give a sufficient and necessary condition for a probability measure μ on the real line to satisfy the logarithmic Sobolev inequality for convex functions. The condition is expressed in terms of the unique left-continuous and non-decreasing map transporting the symmetric exponential measure onto μ. The main tool in the proof is the theory of weak transport costs.

As a consequence, we obtain dimension-free concentration bounds for the lower and upper tails of convex functions of independent random variables which satisfy the convex log-Sobolev inequality.

Résumé. Nous proposons une condition nécessaire et suffisante pour qu'une mesure de probabilité μ sur la droite réelle satisfasse une condition de Sobolev logarithmique sur les fonctions convexes. Cette condition est exprimée en termes de l'unique plan de transport optimal croissant et continu à gauche entre la mesure exponentielle symétrique et la mesure μ. L'outil principal vient de la théorie du transport faible.

Comme conséquence, nous obtenons un résultat de concentration adimensionnelle sur les estimées de queue de fonctions convexes de variables aléatoires indépendantes, lié à l'inégalité de Sobolev logarithmique convexe.

MSC: Primary 60E15; secondary 26A51; 26D10

Keywords: Concentration of measure; Convex functions; Log-Sobolev inequality; Weak transport-entropy inequalities

1. Introduction and main results

Let μ be a Borel probability measure on $\mathbb{R}^n$. We say that μ satisfies the (modified) log-Sobolev inequality for a class of functions $\mathcal{F}$ (with cost function $H : \mathbb{R}^n \to [0, \infty)$ and constant $c < \infty$) if for every $f \in \mathcal{F}$ we have

$$\mathrm{Ent}_\mu\big(e^f\big) \leq \int_{\mathbb{R}^n} H(c\nabla f)e^f \, d\mu. \tag{1.1}$$

Here $\mathrm{Ent}_\mu(g)$ denotes the usual entropy of a non-negative function g, i.e.

$$\mathrm{Ent}_\mu(g) = \int_{\mathbb{R}^n} g \ln(g) \, d\mu - \int_{\mathbb{R}^n} g \, d\mu \ln\left(\int_{\mathbb{R}^n} g \, d\mu\right) \tag{1.2}$$

if $\int_{\mathbb{R}^n} g \ln(g) \, d\mu < \infty$ and $\mathrm{Ent}_\mu(g) = \infty$ otherwise.

---

[1] Research partially supported by the National Science Centre, Poland, grant no. 2015/19/N/ST1/00891.

In the most classical setting where $H(x) = |x|^2$ and $\mathcal{F}$ is the class of C^1 functions this inequality was first introduced by Gross in [20]. In this case it can be rewritten in the form

$$\mathrm{Ent}_\mu\big(g^2\big) \le 4C \int_{\mathbb{R}^n} |\nabla g|^2 \, d\mu, \tag{1.3}$$

where $C = c^2$, or in yet another form which states that the entropy of a positive function g is bounded by its Fisher information:

$$\mathrm{Ent}_\mu(g) \le C \int_{\mathbb{R}^n} \frac{|\nabla g|^2}{g} \, d\mu.$$

Due to its tensorization property the log-Sobolev inequality is a powerful tool and can be used to obtain dimension-free concentration bounds (via the so-called Herbst argument). It has been investigated also in more general settings of Riemannian manifolds and in the context of applications to the study of Markov chains (see, e.g., the monographs [4,5] and the expository article [12]).

Inequality (1.1) with $\mathcal{F} = C^1$ and non-quadratic functions H has been studied by Bobkov and Ledoux in [8], who considered

$$H(x) = |x|^2 1_{\{|x| \le \delta\}} + \infty 1_{\{|x| \ge \delta\}}, \quad x \in \mathbb{R}^n,$$

and Gentil, Guillin, and Miclo in [15,16], with H essentially of the form

$$\sum_{i=1}^n \max\big\{|x_i|^2, |x_i|^p\big\}, \quad x = (x_1, \ldots, x_n) \in \mathbb{R}^n, \, p \ge 2.$$

Also, in the case of modified log-Sobolev inequalities on the real line, Barthe and Roberto in [7] characterized (under some mild technical conditions) measures satisfying inequality (1.1) with

$$H(x) = x^2 1_{\{|x| \le 1\}} + \frac{\Phi(|x|)}{\Phi(1)} 1_{\{|x| > 1\}}, \quad x \in \mathbb{R},$$

where $\Phi : [0, \infty) \to [0, \infty)$ is a sufficiently nice Young function (cf. [6]; the quadratic case was handled earlier by Bobkov and Götze in [11]).

In this paper we are primarily interested in the case when $\mathcal{F}$ is the class of *convex* functions.[2] The restriction of the class of functions allows us to work with measures which satisfy much weaker regularity conditions. Most importantly, their supports do not need to be connected, as opposed to many classical functional inequalities such as the Poincaré, transport-information, or log-Sobolev inequality. On the other hand, a disturbing issue arises: the log-Sobolev inequality for convex functions yields via standard reasonings only deviation inequalities for the upper tail of functions, i.e.

$$\mu^{\otimes N}\left(\left\{x \in \big(\mathbb{R}^n\big)^N : f(x) \ge \int_{(\mathbb{R}^n)^N} f \, d\mu^{\otimes N} + t\right\}\right), \quad t \ge 0$$

(in the classical setting of smooth functions one obtains bounds on the lower tail simply by working with $-f$ instead of f, but this is precluded in our situation because $-f$ is usually not convex).

For brevity in what follows we shall slightly informally refer to the log-Sobolev inequality for convex functions simply as "the convex log-Sobolev inequality".

Our goal is to give an intrinsic characterization of probability measures on the real line for which the convex log-Sobolev inequality holds. As a corollary we will obtain dimension-free concentration bounds for upper *and* lower tails of convex functions of independent random variables satisfying the convex log-Sobolev inequality. Before stating our main result let us outline what is known in the convex setting.

[2] Note that in this case the dimension-free tensorization property still holds, but the alternative formulations (1.2) and (1.3) – with g being convex, respectively convex and non-negative – are no longer equivalent to (1.1).

In [1] Adamczak found a sufficient condition for a probability measure on the real line to satisfy the convex log-Sobolev inequality with $H(x) = x^2$, $x \in \mathbb{R}$. This has been extended to functions of the form $H(x) = \max\{x^2, x^{(\beta+1)/\beta}\}$, where $\beta \in [0, 1]$, by Adamczak and the second named author in [3].

Recall that in the classical setting there are strong links between the log-Sobolev inequality, transport-entropy inequalities (introduced by Talagrand in [24] and subsequently widely studied, see, e.g., [25] for a complete and detailed introduction) and the infimum convolution inequality (first introduced by Maurey in [21]). Specifically, Bobkov and Götze in [11] showed that the transport-entropy inequality is equivalent to some kind of infimum-convolution inequality, and later Bobkov, Gentil, and Ledoux in [9] established a condition equivalent to the log-Sobolev inequality in terms of the infimum-convolution inequality (these results are closely connected with the celebrated Otto–Villani theorem, see [22,23]).

Similar connections have been observed by Gozlan, Roberto, Samson, and Tetali in [19] for convex log-Sobolev inequalities, *weak* transport-entropy inequalities (introduced therein, see Section 2 for definitions and more details), and the convex infimum convolution inequality. Later, Gozlan, Roberto, Samson, Tetali, and the first named author established in [18] a condition sufficient for the convex modified log-Sobolev inequality to hold on the real line. This condition is expressed in terms of the unique left-continuous and non-decreasing map transporting the symmetric exponential measure onto μ (more precisely, in terms of the quantity $\Delta_\mu(h)$ defined below in (1.4); see Section 4 for a precise statement of the result). In fact, their sufficient condition is weaker than the condition considered in [3] (and leads to a result formally stronger than the convex log-Sobolev inequality, cf. Proposition 4.1).

On the other hand, it follows from [18] and the independent work of Feldheim, Marsiglietti, Nayar, and Wang in [14] that in the case when H is quadratic on an interval near zero and then infinite the following are equivalent:

- the condition on the tail of the measure μ from [3] in the case $\beta = 0$,
- the condition on monotone transport map obtained in [18],
- the log-Sobolev inequality for convex functions

(and further: the convex Poincaré inequality, the convex infimum convolution inequality with a quadratic-linear cost function). In what follows we extend this result to more general choices of the function H.

In order to formulate our main result we need to introduce some notation. Let τ be the symmetric exponential measure on $\mathbb{R}$ with density $\frac{1}{2}e^{-|x|}$. For a Borel probability measure μ on $\mathbb{R}$ we denote by U_μ the unique left-continuous and non-decreasing map transporting τ onto the reference measure μ. More precisely,

$$U_\mu(x) = F_\mu^{-1} \circ F_\tau(x) = \begin{cases} F_\mu^{-1}(\frac{1}{2}e^{-|x|}) & \text{if } x < 0, \\ F_\mu^{-1}(1 - \frac{1}{2}e^{-|x|}) & \text{if } x \geq 0, \end{cases}$$

where

$$F_\mu^{-1}(t) = \inf\{y \in \mathbb{R} : F_\mu(y) \geq t\} \in \mathbb{R} \cup \{\pm\infty\}, \quad t \in [0, 1],$$

is the generalized inverse of the cumulative distribution function defined as

$$F_\mu(x) = \mu((-\infty, x]), \quad x \in \mathbb{R}.$$

Denote moreover by

$$\Delta_\mu(h) = \sup_{x \in \mathbb{R}}\{U_\mu(x+h) - U_\mu(x)\}, \quad h > 0, \tag{1.4}$$

the modulus of continuity of U_μ (by all means we can have $\lim_{h\to 0^+} \Delta_\mu(h) > 0$).

Our main result is the following. The results are new already in the case of the quadratic function $H(x) = \frac{1}{4}x^2$.

Theorem 1.1. *Let $H : \mathbb{R} \to [0, \infty)$ be a symmetric convex function, such that $H(x) = \frac{1}{4}x^2$ for $x \in [-2t_0, 2t_0]$ for some $t_0 > 0$, and let $H^* : \mathbb{R} \to [0, \infty)$ be its Fenchel–Legendre transform. Suppose moreover that $\lim_{x\to\infty} H^*(x)/x = \infty$ and that there exist $A \in [1, \infty)$ and $\alpha \in (1, 2]$ such that*

$$\forall_{x \in \mathbb{R}}\ \forall_{s \in [0,1]} \quad H(sx) \leq As^\alpha H(x). \tag{1.5}$$

Denote $\theta(t) = H^*(t)$, $t \geq 0$. For a probability measure μ on the real line the following conditions are equivalent.

(i) *For every $s > 0$ we have $\int_{\mathbb{R}} e^{s|x|}\, d\mu(x) < \infty$ and there exists $c > 0$ such that*

$$\mathrm{Ent}_\mu\big(e^\varphi\big) \le \int_{\mathbb{R}} H\big(c\varphi'\big) e^\varphi \, d\mu$$

for every smooth convex Lipschitz function $\varphi : \mathbb{R} \to \mathbb{R}$.

(ii) *There exists $b > 0$ such that for all $h > 0$,*

$$\Delta_\mu(h) \le \frac{1}{b}\theta^{-1}\big(t_0^2 + h\big).$$

Remark 1.2. The dependence of constants is explicit but complicated and hence we shall only specify it throughout the proof of the theorem. However, in the case when $H(x) = \frac{x^2}{4}$ the dependence of constants can be simplified: (ii) implies (i) with $c = \frac{1}{\kappa b}$ where

$$\kappa = \max_{t_0 > 0}\left\{ \frac{\min(1, t_0)}{210\sqrt{2 + t_0^2}} \right\} = \frac{1}{210\sqrt{3}};$$

(i) implies that $\Delta_\mu(h) \le 16c(\frac{2}{3} + \sqrt{h/2})$.

Remark 1.3. The condition (1.5) is stable under taking convex combinations or maxima of functions and, for $1 < p < \infty$, the function

$$H(x) = H_p(x) = \begin{cases} \frac{1}{4}x^2 & \text{if } |x| \le 2, \\ \frac{2}{p}(|x/2|^p - 1) + 1 & \text{if } |x| > 2, \end{cases}$$

satisfies (1.5) with $\alpha = \min\{p, 2\}$ and $A = 1$. Indeed, if $x > 0$ and $s \in (0, 1)$, then by Cauchy's mean value theorem

$$\frac{H(sx)}{H(x)} = \frac{sH'(s\xi)}{H'(\xi)} = \begin{cases} s^2 & \text{if } 0 \le s\xi \le \xi \le 2, \\ s^2(\xi/2)^{2-p} = s^p(s\xi/2)^{2-p} & \text{if } 0 \le s\xi \le 2 < \xi, \\ s^p & \text{if } 2 < s\xi \le \xi \end{cases}$$

for some $\xi \in (0, x)$. In either case, $H(sx) \le \max\{s^2, s^p\}H(x) = s^{\min\{p,2\}}H(x)$.

Remark 1.4. The condition on Δ_μ from (ii) is to some extent related to the conditions considered in [1,3,14] (see also Lemma 4.5 below). Indeed, suppose for simplicity that μ is symmetric and has a nowhere vanishing density. By the definition of U_μ we have

$$\mu\big(\big[U_\mu(x + h), \infty\big)\big) = \tau\big([x + h, \infty)\big) = e^{-h}\tau\big([x, \infty)\big) = e^{-h}\mu\big(\big[U_\mu(x), \infty\big)\big)$$

for $x, h \ge 0$. This easy computation shows that (ii) implies:

(ii$'$) There exists $b > 0$ such that for every $h > 0$,

$$\mu\big(\big[x + g(h), \infty\big)\big) \le e^{-h}\mu\big([x, \infty)\big) \quad \forall x \ge 0,$$

where $g(h) = \frac{1}{b}\theta^{-1}(t_0^2 + h)$.

The last inequality is similar to the inequalities which appear in the definitions of the classes $\mathcal{M}_\beta$ from [1,3], and [14] (but there h is fixed and x is assumed to be sufficiently large; moreover $g = g(x)$ in [1,3]).

Remark 1.5. In (i) the assumption about exponential integrability is added in order to exclude very heavy-tailed measures for which the only exponentially integrable convex Lipschitz functions are constants and hence the convex log-Sobolev inequality is trivially satisfied, whereas (ii) cannot hold.

Remark 1.6. In the definition of the log-Sobolev inequality the constant c is introduced as a scaling of the argument of the function H rather than as a multiplicative constant outside of the integral. We decided to use this form because it simplifies some of the calculations in Section 3. Clearly, in the most common cases, e.g., in the case of the functions H_p from Remark 1.3, the two formulations are equivalent up to numerical constants (for the functions H_p those constants depend on p).

Since the condition (ii) is in fact equivalent to the infimum convolution inequality for convex functions considered in [18] (see also Proposition 4.1 and Proposition 2.1 below) we immediately obtain concentration bounds for the upper and lower tails of convex Lipschitz functions. For simplicity we state them in the case $H(x) = \frac{1}{4}x^2$, but one can obtain appropriate results also in the case when H is not quadratic. We refer to, e.g., [14, Corollary 3] for a similar statement (with H quadratic on an interval near zero and then infinite) and to [19, Corollary 5.11] for a general overview of concentration properties implied by weak transportation inequalities.

Corollary 1.7. *Let μ be a probability measure on $\mathbb{R}$ such that for every $s > 0$ we have $\int_{\mathbb{R}} e^{s|x|}\, d\mu(x) < \infty$ and*

$$\operatorname{Ent}_\mu\big(e^\varphi\big) \le C \int_{\mathbb{R}} \big|\varphi'\big|^2 e^\varphi \, d\mu$$

for every smooth convex Lipschitz function $\varphi : \mathbb{R} \to \mathbb{R}$. Then there exist $A, B < \infty$ (depending only on C), such that for any convex (or concave) function $\varphi : \mathbb{R}^N \to \mathbb{R}$ which is 1-Lipschitz (with respect to the Euclidean norm on $\mathbb{R}^N$) we have

$$\mu^{\otimes N}\big(\big\{x \in \mathbb{R}^N : \big|\varphi(x) - \operatorname{Med}_{\mu^{\otimes N}}(\varphi)\big| \ge t\big\}\big) \le Be^{-t^2/A}, \quad t \ge 0,$$

where

$$\operatorname{Med}_{\mu^{\otimes N}}(\varphi) = \inf\big\{s \in \mathbb{R} : \mu^{\otimes N}\big(\big\{x \in \mathbb{R}^N : \varphi(x) \le s\big\}\big) \ge 1/2\big\}$$

is the median of φ.

The article is organized as follows. In Section 2 we recall the definitions of the weak transport-entropy inequalities and provide preliminary results about infimum convolution inequalities from [19] and [18]. In Sections 3 and 4 we prove that the conditions (i) and (ii) both are equivalent to a weak transport-entropy inequality. Finally, in Section 5 we summarize the results of the previous sections and give the proof of Theorem 1.1 (and Corollary 1.7). We also recapitulate all conditions equivalent to the convex log-Sobolev inequality in the quadratic case and pose some open questions.

2. A reminder about weak transport-entropy inequalities

In this section, we recall the definition of the weak transport problem and some preliminary results. Although throughout the most part of the paper we work with measures on the real line we introduce the notation in a slightly greater generality. Below $|\cdot|$ denotes the standard Euclidean norm in $\mathbb{R}^n$.

2.1. *Weak transport costs*

For two probability measures μ_1, μ_2 on $\mathbb{R}^n$ the weak (barycentric) transport cost associated to the convex cost function $\theta : [0, \infty) \to [0, \infty]$ is defined by the formula

$$\overline{\mathcal{T}}_\theta(\mu_2|\mu_1) = \inf_\pi \int_{\mathbb{R}^n} \theta\left(\left|x - \int_{\mathbb{R}^n} y p(x, dy)\right|\right)\mu_1(dx),$$

where the infimum is taken over all couplings $\pi(dx\,dy) = p(x, dy)\mu_1(dx)$ of μ_1 and μ_2, and $p(x, \cdot)$ denotes the disintegration kernel of π with respect to its first marginal. Using probabilistic notation we can write

$$\overline{\mathcal{T}}_\theta(\mu_2|\mu_1) = \inf \mathbb{E}\theta\big(\big|X_1 - \mathbb{E}(X_2|X_1)\big|\big),$$

where the infimum is taken over all random variables $X_1 \sim \mu_1$, $X_2 \sim \mu_2$. The adjective *weak* stands for the fact that, by Jensen's inequality, it is smaller than the classical transport cost,

$$\mathcal{T}_\theta(\mu_2, \mu_1) = \inf_\pi \int_{\mathbb{R}^n \times \mathbb{R}^n} \theta\big(|x-y|\big)\pi(dx, dy), \tag{2.1}$$

considered in the Monge–Kantorovich transport problem. Note also that in contrast to $\mathcal{T}_\theta$ the weak transport cost $\overline{\mathcal{T}}_\theta$ is not symmetric.

2.2. *From weak transport inequality to infimum convolution inequality*

Recall that for two probability measures the relative entropy of ν with respect to μ is given by the formula

$$H(\nu|\mu) = \int_{\mathbb{R}^n} \log\left(\frac{d\nu}{d\mu}\right) d\nu \in [0, +\infty]$$

if ν is absolutely continuous with respect to μ; otherwise we set $H(\nu|\mu) = +\infty$.

We say that a probability measure μ on $\mathbb{R}^n$ *satisfies the weak transport-entropy inequality* $\overline{\mathbf{T}}^-(\theta)$ (respectively, $\overline{\mathbf{T}}^+(\theta)$) if

$$\overline{\mathcal{T}}_\theta(\mu|\nu) \le H(\nu|\mu) \quad \big(\text{respectively, } \overline{\mathcal{T}}_\theta(\nu|\mu) \le H(\nu|\mu)\big)$$

for every probability measure ν on $\mathbb{R}^n$ having a finite first moment. We say that μ satisfies $\overline{\mathbf{T}}(\theta)$ if μ satisfies both $\overline{\mathbf{T}}^-(\theta)$ and $\overline{\mathbf{T}}^+(\theta)$.

A Bobkov–Götze type criterion for the weak-transport inequality was given in [19, Proposition 4.5] (cf. also [18, Lemma 4.1], where a formulation more similar to the below is given). It is expressed in terms of the infimum convolution operator Q_t^θ defined by the formula

$$Q_t^\theta f(x) = \inf_{y \in \mathbb{R}^n} \big\{ f(y) + t\theta\big(|x-y|/t\big)\big\}, \quad x \in \mathbb{R}^n,$$

for $t > 0$ and $f : \mathbb{R}^n \to \mathbb{R}$.

Proposition 2.1 ([19, Proposition 4.5]). *Let μ be a probability measure on $\mathbb{R}^n$ and let $\theta : [0, \infty) \to [0, \infty]$ be a convex cost function. Then the following holds.*

(a) *The measure μ satisfies $\overline{\mathbf{T}}^-(\theta)$ if and only if*

$$\int_{\mathbb{R}^n} \exp\big(Q_1^\theta f\big)\, d\mu \exp\left(-\int_{\mathbb{R}^n} f\, d\mu\right) \le 1 \tag{2.2}$$

for every convex function $f : \mathbb{R}^n \to \mathbb{R}$ bounded from below.

(b) *The measure μ satisfies $\overline{\mathbf{T}}^+(\theta)$ if and only if*

$$\exp\left(\int_{\mathbb{R}^n} Q_1^\theta f\, d\mu\right) \int_{\mathbb{R}^n} \exp(-f)\, d\mu \le 1$$

for every convex function $f : \mathbb{R}^n \to \mathbb{R}$ bounded from below.

(c) *If the measure μ satisfies $\overline{\mathbf{T}}(\theta)$, then*

$$\int_{\mathbb{R}^n} \exp\big(Q_t^\theta f\big)\, d\mu \int_{\mathbb{R}^n} \exp(-f)\, d\mu \le 1$$

for $t = 2$ and every convex function $f : \mathbb{R}^n \to \mathbb{R}$ bounded from below. Conversely, if the above inequality holds for some $t > 0$, then μ satisfies $\overline{\mathbf{T}}(t\theta(\cdot/t))$.

3. Equivalence of the convex log-Sobolev inequality and the weak transportation inequality

In this section we establish the equivalence of the convex log-Sobolev inequality and the weak transport-entropy inequality. In the case of the quadratic cost this was done in [19] (see also [3] for related results for other cost functions). Using the techniques developed therein, especially in the dual formulation (2.2), we extend this result to a wider class of cost functions. We work with measures on the real line, but in contrast to Section 4 there are no essential problems with extending the results of this section to a higher dimensional setting (cf. [3,19]).

Let $H : \mathbb{R} \to [0, \infty)$ be a symmetric convex function, such that $H(x) = \frac{1}{4}x^2$ for $x \in [-2t_0, 2t_0]$ for some $t_0 > 0$. Note that then the Fenchel–Legendre transform of H, given by the formula

$$H^*(x) = \sup_{y\in\mathbb{R}}\{xy - H(y)\}, \quad x \in \mathbb{R},$$

is also quadratic near zero (namely, $H^*(x) = x^2$ for $x \in [-t_0, t_0]$, since for such x the supremum in the definition of H^* is attained at $y = 2x$). Moreover, we assume that there exist $A \in [1, \infty)$ and $\alpha \in (1, 2]$ such that

$$\forall_{x\in\mathbb{R}}\ \forall_{s\in[0,1]} \quad H(sx) \le As^\alpha H(x). \tag{3.1}$$

Finally, we assume that $\lim_{x\to\infty} H^*(x)/x = \infty$.

Note that we have

$$H(x) \ge (4A)^{-1} t_0^{2-\alpha} x^\alpha \quad \text{for } x \ge t_0 \tag{3.2}$$

(this follows immediately by taking $s = t_0/x$ in condition (3.1)).

The main result of this section is the following.

Proposition 3.1. *For a probability measure μ on the real line the following conditions are equivalent.*

(i) *There exists $a > 0$ such that the measure μ satisfies $\overline{\mathbf{T}}^-(H^*(a\cdot))$.*
(ii) *For every $s > 0$ we have $\int_\mathbb{R} e^{s|x|}\, d\mu(x) < \infty$ and there exists $c > 0$ such that*

$$\operatorname{Ent}_\mu(e^\varphi) \le \int_\mathbb{R} H(c\varphi')e^\varphi\, d\mu$$

for every smooth convex Lipschitz function $\varphi : \mathbb{R} \to \mathbb{R}$.

The dependence of the constants is the following: (i) *implies* (ii) *with* $c = 2/a$; (ii) *implies* (i) *with* $a = ((\alpha-1)/A)^{1/\alpha}c^{-1}$.

The implication (i) $\implies$ (ii) is a general fact and no special assumptions are used in the proof. For the sake of completeness we sketch the main argument here.

Proof of Proposition 3.1, (i) $\implies$ (ii). The exponential integrability follows from the dual formulation of $\overline{\mathbf{T}}^-(H^*(a\cdot))$ tested with the function $x \mapsto s|x|$ (cf. [3, p. 86]).

According to [19, Proposition 8.3], (i) implies that the so-called (τ)-log-Sobolev inequality holds: for all functions $f : \mathbb{R} \to \mathbb{R}$ with $\int_\mathbb{R} fe^f\, d\mu < \infty$ we have

$$\operatorname{Ent}_\mu(e^f) \le \frac{1}{1-\lambda}\int_\mathbb{R} (f - R^\lambda f)e^f\, d\mu$$

for every $\lambda \in (0, 1)$. Here

$$R^\lambda f(x) := \inf_p\left\{\int_\mathbb{R} f(y)p(dy) + \lambda H^*\left(a\left|x - \int_\mathbb{R} yp(dy)\right|\right)\right\},$$

where the infimum is taken over all probability measures p on $\mathbb{R}$ (note that we skip the dependence on H^* in the notation). Since H^* is convex, for convex functions f the infimum above is achieved at some Dirac measure:

$$R^\lambda f(x) = \inf_{y\in\mathbb{R}} \{ f(y) + \lambda H^*\big(a(x-y)\big)\}$$

(indeed, if we replace the measure p by a Dirac mass at the point $y_p = \int_{\mathbb{R}} y p(dy)$, then, by Jensen's inequality, the expression under the infimum in the definition of R^λ will not increase; cf. Theorem 2.11 and Section 8.2 in [19]). Now, by convexity of f,

$$\begin{aligned} f(x) - R^\lambda f(x) &= f(x) + \sup_{y\in\mathbb{R}} \{ -f(y) - \lambda H^*\big(a(x-y)\big)\} \\ &\le f(x) + \sup_{y\in\mathbb{R}} \{ -f(x) - f'(x)(y-x) - \lambda H^*\big(a(x-y)\big)\} = \lambda H\left(\frac{f'(x)}{a\lambda}\right), \end{aligned}$$

where we have used the fact that $H^{**} = H$. Thus, after taking $\lambda = 1/2$, we arrive at the assertion (with $c = 2/a$). □

For the proof of the second implication we need the following simple lemma. It is based on an argument of Maurey (cf. [21, Proof of Theorem 3]), but takes into account the observation that for compactly supported measures it doesn't matter whether for large arguments the cost function is quadratic or equal to $+\infty$. Recall that $|\cdot|$ stands for the Euclidean norm.

Lemma 3.2. *Let μ be a probability measure on $\mathbb{R}^n$ such that the diameter of the support of μ is not greater than D and denote*

$$\theta_D(x) = \begin{cases} \frac{1}{4D^2}|x|^2 & \text{if } |x| \le D, \\ +\infty & \text{if } |x| > D. \end{cases}$$

Then for any convex function $\varphi : \mathbb{R}^n \to \mathbb{R}$ bounded from below

$$\int_{\mathbb{R}^n} e^{Q_1^{\theta_D}\varphi}\, d\mu \int_{\mathbb{R}^n} e^{-\varphi}\, d\mu \le 1, \tag{3.3}$$

where $Q_1^{\theta_D}\varphi(x) = \inf\{\varphi(y) + \theta_D(x-y) : y \in \mathbb{R}^n\}$, $x \in \mathbb{R}^n$, stands for the infimum convolution.

Conversely, if inequality (3.3) holds for some D, then the support of μ is bounded: $x, y \in \operatorname{Supp}\mu \implies |x-y| \le D$.

Proof. Assume that the diameter of the support of μ is bounded by D. Take a convex function φ, bounded from below. By adding a constant to φ, we may assume that $\inf_{\operatorname{Supp}\mu}\varphi = 0$. Take any $\varepsilon > 0$, any $x \in \operatorname{Supp}\mu$, and let $z \in \operatorname{Supp}\mu$ be such that $\varphi(z) < \varepsilon$. Moreover, define $y = (1-\lambda)x + \lambda z$, where $\lambda \in [0,1]$. Then $|x-y| \le D$ and hence

$$\begin{aligned} Q_1^{\theta_D}\varphi(x) &\le \varphi(y) + \frac{1}{4D^2}|x-y|^2 \le (1-\lambda)\varphi(x) + \lambda\varphi(z) + \frac{\lambda^2}{4D^2}|x-z|^2 \\ &\le (1-\lambda)\varphi(x) + \lambda\varepsilon + \frac{\lambda^2}{4}. \end{aligned}$$

We now let $\varepsilon \to 0^+$, and then optimize with respect to $\lambda \in [0,1]$: if $\varphi(x) \ge 1/2$ we take $\lambda = 1$, and if $0 \le \varphi(x) \le 1/2$ we take $\lambda = 2\varphi(x)$. This gives $Q_1^{\theta_D}\varphi(x) \le k(\varphi(x))$, where

$$k(u) = \big(u - u^2\big) \cdot 1_{\{u\in[0,1/2)\}} + \frac{1}{4} \cdot 1_{\{u \ge 1/2\}}.$$

Note that we have $e^{k(u)} \le 2 - e^{-u}$. Indeed, for $u = 1/2$ (or larger) the inequality holds, and for $u \in [0, 1/2)$ we have

$$\big(e^{u-u^2} + e^{-u}\big)/2 \le e^{-u^2/2}\cosh\big(u - u^2/2\big) \le e^{-u^2/2}\cosh(u) \le 1.$$

Hence

$$\int e^{Q_1^{\theta_D}\varphi}\mu \le \int e^{k(\varphi)}\,d\mu \le 2 - \int e^{-\varphi}\,d\mu \le \left(\int e^{-\varphi}\,d\mu\right)^{-1}.$$

Conversely, assume that inequality (3.3) holds, but there exist $x_0, y_0 \in \operatorname{Supp}\mu$ such that $|x_0 - y_0| > D$. Then there exist $\varepsilon, \delta > 0$, such that $\mu(B(x_0,\varepsilon)) > \delta$ and $\mu(\mathbb{R}^n \setminus B(x_0, D+2\varepsilon)) > \delta$. Consider now $\varphi_a : \mathbb{R}^n \to \mathbb{R}$ defined by the formula $\varphi_a(x) = a\,\mathrm{dist}(x, B(x_0,\varepsilon))$ for $a > 0$. For $x \in \mathbb{R}^n \setminus B(x_0, D+2\varepsilon)$ we have

$$Q_1^{\theta_D}\varphi_a(x) = \inf_{y\in\mathbb{R}^n : |x-y|\le D}\left\{a\,\mathrm{dist}\bigl(y, B(x_0,\varepsilon)\bigr) + \frac{1}{4D^2}|x-y|^2\right\} \ge a\varepsilon.$$

Moreover, $\varphi_a = 0$ on $B(x_0,\varepsilon)$. Thus for sufficiently large $a > 0$,

$$\int_{\mathbb{R}^n} e^{Q_1^{\theta_D}\varphi_a}\,d\mu \int_{\mathbb{R}^n} e^{-\varphi_a}\,d\mu \ge \int_{\mathbb{R}^n\setminus B(x_0,D+2\varepsilon)} e^{Q_1^{\theta_D}\varphi_a}\,d\mu \int_{B(x_0,\varepsilon)} e^{-\varphi_a}\,d\mu \ge \delta^2\exp(a\varepsilon) > 1,$$

which contradicts inequality (3.3). □

Proof of Proposition 3.1, (ii) $\Longrightarrow$ (i). Assume that (ii) holds. Without loss of generality we can assume that μ is absolutely continuous with respect to the Lebesgue measure. Indeed, if γ is a uniform probability distribution on $[0,\delta]$, then by Lemma 3.2, Proposition 2.1(c), and the already proved implication (i) $\Longrightarrow$ (ii) of Proposition 3.1 it satisfies the convex log-Sobolev inequality with a quadratic-linear function

$$H_0(x) = \delta^2 x^2 1_{\{|x|\le 1/(2\delta)\}} + \bigl(\delta|x| - 1/4\bigr)1_{\{|x|>1/(2\delta)\}}$$

(and constant $c = 2$). Hence by (3.2) the product measure $\mu\otimes\gamma$ on $\mathbb{R}^2$ satisfies (for sufficiently small $\delta > 0$)

$$\mathrm{Ent}_{\mu\otimes\gamma}\bigl(e^{\phi}\bigr) \le \int_{\mathbb{R}^2}\bigl(H\bigl(c\phi'_x\bigr) + H\bigl(c\phi'_y\bigr)\bigr)e^{\phi}\,d\mu\otimes\gamma \tag{3.4}$$

for all smooth convex Lipschitz functions $\phi : \mathbb{R}^2 \to \mathbb{R}$. Let $\varphi : \mathbb{R}\to\mathbb{R}$ be a smooth convex Lipschitz function and let $\phi : \mathbb{R}^2 \to \mathbb{R}$ be defined by the formula $\phi(x,y) = \varphi(x+\varepsilon y)$, $x, y \in \mathbb{R}$. Applying (3.4) to the function ϕ and using the assumption (3.1), we see that the convolution $\mu * \gamma_\varepsilon$, where $\gamma_\varepsilon(\cdot) = \gamma(\cdot/\varepsilon)$, satisfies, up to a multiplicative constant which tends to 1 as $\varepsilon \to 0^+$, the same modified log-Sobolev inequality as μ:

$$\mathrm{Ent}_{\mu*\gamma_\varepsilon}\bigl(e^{\varphi}\bigr) \le \bigl(1 + A\varepsilon^\alpha\bigr)\int_{\mathbb{R}} H\bigl(c\varphi'\bigr)e^{\varphi}\,d\mu*\gamma_\varepsilon.$$

The reader will easily check that the proof below shows that $\mu*\gamma_\varepsilon$ satisfies (i) with $a_\varepsilon = ((\alpha-1)/A_\varepsilon)^{1/\alpha}c^{-1}$, where $A_\varepsilon = A\cdot(1 + A\varepsilon^\alpha)$ (the multiplicative constant $1 + A\varepsilon^\alpha$ will appear in one place in the estimate of $F'(t)$). By the Lebesgue dominated convergence theorem applied to the dual formulation this implies that μ satisfies (i) with $a = ((\alpha-1)/A)^{1/\alpha}c^{-1}$ (note that it suffices to consider the dual formulation for convex Lipschitz functions only and that we use the estimate

$$Q_1^{H^*(a_\varepsilon\cdot)}f(x) \le f(x) \le f(0) + \mathrm{Lip}(f)|x|$$

and the assumption that $\int_{\mathbb{R}} e^{s|x|}\,d\mu(x) < \infty$ for every $s > 0$).

Note that if μ is absolutely continuous, then standard approximation shows that (ii) holds for all convex Lipschitz functions (by the Rademacher theorem the gradient is then almost surely well defined).

Denote for brevity

$$Q_t f(x) = Q_t^{H^*(\cdot)}f(x) = \inf_{y\in\mathbb{R}}\left\{f(y) + tH^*\left(\frac{|x-y|}{t}\right)\right\}$$

and set $F(t) = \int_{\mathbb{R}} e^{k(t)Q_t\varphi(x)}\,d\mu(x)$ for $t > 0$ (for some non-decreasing function k yet to be determined). Using first the fact that $\partial_t Q_t\varphi + H(\partial_x Q_t\varphi) = 0$ almost surely on $(0,\infty)\times\mathbb{R}$ (see, e.g., Chapter 3 of [13]), then the log-

Sobolev inequality, and finally the estimate $H(ck(t)\cdot) \le Ac^\alpha k(t)^\alpha H(\cdot)$ which follows from the assumption (3.1) if only $ck(t) \le 1$, we arrive at

$$\begin{aligned}
k(t)F'(t) &= k(t)\int_{\mathbb{R}} e^{k(t)Q_t\varphi(x)}\big(k'(t)Q_t\varphi(x) + k(t)\partial_t Q_t\varphi(x)\big)\,d\mu(x)\\
&= k(t)\int_{\mathbb{R}} e^{k(t)Q_t\varphi(x)}\big(k'(t)Q_t\varphi(x) - k(t)H\big(\partial_x Q_t\varphi(x)\big)\big)\,d\mu(x)\\
&= k'(t)F(t)\log F(t) + k'(t)\,\mathrm{Ent}_\mu\big(e^{k(t)Q_t\varphi}\big)\\
&\quad - k^2(t)\int_{\mathbb{R}} e^{k(t)Q_t\varphi(x)} H\big(\partial_x Q_t\varphi(x)\big)\,d\mu(x)\\
&\le k'(t)F(t)\log F(t) + k'(t)\int_{\mathbb{R}} e^{k(t)Q_t\varphi(x)} H\big(ck(t)\partial_x Q_t\varphi(x)\big)\,d\mu(x)\\
&\quad - k^2(t)\int_{\mathbb{R}} e^{k(t)Q_t\varphi(x)} H\big(\partial_x Q_t\varphi(x)\big)\,d\mu(x)\\
&\le k'(t)F(t)\log F(t) + \big[Ac^\alpha k'(t)k(t)^\alpha - k^2(t)\big]\cdot\int_{\mathbb{R}} e^{k(t)Q_t\varphi(x)} H\big(\partial_x Q_t\varphi(x)\big)\,d\mu(x).
\end{aligned}$$

Denote $\widetilde{A} = A^{1/\alpha}$ and take

$$k(t) = (\widetilde{A}c)^{-\alpha/(\alpha-1)}\big((\alpha-1)t\big)^{1/(\alpha-1)}.$$

Then $k(0)=0$, $ck(t) \le 1$ for $t \in [0, \widetilde{A}^\alpha c/(\alpha-1)]$, and $Ac^\alpha k'(t)k(t)^\alpha - k^2(t) = 0$. Thus the above differential inequality is equivalent to $(\log(F(t))/k(t))' \le 0$ for almost all $t \in (0, \widetilde{A}^\alpha c/(\alpha-1))$, which, since $Q_t\varphi \le \varphi$, yields

$$\frac{\log F(t)}{k(t)} \le \liminf_{s\to 0^+}\frac{\log F(s)}{k(s)} \le \lim_{s\to 0^+}\frac{\log(\int_{\mathbb{R}} e^{k(s)\varphi(x)}\,d\mu(x))}{k(s)} = \int_{\mathbb{R}}\varphi\,d\mu$$

for $t \in (0, \widetilde{A}^\alpha c/(\alpha-1))$. This is exactly the dual formulation of $\overline{\mathbf{T}}^-(tk(t)H^*(\cdot/t))$ (see Proposition 2.1(a)). Taking $t = t_* = \widetilde{A}c/(\alpha-1)^{1/\alpha}$ we see that $t_*k(t_*) = 1$, $t_* \le \widetilde{A}^\alpha c/(\alpha-1)$ (recall that $A \ge 1$ and $1 < \alpha \le 2$), and thus also $ck(t_*) \le 1$. We conclude that μ satisfies $\overline{\mathbf{T}}^-(H^*(a\cdot))$ with $a = 1/t_* = ((\alpha-1)/A)^{1/\alpha}c^{-1}$. □

4. Equivalence of the weak transportation inequality and the condition on U_μ

In the previous section we showed the equivalence of the convex log-Sobolev inequality and the weak transport-entropy inequality. In this section, working towards the proof of the main theorem, we deal with weak transport-entropy inequalities.

Throughout this section let μ be a measure on the real line (which is not a Dirac mass) with median $m = F_\mu^{-1}(1/2)$. Denote $s_\mu = \inf \mathrm{Supp}(\mu) \in [-\infty, \infty)$, $t_\mu = \sup \mathrm{Supp}(\mu) \in (-\infty, \infty]$.

Let $\theta : [0,\infty) \to [0,\infty)$ be a convex cost function such that $\theta(t) = t^2$ for $t \in [0, t_0]$ for some $t_0 > 0$. Note that by convexity θ is increasing. We moreover assume that $\int_0^\infty \theta(x)e^{-\lambda x}\,dx < \infty$ for any $\lambda > 0$. Recall that the following result is proved in [18].

Proposition 4.1 ([18, Theorem 1.3]). *The following conditions are equivalent.*

(i) *There exists $a > 0$ such that μ satisfies $\overline{\mathbf{T}}(\theta(a\cdot))$.*
(ii) *There exists $b > 0$ such that for all $h > 0$ we have*

$$\Delta_\mu(h) \le \frac{1}{b}\theta^{-1}\big(t_0^2 + h\big).$$

The dependence of the constants is the following: (i) *implies* (ii) *with $b = \kappa_1 a$ and* (ii) *implies* (i) *with $a = \kappa_2 b$, where* $\kappa_1 = \frac{t_0}{8\theta^{-1}(\log(3)+t_0^2)}$, $\kappa_2 = \frac{\min(1,t_0)}{210\theta^{-1}(2+t_0^2)}$.

The goal of this section is to provide a proof of the following stronger version of the above proposition, where $\overline{\mathbf{T}}(\theta(a\cdot))$ is replaced by the formally weaker inequality $\overline{\mathbf{T}}^-(\theta(a\cdot))$. Note that this in particular means that the inequalities $\overline{\mathbf{T}}$ and $\overline{\mathbf{T}}^-$ are equivalent.

Proposition 4.2. *The following conditions are equivalent.*

(i) *There exists $a > 0$ such that μ satisfies $\overline{\mathbf{T}}^-(\theta(a\cdot))$.*
(ii) *There exists $b > 0$ such that for all $h > 0$ we have*

$$\Delta_\mu(h) \le \frac{1}{b}\theta^{-1}\big(t_0^2 + h\big).$$

The dependence of the constants is the following: (i) *implies* (ii) *with*

$$b = \frac{\min(a, 1)}{16}\left(1 + \frac{1}{at_0}\theta^{-1}\left(\frac{\log(2e^{C_\theta/2} - 1)}{2}\right)\right)^{-1},$$

where $C_\theta = \int_0^\infty \theta(2 + \frac{1}{\log 2}t)e^{-t}\,dt$, and (ii) *implies* (i) *with $a = \kappa b$, where $\kappa = \frac{\min(1,t_0)}{210\theta^{-1}(2+t_0^2)}$.*

For the proof we need the following lemma which is an immediate consequence of Theorem 2.2 from [17] (cf. [18, Theorem 6.1]), where the connection between the condition (ii) satisfied by the map U_μ and transport-entropy inequalities connected to transport cost which are equal to zero in a neighborhood of zero is explained in detail. In what follows the symbol $\int_{t_1}^{t_2}$ always denotes an integral over the *open* interval (t_1, t_2).

Lemma 4.3 ([17, Theorem 2.2]). *Let $\beta : [0, \infty) \to [0, \infty)$ be a function which is equal to zero on the interval* $[0, t_0]$ *and then strictly increasing*; *denote its inverse by $\beta^{-1} : [0, \infty) \to [t_0, \infty)$. The following conditions are equivalent.*

(i) *There exists $d > 0$ such that for all $h > 0$ and $x \in \mathbb{R}$,*

$$\Delta_\mu(h) \le \frac{1}{d}\beta^{-1}(h).$$

(ii) *There exist $k > 0$, $K < \infty$ such that*

$$\sup_{x\in[m,t_\mu)} \frac{1}{\mu((x,\infty))}\int_x^\infty \exp\big(\beta\big(k(u-x)\big)\big)\mu(du) \le K,$$
$$\sup_{x\in(s_\mu,m]} \frac{1}{\mu((-\infty,x))}\int_{-\infty}^x \exp\big(\beta\big(k(x-u)\big)\big)\mu(du) \le K.$$

The dependence of the constants is the following: (i) *implies* (ii) *with $K = 3$ and $k = d\frac{t_0}{18\beta^{-1}(2)}$*; (ii) *implies* (i) *with $d = k\frac{t_0}{4\beta^{-1}(\log K)}$.*

We also need the following preparatory result. Note that for the assertion to hold it suffices to assume only that μ satisfies the convex Poincaré inequality.

Lemma 4.4. *If μ satisfies $\overline{\mathbf{T}}^-(\theta(a\cdot))$, then*

$$\frac{1}{\mu((x,\infty))}\int_x^\infty \theta\big(a(u-x)\big)\mu(du) \le C_\theta \quad \textit{for } x \in [m, t_\mu),$$
$$\frac{1}{\mu((-\infty,x))}\int_{-\infty}^x \theta\big(a(x-u)\big)\mu(du) \le C_\theta \quad \textit{for } x \in (s_\mu, m],$$

where $C_\theta = \int_0^\infty \theta(2 + \frac{1}{\log 2}t)e^{-t}\,dt$. Moreover, if $\theta(x) = x^2$ then one can choose $C_\theta = 1$.

Proof of Lemma 4.4. First note that μ satisfies the convex Poincaré inequality: for any (smooth) convex Lipschitz $f:\mathbb{R}\to\mathbb{R}$ we have

$$\operatorname{Var}_\mu(f) \le \frac{1}{2a^2}\int_{\mathbb{R}} |f'|^2\,d\mu. \tag{4.1}$$

This follows by a standard Taylor expansion argument from the dual formulation of the transport entropy inequality (see Proposition 2.1(a)): we plug in εf instead of f, use the estimate

$$Q_1^\theta(\varepsilon f)(x) = \inf_{y\in\mathbb{R}}\bigl\{\varepsilon f(x-y)+\theta\bigl(|y|\bigr)\bigr\} \ge \varepsilon f(x) + \inf_{y\in\mathbb{R}}\bigl\{-\varepsilon f'(x)y+\theta\bigl(|y|\bigr)\bigr\} \ge \varepsilon f(x) - \theta^*\bigl(\varepsilon\bigl|f'(x)\bigr|\bigr)$$

(valid for convex functions), and take $\varepsilon\to 0^+$ (alternatively, one could use Proposition 3.1 and deduce the Poincaré inequality with a slightly worse constant from the log-Sobolev inequality by a similar argument).

We only need to prove the first inequality (where $x\in[m,t_\mu)$), the second one can be taken care of in a similar way.

First, we deal with the case $\theta(x) := x^2$. Inequality (4.1) implies that

$$A - B \le \frac{1}{2}\mu\bigl((x,\infty)\bigr),$$

where $A=\int_x^\infty (a(u-x))^2\mu(du)$ and $B=(\int_x^\infty a(u-x)\mu(du))^2$. (This is obtained by testing (4.1) with $u\mapsto a\max\{u-x,0\}$ – note that even though μ can have atoms there are no problems with non-differentiability at $u=x$, see [17, Proposition 4.6].) By the Cauchy–Schwarz inequality $B\le A\mu((x,\infty))$ and thus

$$A \le \frac{1}{2}\mu\bigl((x,\infty)\bigr) + A\mu\bigl((x,\infty)\bigr),$$

which, since $x\ge m$, leads to

$$\frac{1}{\mu((x,\infty))}\int_x^\infty \theta\bigl(a(u-x)\bigr)\mu(du) \le \frac{1}{2(1-\mu((x,\infty)))} \le 1.$$

Now we turn to the general θ. By the characterization of Bobkov and Götze [10, Theorem 4.2] (cf. [18, Theorem 1.5]), there exist $D_1, D_2>0$ such that $\Delta_\mu(h)\le D_1+D_2h$ for all $h\ge 0$. Following the proof of [14, Theorem 1] (see Lemma 4.5 below for more details), we see that one can choose $D_1=\frac{2}{a}$ and $D_2=\frac{1}{a\log 2}$.

Fix $x\ge m$ and define $v := \sup\{u : U_\mu(u)\le x\}$. Since the map U_μ is left-continuous we have $U_\mu(v)\le x < U_\mu(v+\varepsilon)$ for any $\varepsilon>0$; also $v\ge 0$ since $x\ge m$. Recall that τ denotes the exponential measure and that

$$\int_{\mathbb{R}} f\,d\mu = \int_{\mathbb{R}} f\circ U_\mu\,d\tau$$

for any measurable function f. Thus (note that $\mu((U_\mu(v),x])=0$ if we have $U_\mu(v)<x$)

$$\begin{aligned}
\frac{1}{\mu((x,\infty))}\int_x^\infty \theta\bigl(a(u-x)\bigr)\mu(du) &= \frac{1}{\tau((v,\infty))}\int_v^\infty \theta\bigl(a\bigl(U_\mu(u)-x\bigr)\bigr)\tau(du)\\
&\le e^v\int_v^\infty \theta\bigl(a\bigl(U_\mu(u)-U_\mu(v)\bigr)\bigr)e^{-u}\,du\\
&\le \int_v^\infty \theta\bigl(a\bigl(D_1+D_2(u-v)\bigr)\bigr)e^{-(u-v)}\,du\\
&= \int_0^\infty \theta\bigl(a(D_1+D_2t)\bigr)e^{-t}\,dt\\
&= \int_0^\infty \theta\Bigl(2+\frac{1}{\log 2}t\Bigr)e^{-t}\,dt = C_\theta < \infty,
\end{aligned}$$

where the last inequality follows from the integrability condition placed on θ. $\square$

Let us clarify here what we mean above by "Following the proof of [14, Theorem 1]", especially since in the proof of [14, Theorem 1] the reference measure μ is assumed to be symmetric and the function U_μ does not appear explicitly.

Lemma 4.5. *If μ satisfies the convex Poincaré inequality, then we have $U_\mu(x+h)-U_\mu(x)\le \frac{4}{a}+\frac{1}{a\log(2)}h$. However, the constant $4/a$ may be replaced by $2/a$ if we know that x and $x+h$ are of the same sign. Thus in the above proof one can take $D_1=\frac{2}{a}$ and $D_2=\frac{1}{a\log 2}$.*

Proof. Let X, X' be two independent random variables with distribution μ. Fix $u\ge m$ and plug the function $f(x)=\max\{x-u,0\}$ into (4.1):

$$\begin{aligned}\frac{1}{2a^2}\mu\bigl([u,\infty)\bigr)\ge \operatorname{Var}_\mu(f)&=\frac{1}{2}\mathbb{E}\bigl(f(X)-f\bigl(X'\bigr)\bigr)^2\\&\ge\frac{1}{2}\mathbb{E}\bigl(f(X)-f\bigl(X'\bigr)\bigr)^2\bigl[1_{\{X'\le m\}}1_{\{X\ge u+2/a\}}+1_{\{X\le m\}}1_{\{X'\ge u+2/a\}}\bigr]\\&\ge\frac{1}{2}\mathbb{E}f(X)^2 1_{\{X\ge u+2/a\}}\ge\frac{2}{a^2}\mu\bigl([u+2/a,\infty)\bigr).\end{aligned}$$

Thus,

$$\mu\bigl([u+2/a,\infty)\bigr)\le\frac{1}{4}\mu\bigl([u,\infty)\bigr),\quad u\ge m,$$

and similarly

$$\mu\bigl((-\infty,u-2/a]\bigr)\le\frac{1}{4}\mu\bigl((-\infty,u]\bigr),\quad u\le m.$$

By the definition of U_μ (recall that this function is left-continuous) for $x\in\mathbb{R}$ and all $\varepsilon>0$,

$$F_\mu\bigl(U_\mu(x)-\varepsilon\bigr)\le F_\tau(x)\le F_\mu\bigl(U_\mu(x)\bigr).$$

Denote now $h_0=2\ln(2)$ and let $x, x+h_0\le 0$. Then, for $\varepsilon>0$,

$$\begin{aligned}\mu\bigl(\bigl(-\infty,U_\mu(x+h_0)-2/a-\varepsilon\bigr]\bigr)&\le\frac{1}{4}\mu\bigl(\bigl(-\infty,U_\mu(x+h_0)-\varepsilon\bigr]\bigr)\\&\le e^{-h_0}\tau\bigl((-\infty,x+h_0]\bigr)=\tau\bigl((-\infty,x]\bigr)\\&\le\mu\bigl(\bigl(-\infty,U_\mu(x)\bigr]\bigr).\end{aligned}$$

Hence $U_\mu(x+h_0)-U_\mu(x)\le 2/a$ for $x, x+h_0\le 0$ (since $\varepsilon>0$ was arbitrary). Similarly $U_\mu(x+h_0)-U_\mu(x)\le 2/a$ for $x, x+h_0\ge 0$, since

$$\begin{aligned}\mu\bigl(\bigl[U_\mu(x)+2/a+\varepsilon,\infty\bigr)\bigr)&\le\frac{1}{4}\mu\bigl(\bigl[U_\mu(x)+\varepsilon,\infty\bigr)\bigr)\le\frac{1}{4}\bigl(1-F_\mu\bigl(U_\mu(x)\bigr)\bigr)\\&\le e^{-h_0}\bigl(1-F_\tau(x)\bigr)\le 1-F_\tau(x+h_0)\\&\le\mu\bigl(\bigl[U_\mu(x+h_0)-\varepsilon,\infty\bigr)\bigr).\end{aligned}$$

Using these inequalities in a telescoping manner at most $\lceil h/h_0\rceil\le 1+h/h_0$ times we conclude that

$$U_\mu(x+h)-U_\mu(x)\le\frac{2}{a}+\frac{1}{a\log(2)}h$$

for any $x\in\mathbb{R}$ and $h\ge 0$ such that x and $x+h$ are of the same sign . If $x<0<x+h$, then the additive constant $2/a$ in the last estimate has to be replaced by $4/a$. □

Now we are ready to prove Proposition 4.2.

Proof of Proposition 4.2. Due to Proposition 4.1 we only need to check that if μ satisfies the inequality $\overline{\mathbf{T}}^{-}(\theta(a\cdot))$, then the condition from (ii) is satisfied by U_μ.

Fix $x > m$ and consider the function $f(t) = \theta(a[t-x]_+)$. Then $Q_1^{\theta(a\cdot)} f(t) = 0$ if $t \le x$. For $t > x$,

$$Q_1^{\theta(a\cdot)} f(t) = \inf_{y\in\mathbb{R}} \bigl\{\theta\bigl(a[y-x]_+\bigr) + \theta\bigl(a|t-y|\bigr)\bigr\} = \inf_{y\in[x,t]} \bigl\{\theta\bigl(a(y-x)\bigr) + \theta\bigl(a(t-y)\bigr)\bigr\} = 2\theta\bigl(a(t-x)/2\bigr),$$

where the last equality follows from the fact that we have an inequality due to the convexity of θ and on the other hand the infimum is attained at $y = (x+t)/2$. Hence the dual formulation of the weak transport inequality (see Proposition 2.1(a)) implies that

$$\mu\bigl((-\infty,x]\bigr) + \int_x^\infty \exp\bigl(2\theta\bigl(a(t-x)/2\bigr)\bigr)\mu(dt) \le \exp\biggl(\int_x^\infty \theta\bigl(a(t-x)\bigr)\mu(dt)\biggr).$$

Denote $k = 1/2$ and $\beta(u) = 2\theta(a[u-t_0]_+)$ for $u > 0$. Since θ is increasing we have $\beta(ku) \le 2\theta(au/2)$. Therefore,

$$\int_x^\infty \exp\bigl(\beta\bigl(k(t-x)\bigr)\bigr)\mu(dt) \le \exp\biggl(\int_x^\infty \theta\bigl(a(t-x)\bigr)\mu(dt)\biggr) - 1 + \mu\bigl((x,\infty)\bigr). \tag{4.2}$$

By Lemma 4.4 there exists $C_\theta < \infty$ such that $\int_x^\infty \theta(a(t-x))\mu(dt) \le C_\theta \mu((x,\infty))$. Hence, since $\mu((x,\infty)) \in [0,1/2]$,

$$\frac{\int_x^\infty \exp(\beta(k(t-x)))\mu(dt)}{\mu((x,\infty))} \le \frac{\exp(C_\theta\mu((x,\infty))) - 1 + \mu((x,\infty))}{\mu((x,\infty))} \le 2e^{C_\theta/2} - 1.$$

One can deal with $x \le m$ similarly. Since

$$\beta^{-1}(h) = \biggl(t_0 + \frac{1}{a}\theta^{-1}(h/2)\biggr),$$

Lemma 4.3 implies that

$$\begin{aligned}\Delta_\mu(h) &\le \frac{1}{d}\beta^{-1}(h) = \frac{1}{d}\biggl(t_0 + \frac{1}{a}\theta^{-1}(h/2)\biggr) \le \frac{1}{d\min(a,1)}\bigl(t_0 + \theta^{-1}(h/2)\bigr) \\ &\le \frac{2}{d\min(a,1)}\theta^{-1}\bigl(t_0^2 + h\bigr),\end{aligned} \tag{4.3}$$

where

$$d = \frac{t_0}{8\beta^{-1}(\log(2e^{C_\theta/2}-1))} = \frac{t_0}{8(t_0 + \frac{1}{a}\theta^{-1}(\frac{1}{2}\log(2e^{C_\theta/2}-1)))}$$

(recall that $k = 1/2$). This finishes the proof of the implication (i) $\Longrightarrow$ (ii). $\square$

5. Summary

5.1. *Proof of the main result and dependence of constants for* $H(x) = \frac{1}{4}x^2$

The results of the two preceding sections allow us to prove Theorem 1.1. The proof of Corollary 1.7 is postponed to the next subsection.

Proof of Theorem 1.1. The implication (ii) $\Longrightarrow$ (i) has been proved in [18]. The implication (i) $\Longrightarrow$ (ii) follows immediately by combining Propositions 3.1 and 4.2. The only assumption we need to check is that $\int_0^\infty \theta(x)e^{-\lambda x}\,dx < \infty$ for any $\lambda > 0$, but this follows from the scaling condition placed on H. Indeed, for $s \in (0,1]$,

$$H^*(y/s) = \big(H(s\cdot)\big)^*(y) \geq \big(As^\alpha H(\cdot)\big)^*(y) = As^\alpha H^*\big(y/\big(As^\alpha\big)\big).$$

Taking $z \geq 1$ and substituting into the above inequality $s = z^{-1/(\alpha-1)}$ and $y = As = Az^{-1/(\alpha-1)}$ we arrive at $\theta(z) = H^*(z) \leq H^*(A)A^{-1}z^{\alpha/(\alpha-1)}$, which implies the claim. $\square$

As for the dependence of constants, in the case $H(x) = \frac{1}{4}x^2$ one can take $A = 1$ and $\alpha = 2$ in (1.5). Let us consider the implication (i) $\Longrightarrow$ (ii) from Theorem 1.1. In Proposition 3.1 we have $a = 1/c$ and moreover we can take $C_\theta = 1$ in Lemma 4.4. Therefore, inequality (4.3) reads

$$\Delta_\mu(h) \leq \frac{1}{d}(t_0 + c\sqrt{h/2}) \leq 8\frac{t_0 + \frac{2}{3}c}{t_0}(t_0 + c\sqrt{h/2}), \tag{5.1}$$

since

$$d = \frac{t_0}{8(t_0 + \frac{1}{a}\sqrt{\frac{1}{2}\log(2e^{1/2} - 1)})} \geq \frac{t_0}{8(t_0 + \frac{2}{3}c)}.$$

Taking $t_0 = \frac{2}{3}c$ we obtain the result announced in Remark 1.2 (the dependence of constants for the implication (ii) $\Longrightarrow$ (i) follows directly from Proposition 3.1 and Proposition 4.1). In fact, we can take $t_0 = c\sqrt[4]{2h/9}$ (which minimizes the right-hand side of (5.1)) to obtain a slightly better estimate

$$\Delta_\mu(h) \leq 8c\big(2/3 + \sqrt{h/2} + 2\sqrt[4]{2h/9}\big).$$

5.2. *Conditions equivalent to the convex log-Sobolev inequality*

Taking into account the results from [18], we can state a handful of conditions equivalent to the convex log-Sobolev inequality on the real line. For simplicity we work with the quadratic cost.

Theorem 5.1. *Let $\theta(t) = t^2$ for $t \geq 0$. For a probability measure μ on the real line the following conditions are equivalent.*

(i) *For every $s > 0$ we have $\int_{\mathbb{R}} e^{s|x|}\,d\mu(x) < \infty$ and there exists $C > 0$ such that*

$$\mathrm{Ent}_\mu\big(e^\varphi\big) \leq C\int_{\mathbb{R}} \big|\varphi'\big|^2 e^\varphi\,d\mu$$

for every smooth convex Lipschitz function $\varphi : \mathbb{R} \to \mathbb{R}$.

(ii) *There exist $a, b > 0$ such that for all $h > 0$,*

$$\Delta_\mu(h) \leq \sqrt{a + bh}.$$

(iii) *There exists $a_1 > 0$ such that μ satisfies the inequality $\overline{\mathbf{T}}^-(\theta(a_1\cdot))$.*

(iv) *There exists $a_2 > 0$ such that μ satisfies the inequality $\overline{\mathbf{T}}(\theta(a_2\cdot))$.*

(v) *There exists $t > 0$ such that μ satisfies the infimum convolution inequality*

$$\int_{\mathbb{R}} \exp\big(Q_t^\theta f\big)\,d\mu \int_{\mathbb{R}} \exp(-f) \leq 1$$

for every convex function $f : \mathbb{R} \to \mathbb{R}$ bounded from below.

In each of the implications the constants in the conclusion depend only on the constants in the premise.

With this theorem we can easily give the proof of the concentration inequalities from Corollary 1.7.

Proof of Corollary 1.7. By Theorem 5.1 the measure μ satisfies the inequality $\overline{\mathbf{T}}$ for the quadratic cost. An application of, e.g., [19, Corollary 5.11] completes the proof. Alternatively, for a more self-contained reasoning, one can use item (v) of the above theorem and adapt the approach of [21]. □

5.3. *Relation to Talagrand's inequality*

Let $\theta(t) = t^2$ for $t \geq 0$. We say that a probability measure μ on the real line satisfies Talagrand's inequality (with constant C) if

$$\mathcal{T}_\theta(\mu, \nu) \leq C H(\nu|\mu)$$

for every probability measure ν, where $\mathcal{T}_\theta$ was defined in (2.1). Recall that in the classical setting of smooth functions we have the implication chain

$$\text{log-Sobolev inequality} \implies \text{Talagrand's inequality} \implies \text{Poincaré inequality}$$

and these implications are strict (see [17, Section 4.3] for a nice discussion of counterexamples). From [18] we also know that Talagrand's inequality is strictly stronger than the convex log-Sobolev inequality. The following corollary explains what additional information is carried by the Talagrand inequality. It is an immediate consequence of Theorem 1.1 and [17, Theorem 1.1].

Corollary 5.2. *A probability measure μ on the real line satisfies the Talagrand inequality if and only if it satisfies the Poincaré inequality for smooth functions and the log-Sobolev inequality for convex functions.*

5.4. *Open questions*

We conclude with three open questions, which to the best of our knowledge are open even in the case $\theta(t) = t^2$. They address the possibility of extending some of the results of this paper.

Roughly speaking, Theorem 1.1 together with the results from [19] implies that if a measure μ on the real line satisfies the log-Sobolev inequality for convex functions (i.e. the inequality $\overline{\mathbf{T}}^-$), then it also satisfies the log-Sobolev inequality for (almost all) concave functions (i.e. the inequality $\overline{\mathbf{T}}^+$). Does the reverse implication hold true? Which of those implications hold in a higher dimensional space?

In terms of the weak transport-entropy inequalities the above questions read:

1. Suppose that a probability measure μ on $\mathbb{R}^n$, $n \geq 2$, satisfies the inequality $\overline{\mathbf{T}}^-(\theta(a\cdot))$ for some $a > 0$. Does it satisfy the inequality $\overline{\mathbf{T}}(\theta(a'\cdot))$ for some $a' > 0$?
2. Suppose that a probability measure μ on the real line satisfies the inequality $\overline{\mathbf{T}}^+(\theta(a\cdot))$ for some $a > 0$. Does it satisfy the inequality $\overline{\mathbf{T}}^-(\theta(a'\cdot))$, and thus $\overline{\mathbf{T}}(\theta(\min\{a', a\}\cdot))$, for some $a' > 0$?

We refer to [19, Theorem 8.15] and [19, Remark 8.18] for details and subtleties concerning the log-Sobolev inequality for concave functions. We also point out that in the case of the quadratic-linear cost a partial answer to the above two questions has been recently obtained by Adamczak and the second named author [2].

The last question is deliberately somewhat vague. An intuitive way to understand Theorem 1.1 is the following: on the real line one can obtain a measure which satisfies the convex log-Sobolev inequality by a local perturbation of another measure which satisfies this inequality. (Indeed, if μ is a measure on the real line which satisfies the convex log-Sobolev inequality and $\widetilde{\mu}$ coincides with μ on the complement of a bounded inteval, then, using the condition (ii) from Theorem 1.1, it is not hard to see that $\widetilde{\mu}$ also satisfies the convex log-Sobolev inequality.) Now take a measure satisfying the convex log-Sobolev inequality in $\mathbb{R}^n$ and rearrange the mass locally. Does the convex log-Sobolev inequality still hold? More generally, suppose that a probability measure μ on $\mathbb{R}^n$, $n \geq 2$, satisfies the convex log-Sobolev inequality. Can one state this fact with a condition which is not expressed in terms of quantifiers ranging over families of functions or measures (like, e.g., the conditions in Section 3) but rather in terms of the measure μ itself (like the condition (ii) in Theorem 1.1)? To put it briefly:

3. Find an intrinsic characterization of probability measures on $\mathbb{R}^n$, $n \geq 2$, which satisfy the convex log-Sobolev inequality.

Acknowledgements

We thank Radosław Adamczak for numerous conversations and many helpful comments, especially for pointing out that the approach of [21] can be adapted to the setting of Lemma 3.2. We also thank two Anonymous Referees for a careful reading of the first version of this article and for many helpful comments.

References

[1] R. Adamczak. Logarithmic Sobolev inequalities and concentration of measure for convex functions and polynomial chaoses. *Bull. Pol. Acad. Sci. Math.* **53** (2) (2005) 221–238. MR2163396
[2] R. Adamczak and M. Strzelecki. On the convex Poincaré inequality and weak transportation inequalities. *Bernoulli*. To appear, 2018. Available at arXiv:1703.01765v2.
[3] R. Adamczak and M. Strzelecki. Modified log-Sobolev inequalities for convex functions on the real line. Sufficient conditions. *Studia Math.* **230** (1) (2015) 59–93. MR3456588
[4] C. Ané, S. Blachère, D. Chafaï, P. Fougères, I. Gentil, F. Malrieu, C. Roberto and G. Scheffer. *Sur les inégalités de Sobolev logarithmiques*. *Panoramas et Synthèses [Panoramas and Syntheses]* **10**. Société Mathématique de France, Paris, 2000. With a preface by Dominique Bakry and Michel Ledoux. MR1845806
[5] D. Bakry, I. Gentil and M. Ledoux. *Analysis and Geometry of Markov Diffusion Operators*. *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]* **348**. Springer, Cham, 2014. MR3155209
[6] F. Barthe and C. Roberto. Sobolev inequalities for probability measures on the real line. *Studia Math.* **159** (3) (2003) 481–497. Dedicated to Professor Aleksander Pełczyński on the occasion of his 70th birthday (Polish). MR2052235
[7] F. Barthe and C. Roberto. Modified logarithmic Sobolev inequalities on $\mathbb{R}$. *Potential Anal.* **29** (2) (2008) 167–193. MR2430612
[8] S. Bobkov and M. Ledoux. Poincaré's inequalities and Talagrand's concentration phenomenon for the exponential distribution. *Probab. Theory Related Fields* **107** (3) (1997) 383–400. MR1440138
[9] S. G. Bobkov, I. Gentil and M. Ledoux. Hypercontractivity of Hamilton–Jacobi equations. *J. Math. Pures Appl. (9)* **80** (7) (2001) 669–696. MR1846020
[10] S. G. Bobkov and F. Götze. Discrete isoperimetric and Poincaré-type inequalities. *Probab. Theory Related Fields* **114** (2) (1999) 245–277. MR1701522
[11] S. G. Bobkov and F. Götze. Exponential integrability and transportation cost related to logarithmic Sobolev inequalities. *J. Funct. Anal.* **163** (1) (1999) 1–28. MR1682772
[12] P. Diaconis and L. Saloff-Coste. Logarithmic Sobolev inequalities for finite Markov chains. *Ann. Appl. Probab.* **6** (3) (1996) 695–750. MR1410112
[13] L. C. Evans. *Partial Differential Equations*, 2nd edition. *Graduate Studies in Mathematics* **19**. American Mathematical Society, Providence, RI. MR2597943
[14] N. Feldheim, A. Marsiglietti, P. Nayar and J. Wang. A note on the convex infimum convolution inequality. *Bernoulli* **24** (1) (2018) 257–270. MR3706756
[15] I. Gentil, A. Guillin and L. Miclo. Modified logarithmic Sobolev inequalities and transportation inequalities. *Probab. Theory Related Fields* **133** (3) (2005) 409–436. MR2198019
[16] I. Gentil, A. Guillin and L. Miclo. Modified logarithmic Sobolev inequalities in null curvature. *Rev. Mat. Iberoam.* **23** (1) (2007) 235–258. MR2351133
[17] N. Gozlan. Transport-entropy inequalities on the line. *Electron. J. Probab.* **17** (2012) Paper no. 49. MR2946156
[18] N. Gozlan, C. Roberto, P. Samson, Y. Shu and P. Tetali. Characterization of a class of weak transport-entropy inequalities on the line. *Ann. Inst. Henri Poincaré Probab. Stat.* **54** (2018) 1667–1693. MR3825894
[19] N. Gozlan, C. Roberto, P.-M. Samson and P. Tetali. Kantorovich duality for general transport costs and applications. *J. Funct. Anal.* **273** (11) (2017) 3327–3405. MR3706606
[20] L. Gross. Logarithmic Sobolev inequalities. *Amer. J. Math.* **97** (4) (1975) 1061–1083. MR0420249
[21] B. Maurey. Some deviation inequalities. *Geom. Funct. Anal.* **1** (2) (1991) 188–197. MR1097258
[22] F. Otto and C. Villani. Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality. *J. Funct. Anal.* **173** (2) (2000) 361–400. MR1760620
[23] F. Otto and C. Villani. Comment on: "Hypercontractivity of Hamilton–Jacobi equations" [J. Math. Pures Appl. (9) **80** (2001), no. 7, 669–696; MR1846020 (2003b:47073)] by S. G. Bobkov, I. Gentil and M. Ledoux. *J. Math. Pures Appl. (9)* **80** (7) (2001) 697–700. MR1846021
[24] M. Talagrand. Transportation cost for Gaussian and other product measures. *Geom. Funct. Anal.* **6** (3) (1996) 587–600. MR1392331
[25] C. Villani. *Optimal Transport: Old and New*. *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]* **338**. Springer-Verlag, Berlin, 2009. MR2459454