CHARACTERIZATION OF TALAGRAND'S TRANSPORT-ENTROPY INEQUALITIES IN METRIC SPACES.

Nathaël Gozlan, Cyril Roberto, Paul-Marie Samson

The Annals of Probability
2013, Vol. 41, No. 5, 3112–3139
DOI: 10.1214/12-AOP757

CHARACTERIZATION OF TALAGRAND'S TRANSPORT-ENTROPY INEQUALITIES IN METRIC SPACES[1]

By N. Gozlan, C. Roberto[2] and P.-M. Samson

Université Paris Est Marne la Vallée, Université Paris Est Marne la Vallée and Université Paris Ouest Nanterre la Défense, and Université Paris Est Marne la Vallée

We give a characterization of transport-entropy inequalities in metric spaces. As an application we deduce that such inequalities are stable under bounded perturbation (Holley–Stroock perturbation lemma).

1. Introduction. In their celebrated paper [24], Otto and Villani proved that, in a smooth Riemannian setting, the log-Sobolev inequality implies the Talagrand transport-entropy inequality $\mathbf{T}_2$. Later, Bobkov, Gentil and Ledoux [3] proposed an alternative proof of this result. Both approaches are based on semi-group arguments. More recently, the first named author of this paper gave a new proof, based on large deviation theory, valid on metric spaces [11].

In this paper, on the one hand, we give yet another proof of Otto and Villani's theorem. This proof does not use any semi-group arguments nor large deviations, and it requires very few structures on the space. We are thus able to recover and extend the result of [11] in a general metric space framework.

On the other hand, we recently introduced in [15] a new family of functional inequalities, called inf-convolution log-Sobolev inequalities. In a Euclidean framework, we proved that these inequalities are equivalent to Talagrand transport-entropy inequalities $\mathbf{T}_\alpha$, associated to cost functions α

Received July 2011; revised March 2012.
[1]Supported in part by the "Agence Nationale de la Recherche" through the Grants ANR 2011 BS01 007 01 and ANR 10 LABX-58.
[2]Supported in part by the European Research Council through the "Advanced Grant" PTRELSS 228032.
AMS 2000 subject classifications. 60E15, 26D10.
Key words and phrases. Transport-entropy inequalities, logarithmic-Sobolev inequalities, metric spaces, concentration of measure.

between linear and quadratic. This led to a new characterization of $\mathbf{T}_2$ and other transport-entropy inequalities. The present paper establishes that this equivalence is true in a general metric space framework and for general cost functions α. As a byproduct, we prove that the inequalities $\mathbf{T}_\alpha$ are stable under bounded perturbation (Holley–Stroock perturbation lemma).

Our strategy is very general and applies to a very large class of transport-entropy inequalities.

In order to present our results, we need first to fix some notation.

1.1. *Notation and definitions.* We first introduce the notion of optimal transport cost. Then we give the definition of the transport-entropy inequality and of the (τ)-log-Sobolev inequality.

General assumption. Throughout this paper, (X,d) will always be a complete, separable metric space such that closed balls are compact.

1.1.1. *Optimal transport cost and transport-entropy inequality.* Let $\alpha\colon \mathbb{R}\to\mathbb{R}^+$ be a continuous function. Given two probability measures ν and μ on X, the *optimal transport cost* between ν and μ (with respect to the cost function α) is defined by

$$\mathcal{T}_\alpha(\nu,\mu) := \inf_\pi \left\{ \iint \alpha(d(x,y))\, d\pi(x,y) \right\},$$

where the infimum runs over all the probability measures π on $X\times X$ with marginals ν and μ. The notion of optimal transport cost is very old (it goes back to Monge [23]). It has been intensively studied and it is used in a wide class of problems running from geometry, PDE theory, probability and statistics; see [31]. Here we focus on the following transport-entropy inequality.

Throughout this paper, the cost functions α will be assumed to belong to the class of Young functions.

DEFINITION 1.1 (Young functions[3]). A function $\alpha\colon\mathbb{R}\to\mathbb{R}^+$ is a *Young function* if α is an even, convex, increasing function on $\mathbb{R}^+$ such that $\alpha(0)=0$ and $\alpha'(0)=0$.

DEFINITION 1.2 (Transport-entropy inequality $\mathbf{T}_\alpha$). Let α be a Young function; a probability measure μ on X is said to satisfy the *transport-entropy inequality* ($\mathbf{T}_\alpha(C)$), for some $C>0$ if

$$(\mathbf{T}_\alpha(C)) \qquad \mathcal{T}_\alpha(\nu,\mu) \le C H(\nu|\mu) \qquad \forall \nu\in\mathcal{P}(X),$$

[3]Note that, contrary to the definition of some authors, for us, a Young function cannot take infinite values.

where

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

is the relative entropy of ν with respect to μ, and $\mathcal{P}(X)$ is the set of all probability measures on X.

REMARK 1.3. It can be shown that if $\alpha : \mathbb{R} \to \mathbb{R}^+$ is an even convex function such that $\limsup_{x\to 0} \frac{\alpha(x)}{x^2} = +\infty$, then the only probability measures that satisfy the transport inequality $\mathbf{T}_\alpha$ are Dirac masses; see, for example, [12], Proposition 2. This is the reason why, in our definition of Young functions, we impose that $\alpha'(0) = 0$.

Popular Young functions appearing in the literature, as cost functions in transport-entropy inequalities, are the functions α_{p_1,p_2}, defined by

$$\alpha_{p_1,p_2}(x) := \begin{cases} |x|^{p_1}, & \text{if } |x| \le 1, \\ \dfrac{p_1}{p_2}|x|^{p_2} + 1 - \dfrac{p_1}{p_2}, & \text{if } |x| > 1, \end{cases} \qquad p_1 \ge 2,\ p_2 \ge 1 \tag{1.1}$$

(the case $p_1 < 2$ can be discarded according to the remark above). When $p_1 = p_2 = p$, we use the notation α_p instead of $\alpha_{p,p}$.

Transport-entropy inequalities imply concentration results as shown by Marton [21]; see also [4, 19], and [13] for a full introduction to this notion.

The transport-entropy inequality related to the quadratic cost $\alpha_2(x) = x^2$ is the most studied in the literature. In this case, the transport-entropy inequality is often referred to as the Talagrand transport-entropy inequality and is denoted by $\mathbf{T}_2$. Talagrand [30] proved that, on $(\mathbb{R}^n, |\cdot|_2)$ (where $|\cdot|_2$ stands for the Euclidean norm), the standard Gaussian measure satisfies $\mathbf{T}_2$ with the optimal constant $C = 2$.

1.1.2. *Log-Sobolev-type inequalities.* The second inequality of interest for us is the log-Sobolev inequality and, more generally, modified log-Sobolev inequalities. To define these inequalities properly, we need to introduce additional notation.

Recall that the Fenchel–Legendre transform α^* of a Young function α is defined by

$$\alpha^*(y) = \sup_{x\in\mathbb{R}} \{xy - \alpha(x)\} \in \mathbb{R}^+ \cup \{\infty\} \qquad \forall y \in \mathbb{R}.$$

A function $f : X \to \mathbb{R}$ is said to be *locally Lipschitz* if for all $x \in X$, there exists a ball B centered at point x such that

$$\sup_{y,z\in B,\ y\neq z} \frac{|f(y) - f(z)|}{d(y,z)} < \infty.$$

When f is locally Lipschitz, we define

$$|\nabla^+ f|(x) = \begin{cases} \limsup_{y \to x} \dfrac{[f(y) - f(x)]_+}{d(y,x)}, & \text{if } x \text{ is not an isolated point,} \\ 0, & \text{otherwise,} \end{cases}$$

and

$$|\nabla^- f|(x) = \begin{cases} \limsup_{y \to x} \dfrac{[f(y) - f(x)]_-}{d(y,x)}, & \text{if } x \text{ is not an isolated point,} \\ 0, & \text{otherwise,} \end{cases}$$

where $[a]_+ = \max(a;0)$ and $[a]_- = \max(-a;0)$. Note that $|\nabla^+ f|(x)$ and $|\nabla^- f|(x)$ are finite for all $x \in X$. When f is a smooth function on a smooth manifold, $|\nabla^+ f|$ and $|\nabla^- f|$ equal the norm of the gradient of f.

Finally, if μ is a probability measure on X, recall that the entropy functional $\mathrm{Ent}_\mu(\cdot)$ is defined by

$$\mathrm{Ent}_\mu(g) = \int g \log \frac{g}{\int g \, d\mu} \, d\mu \qquad \forall g > 0.$$

Definition 1.4 (Modified log-Sobolev inequality $\mathbf{LSI}_\alpha^\pm$). Let α be a Young function; a probability measure μ on X is said to satisfy the *modified log-Sobolev inequality plus* ($\mathbf{LSI}_\alpha^+(A)$) for some $A > 0$ if

$$(\mathbf{LSI}_\alpha^+(A)) \qquad \mathrm{Ent}_\mu(e^f) \le A \int \alpha^*(|\nabla^+ f|) e^f \, d\mu$$

for all locally Lipschitz bounded functions $f : X \to \mathbb{R}$.

It verifies the *modified log-Sobolev inequality minus* ($\mathbf{LSI}_\alpha^-(A)$) for some $A > 0$ if

$$(\mathbf{LSI}_\alpha^-(A)) \qquad \mathrm{Ent}_\mu(e^f) \le A \int \alpha^*(|\nabla^- f|) e^f \, d\mu$$

for all locally Lipschitz bounded functions $f : X \to \mathbb{R}$.

Again, the quadratic cost $\alpha_2(x) = x^2$ plays a special role since in this case we recognize the usual log-Sobolev inequality introduced by Gross [16]; see also [28]. In this case, we will use the notation $\mathbf{LSI}^\pm$.

Bobkov and Ledoux [5] introduced first the modified log-Sobolev inequality with the function $\alpha_{2,1}$, in order to recover the celebrated result by Talagrand [29] on the concentration phenomenon for products of exponential measures. In particular these authors proved that, with this special choice of function, the modified log-Sobolev inequality is actually equivalent to the Poincaré inequality. After them, Gentil, Guillin and Miclo [8] established that the probability measure $d\nu_p(x) = e^{-|x|^p}/Z_p$, $x \in \mathbb{R}$ and $p \in (1,2)$ verifies the modified log-Sobolev inequality associated with the function $\alpha_{2,p}$. In a subsequent paper [9], they generalized their results to a large class of

measures with tails between exponential and Gaussian; see also [2, 7, 10] and [25].

Finally, let us introduce the notion of inf-convolution log-Sobolev inequality. In a previous work [15], we proposed the following inequality:

$$\text{(1.2)}\quad \mathrm{Ent}_\mu(e^f) \le \frac{1}{1-\lambda C}\int (f - Q_\alpha^\lambda f)e^f\,d\mu \qquad \forall f : X \to \mathbb{R}, \forall \lambda \in (0, 1/C),$$

where

$$Q_\alpha^\lambda f(x) = \inf_{y\in X}\{f(y) + \lambda\alpha(d(x,y))\} \qquad \forall x \in X.$$

We called it inf-convolution log-Sobolev inequality, and we proved that it is equivalent—in a Euclidean setting—to the transport-entropy inequality $\mathbf{T}_\alpha(C')$, for Young functions α such that α' is concave. Also, we get an explicit comparison between the constants C and C', namely $C \le C' \le 8C$. Our proof relies in part on the Hamilton–Jacobi semi-group approach developed by Bobkov, Gentil and Ledoux [3].

Inequality (1.2) is actually a family of inequalities, with a constant having a specific form [i.e., $1/(1-\lambda C)$] on the right-hand side. In this paper, in order to broaden this notion, we will say (τ)-log-Sobolev inequality, rather than inf-convolution log-Sobolev inequality, in the following inequality.

DEFINITION 1.5 [(τ)-log-Sobolev inequality]. Let α be a Young function; a probability measure μ on X is said to satisfy the (τ)-*log-Sobolev inequality* ($(\tau) - \mathbf{LSI}_\alpha(\lambda, A)$) for some $\lambda, A > 0$ if

$$((\tau) - \mathbf{LSI}_\alpha(\lambda, A)) \qquad \mathrm{Ent}_\mu(e^f) \le A\int (f - Q_\alpha^\lambda f)e^f\,d\mu$$

for all bounded locally Lipschitz functions $f : X \to \mathbb{R}$, where the inf-convolution operator Q_α^λ is defined by

$$\text{(1.3)}\qquad Q_\alpha^\lambda f(x) = \inf_{y\in X}\{f(y) + \lambda\alpha(d(x,y))\} \qquad \forall x \in X.$$

When $\lambda = 1$, we use the notation Q_α instead of Q_α^1.

The notation $(\tau) - \mathbf{LSI}_\alpha$ refers to the celebrated (τ)-Property introduced by Maurey [22] (that uses the inf-convolution operator Q_α and that is also closely related to the transport-entropy inequality; see [13], Section 8.1).

Of course (1.2) implies $(\tau) - \mathbf{LSI}_\alpha(\lambda, 1/(1-\lambda C))$, for any $\lambda \in (0, 1/C)$. The other direction is not clear, a priori (it would trivially be true if $A = 1$), even if the two inequalities have the same flavor. Thanks to Theorem 1.8 below, they appear to be equivalent, under mild assumptions on α.

1.1.3. Δ_2-*condition.* In the next sections, our objective will be to relate the log-Sobolev inequalities $\mathbf{LSI}_\alpha$ and $(\tau) - \mathbf{LSI}_\alpha$ to the transport-entropy inequality $\mathbf{T}_\alpha$. This program works well if we suppose that α verifies the

classical doubling condition Δ_2. Recall that a Young function α is said to satisfy the Δ_2-*condition* if there exists some positive constant K (that must be greater than or equal to 2) such that

$$\alpha(2x) \le K\alpha(x) \qquad \forall x \in \mathbb{R}.$$

The classical functions α_{p_1,p_2} introduced in (1.1) enjoy this condition.

The following observation will be very useful in the sequel.

LEMMA 1.6. *If α is a Young function satisfying the Δ_2-condition, then*

$$r_\alpha := \inf_{x>0} \frac{x\alpha'_-(x)}{\alpha(x)} \ge 1 \quad \text{and} \quad 1 < p_\alpha := \sup_{x>0} \frac{x\alpha'_+(x)}{\alpha(x)} < +\infty, \tag{1.4}$$

where α'_+ (resp., α'_-) denotes the right (resp., left) derivative of α.

The proof of this lemma is in the Appendix. To understand these exponents r_α and p_α, observe that for the function $\alpha = \alpha_{p_1,p_2}$, defined by (1.1), we have $r_\alpha = \min(p_1,p_2)$ and $p_\alpha = \max(p_1,p_2)$. Moreover, if $1 \le r \le p$ are given numbers, and α is a Young function such that $r_\alpha = r$ and $p_\alpha = p$, then it is not difficult to check that

$$\alpha(1)\alpha_{p,r} \le \alpha \le \alpha(1)\alpha_{r,p}.$$

1.2. *Main results.* Our first result states that the modified log-Sobolev inequality (plus or minus) implies the transport-entropy inequality associated with the same α (Otto–Villani theorem).

THEOREM 1.7. *Let μ be a probability measure on X and α a Young function satisfying the Δ_2-condition.*

(i) *If μ satisfies ($\mathbf{LSI}^+_\alpha(A)$) for some $A > 0$, then μ satisfies $\mathbf{T}_\alpha(C^+)$ with*

$$C^+ = \max(((p_\alpha - 1)A)^{r_\alpha - 1}; ((p_\alpha - 1)A)^{p_\alpha - 1}).$$

(ii) *If μ satisfies ($\mathbf{LSI}^-_\alpha(A)$) for some $A > 0$, then μ satisfies $\mathbf{T}_\alpha(C^-)$ with*

$$C^- = (1 + (p_\alpha - 1)A)^{p_\alpha - r_\alpha}((p_\alpha - 1)A)^{r_\alpha - 1}.$$

The numbers $1 \le r_\alpha \le p_\alpha$, $p_\alpha > 1$ are defined by (1.4).

Let us comment on this theorem. First observe that C^+ and C^- are of the same order since

$$C^+ \le C^- \le 2^{p_\alpha - r_\alpha} C^+.$$

For the quadratic case $\alpha_2(x) = x^2$, the constants reduce to $C^+ = C^- = A$. This corresponds (when X is a smooth Riemannian manifold) to the usual Otto–Villani theorem [24]; see also [3]. Let us mention that Lott and Villani [20] generalized the result from Riemannian manifolds to length spaces,

for $\alpha_2(x) = x^2$, with an adaptation of the Hamilton–Jacobi semigroup approach developed by Bobkov, Gentil and Ledoux [3]. But their statement requires additional assumptions, such as a local Poincaré inequality, which are not needed in Theorem 1.7.

Also, in [8] the authors prove that the modified log-Sobolev inequality, in Euclidean setting and with $\alpha = \alpha_{2,p}$, with $1 \le p \le 2$, implies the corresponding transport inequality $\mathbf{T}_\alpha$, again using the Hamilton–Jacobi approach [3].

More recently, in [11], the first named author proved that $\mathbf{LSI}^+(A)$ implies $\mathbf{T}_2(A)$ in the quadratic case $\alpha_2(x) = x^2$ and on an arbitrary complete and separable metric space. His proof can be easily extended to more general functions such as $\alpha_p(x) = x^p$. The scheme of proof is the following. Talagrand's inequality $\mathbf{T}_2$ is first shown to be equivalent to dimension-free Gaussian concentration. According to the well-known Herbst argument (see, e.g., [19]), $\mathbf{LSI}^+$ implies dimension-free Gaussian concentration, so it also implies $\mathbf{T}_2$.

Finally, as shown by Cattiaux and Guillin [6], we mention that the Talagrand transport-entropy inequality $\mathbf{T}_2$ does not imply, in general, the log-Sobolev inequality. Hence, there is no hope to get an equivalence in the above theorem.

However, the (τ)-log-Sobolev inequality appears to be equivalent to the transport-entropy inequality. This is the main result of this paper.

THEOREM 1.8. *Let μ be a probability measure on X, and α a Young function satisfying the Δ_2-condition, and let $p_\alpha > 1$ be defined by (1.4). The following statements are equivalent:*

(1) *there exists C such that μ satisfies $\mathbf{T}_\alpha(C)$;*
(2) *there exist λ, $A > 0$ such that μ satisfies $((\tau) - \mathbf{LSI}_\alpha(\lambda, A))$.*

Moreover, the constants are related in the following way:

$$(1) \quad \Rightarrow \quad (2) \qquad \text{for any } \lambda \in (0, 1/C) \text{ and } A = \frac{1}{1 - \lambda C};$$

$$(2) \quad \Rightarrow \quad (1) \qquad \text{with } C = \frac{1}{\lambda} \kappa_{p_\alpha} \max(A; 1)^{p_\alpha - 1},$$

where $\kappa_{p_\alpha} = \frac{p_\alpha^{p_\alpha(p_\alpha-1)}}{(p_\alpha-1)^{(p_\alpha-1)^2}}$.

Such a characterization appeared for the first time in [15], in a Euclidean setting and with α between linear and quadratic. Here our result is valid, not only for a wider family of Young functions α, but also on very general metric spaces.

Due to its functional form, it is easy to prove a perturbation lemma for the inequality $(\tau) - \mathbf{LSI}_\alpha$. This leads to the following general Holley–Stroock perturbation result for transport-entropy inequalities whose proof is given in Section 5.

THEOREM 1.9. *Let μ be a probability measure on X and α a Young function satisfying the Δ_2-condition, and let $p_\alpha > 1$ be defined by (1.4). Assume that μ satisfies $\mathbf{T}_\alpha(C)$ for some constant $C > 0$. Then, for any bounded function $\varphi : X \to \mathbb{R}$, the measure $d\tilde{\mu} = \frac{1}{Z} e^{\varphi}\, d\mu$ (where Z is the normalization constant) satisfies $\mathbf{T}_\alpha(\widetilde{C})$, with*

$$\widetilde{C} = \widetilde{\kappa}_{p_\alpha} C e^{(p_\alpha - 1)\mathrm{Osc}(\varphi)},$$

where $\mathrm{Osc}(\varphi) := \sup \varphi - \inf \varphi$, *and* $\widetilde{\kappa}_{p_\alpha} = \frac{p_\alpha^{p_\alpha^2}}{(p_\alpha - 1)^{p_\alpha (p_\alpha - 1)}}$.

This theorem fully extends the previous perturbation result [15], Corollary 1.8, obtained in a Euclidean setting and for a Young function α such that α' is concave. Namely, for such an α, the function $\alpha(x)/x^2$ is nonincreasing [15], Lemma 5.6, and so $p_\alpha \le 2$.

The paper is divided into five sections and one Appendix. Section 2 is dedicated to some preliminaries. In particular we will give a characterization of transport-entropy inequalities (close from Bobkov and Götze one) that might be of independent interest, and that is one of the main ingredients in our proofs. For the sake of completeness, we also recall how the transport-entropy inequality $\mathbf{T}_\alpha$ implies the (τ)-log-Sololev inequality (1) $\Rightarrow$ (2) of Theorem 1.8; this argument had been first used in [27] and then in [15]. In Section 4, we prove the other direction: the (τ)-log-Sololev inequality implies the transport-entropy inequality $\mathbf{T}_\alpha$. In Section 3, we give the proof of the generalized Otto–Villani result, Theorem 1.7. The proof of the Holley–Stroock perturbation result is given in Section 5. Finally, most of the technical results needed on Young functions are proved in the Appendix.

2. Preliminaries. In this section, we first recall the proof of the first half of Theorem 1.8, namely $\mathbf{T}_\alpha \Rightarrow (\tau) - \mathbf{LSI}_\alpha$. In a second part, we give a useful "dimensional" refinement of the characterization of transport-entropy inequalities by Bobkov and Götze [4]. This characterization provides sufficient conditions for the transport-entropy inequality to hold. These are the same conditions as those obtained in the proofs of $\mathbf{LSI}_\alpha^+ \Rightarrow \mathbf{T}_\alpha$ and $(\tau) - \mathbf{LSI}_\alpha \Rightarrow \mathbf{T}_\alpha$.

2.1. *From transport entropy to (τ)-log-Sobolev inequality.* In [15], Theorem 2.1, we proved the following result which is the first half [(1) $\Rightarrow$ (2)] of Theorem 1.8. For the sake of completeness, its short proof is recalled below.

THEOREM 2.1 ([15]). *Let μ be a probability measure on X and α a Young function. If μ satisfies $\mathbf{T}_\alpha(C)$ for some constant $C > 0$, then, for all $\lambda \in (0, 1/C)$, μ satisfies $(\tau) - \mathbf{LSI}_\alpha(\lambda, \frac{1}{1-\lambda C})$.*

PROOF. Take $f : X \to \mathbb{R}$ a locally Lipschitz function such that $\int e^f \, d\mu = 1$, and consider the probability ν_f defined by $\nu_f = e^f \mu$. Jensen's inequality implies that $\int f \, d\mu \leq 0$. So, if π is an optimal coupling between $\nu_f(dx)$ and $\mu(dy)$, then it holds

$$H(\nu_f|\mu) = \int f d\nu_f \leq \int f d\nu_f - \int f \, d\mu = \int f(x) - f(y)\pi(dx\, dy).$$

By definition of $Q_\alpha^\lambda f$,

$$f(x) - f(y) \leq f(x) - Q_\alpha^\lambda f(x) + \lambda\alpha(d(x,y)).$$

Since π is optimal, it holds

$$H(\nu_f|\mu) \leq \int (f - Q_\alpha^\lambda f) \, d\nu_f + \lambda \mathcal{T}_\alpha(\nu_f, \mu).$$

Plugging the inequality $\mathcal{T}_\alpha(\nu_f, \mu) \leq CH(\nu_f|\mu)$ into the inequality above with $\lambda < 1/C$ immediately gives $(\tau) - \mathbf{LSI}_\alpha(\lambda, \frac{1}{1-\lambda C})$. □

2.2. *Sufficient conditions for transport-entropy inequality.* In this section, we show that bounds on the exponential moment of the tensorized inf-convolution or sup-convolution operator allow us to recover the transport-entropy inequality; see Proposition 2.3 and Corollary 2.5 below. These results are a key argument to recover the transport-entropy inequality, either from a modified log-Sobolev or from a (τ)-log-Sobolev inequality.

It is known, since the work by Bobkov and Götze [4] (see also [13, 31]), that transport-entropy inequalities have the following dual formulation.

PROPOSITION 2.2 ([4]). *Let μ be a probability measure on a complete and separable metric space (X, d). Then the following are equivalent:*

(i) *the probability measure μ satisfies $\mathbf{T}_\alpha(1/c)$;*
(ii) *for any bounded continuous function $f : X \to \mathbb{R}$, it holds*

$$\int e^{cQ_\alpha f} \, d\mu \leq e^{c\mu(f)}.$$

In the next proposition we show, using the law of large numbers, that the bound in point (ii) can be relaxed as soon as it holds in any dimension.

PROPOSITION 2.3. *Let μ be a probability measure on a complete and separable metric space (X, d). Then the following are equivalent:*

(i) *the probability measure μ satisfies $\mathbf{T}_\alpha(1/c)$;*
(ii) *there exist three constants a, b, $c > 0$ such that for any $n \in \mathbb{N}^*$, for any bounded continuous function $f : X^n \to \mathbb{R}^+$, it holds*

$$\int e^{cQ_{\alpha,n} f} \, d\mu^n \leq a e^{b\mu^n(f)},$$

where

$$Q_{\alpha,n}f(x) = \inf_{y\in X^n}\left\{ f(y) + \sum_{i=1}^{n} \alpha(d(x_i,y_i))\right\} \qquad \forall x = (x_1,\dots,x_n)\in X^n. \tag{2.1}$$

REMARK 2.4. Note that the constants a and b do not play any role. On the other hand, notice that f is only assumed to be nonnegative.

PROOF. Observe that the transport-entropy inequality $\mathbf{T}_\alpha(1/c)$ naturally tensorises; see, e.g., [13]. Applying Bobkov and Götze result above, we see that (i) implies (ii) with $a=1$ and $b=c$.

Now let us prove that (ii) implies (i). For that purpose, fix a bounded continuous function $f: X\to\mathbb{R}$ with mean 0 under μ and, following [14] (see also [11]), define g on X^n as $g(x)=\sum_{i=1}^n f(x_i)$, $x=(x_1,\dots,x_n)\in X^n$. Then,

$$\left(\int e^{cQ_\alpha f}\,d\mu\right)^n = \int e^{cQ_{\alpha,n}g}\,d\mu^n \le \int e^{cQ_{\alpha,n}g_+}\,d\mu^n \le a e^{b\mu^n(g_+)},$$

where, as usual, $g_+ = \max(g,0)$. It follows that

$$\int e^{cQ_\alpha f}\,d\mu \le a^{1/n} e^{b\mu^n(g_+)/n}.$$

Now, according to the strong law of large numbers, $\frac{1}{n}\sum_{i=1}^n f(X_i)\to 0$ in $\mathbb{L}^1$, where the X_i's are i.i.d. random variables with common law μ. Hence,

$$\mu^n\left(\frac{g_+}{n}\right) \le \mathbb{E}\left(\left|\frac{1}{n}\sum_{i=1}^n f(X_i)\right|\right)\to 0$$

when n tends to infinity. We conclude that

$$\int e^{cQ_\alpha f}\,d\mu \le 1 = e^{c\mu(f)}. \tag{2.2}$$

Since the latter is invariant by changing f into $f+e$ for any constant e, we can remove the assumption $\mu(f)=0$. This ends the proof. □

The next corollary will be used in the proofs of Theorems 1.7 and 1.8. It gives a sufficient condition for the transport-entropy inequality $\mathbf{T}_\alpha$ to hold.

COROLLARY 2.5. *Let μ be a probability measure on a complete and separable metric space (X,d). Define, for all $f: X^n\to\mathbb{R}$,*

$$P_{\alpha,n}f(x) = \sup_{y\in X^n}\left\{ f(y) - \sum_{i=1}^{n} \alpha(d(x_i,y_i))\right\} \qquad \forall x = (x_1,\dots,x_n)\in X^n. \tag{2.3}$$

Assume that there exist some constants τ, a, $b>0$ *and* $c\in[0,1)$ *such that, for all integer* $n\in\mathbb{N}^*$ *and all bounded continuous functions* $f:X^n\to\mathbb{R}^+$, *it holds*

$$\int e^{\tau P_{\alpha,n}f}\,d\mu^n \le a e^{b\mu^n(P_{\alpha,n}f)} e^{\tau c\|f\|_\infty}.$$

Then μ *satisfies* $\mathbf{T}_\alpha(\frac{1}{\tau(1-c)})$.

Proof. Let $n\in\mathbb{N}^*$ and take a bounded continuous function $g:X^n\to\mathbb{R}^+$. In order to apply Proposition 2.3, we need to remove the spurious term $\|f\|_\infty$. Observe on the one hand that for any $\beta\in(0,\tau(1-c))$, one has

$$\begin{aligned}\int e^{\beta Q_{\alpha,n}g}\,d\mu^n &= 1+\beta\int_0^{+\infty} e^{\beta r}\mu^n(Q_{\alpha,n}g\ge r)\,dr\\ &= 1+\beta\int_0^{+\infty} e^{\beta r}\mu^n(\min(Q_{\alpha,n}g,r)\ge r)\,dr.\end{aligned}$$

On the other hand, set $f=\min(Q_{\alpha,n}g,r)$. It is bounded, nonnegative and satisfies $\|f\|_\infty\le r$. Moreover, we have $P_{\alpha,n}(Q_{\alpha,n}g)\le g$. Indeed,

$$P_{\alpha,n}(Q_{\alpha,n}g)(x)=\sup_{y\in X^n}\inf_{z\in X^n}\left\{f(z)+\sum_{i=1}^n\alpha(d(y_i,z_i))-\sum_{i=1}^n\alpha(d(x_i,y_i))\right\},$$

and the inequality follows by taking $z=x$. Hence $\mu^n(P_{\alpha,n}f)=\mu^n(P_{\alpha,n}(Q_{\alpha,n}g))\le\mu^n(g)$. Therefore, since $P_{\alpha,n}f\ge f$, by Chebyshev's inequality and the assumption, we have

$$\begin{aligned}\mu^n(\min(Q_{\alpha,n}g,r)\ge r)\le\mu^n(P_{\alpha,n}f\ge r)&\le e^{-\tau r}\int e^{\tau P_{\alpha,n}f}\,d\mu^n\\ &\le a e^{b\mu^n(P_{\alpha,n}f)}e^{-\tau(1-c)r}\\ &\le a e^{b\mu^n(g)}e^{-\tau(1-c)r}.\end{aligned}$$

Consequently, we get

$$\begin{aligned}\int e^{\beta Q_{\alpha,n}g}\,d\mu^n &\le 1+\beta a e^{b\mu^n(g)}\int_0^{+\infty} e^{-(\tau(1-c)-\beta)r}\,dr\\ &= 1+\frac{\beta a}{\tau(1-c)-\beta}e^{b\mu^n(g)}\\ &\le \frac{\tau(1-c)+\beta(a-1)}{\tau(1-c)-\beta}e^{b\mu^n(g)}.\end{aligned}$$

Finally, Proposition 2.3 provides that μ satisfies $\mathbf{T}_\alpha(1/\beta)$. Optimizing over β leads to the expected result. □

3. From modified log-Sobolev inequality to transport-entropy inequality. In this section we prove Theorem 1.7. We have to distinguish between the modified log-Sobolev inequalities plus and minus. As in [11], the proofs of Theorems 1.7 and 1.8 use as a main ingredient the stability of log-Sobolev-type inequalities under tensor products.

Let us recall this tensorisation property. The entropy functional enjoys the following well-known sub-additivity property (see, e.g., [1], Chapter 1): if $h : X^n \to \mathbb{R}^+$,

$$\operatorname{Ent}_{\mu^n}(h) \le \sum_{i=1}^{n} \int \operatorname{Ent}_{\mu}(h_{i,x})\, d\mu^n(x), \tag{3.1}$$

where, for all $x \in X^n$, the application $h_{i,x}$ is the ith partial application defined by

$$h_{i,x}(u) = h(x_1, \dots, x_{i-1}, u, x_{i+1}, \dots, x_n) \qquad \forall u \in X.$$

Let us say that $h : X^n \to \mathbb{R}^+$ is *separately locally Lipschitz*, if all the partial applications $h_{i,x}$ $1 \le i \le n$, $x \in X^n$ are locally Lipschitz on X. Now, suppose that a probability μ on X verifies ($\mathbf{LSI}^+_\alpha(A)$) for some $A > 0$. Then, using (3.1), we easily conclude that μ^n enjoys the following inequality:

$$\operatorname{Ent}_{\mu^n}(e^f) \le A \int \sum_{i=1}^{n} \alpha^*(|\nabla_i^+ f|) e^f \, d\mu^n \tag{3.2}$$

for all functions $f : X^n \to \mathbb{R}$ separately locally Lipschitz, where $|\nabla_i^+ f|(x)$ is defined by

$$|\nabla_i^+ f|(x) = |\nabla^+ f_{i,x}|(x_i) = \limsup_{y \to x_i} \frac{[f(x_1, \dots, x_{i-1}, y, x_{i+1}, \dots, x_n) - f(x)]_+}{d(y, x_i)}.$$

The same property holds for $\mathbf{LSI}^-_\alpha$.

3.1. *Modified log-Sobolev inequality plus.* The first part of Theorem 1.7, that we restate below, says that the modified log-Sobolev inequality $\mathbf{LSI}^+_\alpha$ implies the transport-entropy inequality $\mathbf{T}_\alpha$. In fact we shall prove the following, slightly stronger, result. To any Young function α, we associate a function ξ_α defined by

$$\xi_\alpha(x) := \sup_{u > 0} \frac{\alpha^*(x \alpha'_+(u))}{x \alpha(u)}, \qquad x > 0, \tag{3.3}$$

where α'_+ is the right derivative of α. Note that ξ_α is nondecreasing and may take infinite values.

THEOREM 3.1. *Let μ be a probability measure on X and α a Young function satisfying the Δ_2-condition. If μ satisfies ($\mathbf{LSI}^+_\alpha(A)$) for some constant $A > 0$, then μ satisfies $\mathbf{T}_\alpha(1/t_A)$ with $t_A = \sup\{t \in \mathbb{R}^+; \xi_\alpha(t) < 1/A\}$.*

The following lemma gives an estimation of ξ_α.

LEMMA 3.2. *Let α be a Young function satisfying the Δ_2-condition, and let $1 \le r_\alpha \le p_\alpha$, $p_\alpha > 1$ be the numbers defined by (1.4). Then, it holds*

$$\xi_\alpha(x) \le (p_\alpha - 1)\max(x^{1/(p_\alpha-1)}; x^{1/(r_\alpha-1)}) \qquad \forall x > 0, \tag{3.4}$$

with the convention $x^\infty = 0$ if $x \le 1$ and ∞ otherwise.

The proof of this result is in the Appendix.

Using Lemma 3.2, we easily derive point (i) of Theorem 1.7, with the explicit constant $C^+ = \max(((p_\alpha - 1)A)^{r_\alpha - 1}; ((p_\alpha - 1)A)^{p_\alpha - 1})$.

Before turning to the proof of Theorem 3.1, let us say that estimation (3.4) is satisfactory, at least for the small values of x (corresponding to the large values of A), as we show with the following exact calculation of ξ_α, when α is the function α_{p_1,p_2} defined by (1.1).

LEMMA 3.3. *Let $p_1 \ge 2$ and $p_2 > 1$, and let $\alpha = \alpha_{p_1,p_2}$; then $p_\alpha = \max(p_1, p_2)$, and it holds*

$$\xi_\alpha(x) = (p_\alpha - 1)x^{1/(p_\alpha-1)} \qquad \forall x \le 1.$$

Moreover, for $x \ge 1$, it holds

$$\xi_\alpha(x) = \begin{cases} p_1\left(\dfrac{1}{q_2}x^{1/(p_2-1)} + \left(\dfrac{1}{q_1} - \dfrac{1}{q_2}\right)\dfrac{1}{x}\right), & \text{if } p_1 \ge p_2, \\ \max((p_1 - 1)x^{1/(p_1-1)}; (p_2 - 1)x^{1/(p_2-1)}), & \text{if } p_1 \le p_2, \end{cases}$$

where $q_1 = p_1/(p_1 - 1)$ and $q_2 = p_2/(p_2 - 1)$.

The proof of this lemma is in the Appendix, too.

PROOF OF THEOREM 3.1. Our aim is to use Herbst's argument (see, e.g., [1, 17, 19]) together with Proposition 2.3. Let $n \in \mathbb{N}^*$; according to Lemma 3.4 below, for any bounded function $f : X^n \to \mathbb{R}$, the function $Q_{\alpha,n}f$ is separately locally Lipschitz [recall that the inf-convolution operator $Q_{\alpha,n}$ is defined by (2.1)]. Fix a nonnegative bounded continuous function $f : X^n \to \mathbb{R}^+$. Applying (3.2) to $tQ_{\alpha,n}f$, $t > 0$, and using Lemma 3.5 below together with the fact that $f \ge 0$, one gets

$$\begin{aligned} \mathrm{Ent}_{\mu^n}(e^{tQ_{\alpha,n}f}) &\le A\int \sum_{i=1}^n \alpha^*(t|\nabla_i^+ Q_{\alpha,n}f|)e^{tQ_{\alpha,n}f}\,d\mu^n \\ &\le At\xi_\alpha(t)\int Q_{\alpha,n}f e^{tQ_{\alpha,n}f}\,d\mu^n. \end{aligned}$$

Now, we proceed with the Herbst argument. Set $H(t) = \int e^{tQ_{\alpha,n}f}\, d\mu^n$, $t > 0$. Since $\mathrm{Ent}_{\mu^n}(e^{tQ_{\alpha,n}f}) = tH'(t) - H(t)\log H(t)$, the latter can be rewritten as

$$(t - At\xi_\alpha(t))H'(t) \le H(t)\log H(t), \qquad t > 0.$$

Set $W(t) = \frac{1}{t}\log(H(t))$, $t > 0$, so that the previous differential inequality reduces to

$$W'(t)t(1 - A\xi_\alpha(t)) \le A\xi_\alpha(t)W(t).$$

Since $\lim_{t\to 0} W(t) = \mu^n(Q_{\alpha,n}f)$, we get

$$H(t) \le \exp(tC(t)\mu^n(Q_{\alpha,n}f)) \qquad \forall t \in (0, t_A),$$

where we set $C(t) = \exp\int_0^t \frac{A\xi_\alpha(u)}{u(1-A\xi_\alpha(u))}\, du$; thanks to Lemma 3.2 above, we are guaranteed that $t_A > 0$ and that $C(t) < \infty$ on $(0, t_\alpha)$. Since $Q_{\alpha,n}f \le f$, we finally get

$$\int e^{tQ_{\alpha,n}f}\, d\mu^n \le e^{tC(t)\mu^n(f)} \qquad \forall t \in (0, t_A),$$

which leads to the expected result, thanks to Proposition 2.3 [and after optimization over $t \in (0, t_A)$]. $\square$

LEMMA 3.4. *Let α be a Young function. For any integer $n \in \mathbb{N}^*$, any bounded function $f : X^n \to \mathbb{R}$, the function $Q_{\alpha,n}f$ is separately locally Lipschitz on X^n.*

PROOF. Let $h = Q_{\alpha,n}f$; then, for all $x \in X^n$ and $1 \le i \le n$, it holds

$$\begin{aligned} h_{i,x}(u) &= \inf_{y_i \in X}\left\{ \inf_{y_1,\dots,y_{i-1},y_{i+1},\dots,y_n}\left\{ f(y) + \sum_{j\neq i} \alpha(d(x_j, y_j))\right\} + \alpha(d(u, y_i))\right\} \\ &= Q_\alpha g(u), \end{aligned}$$

where $g : X \to \mathbb{R}$ is defined by the second infimum. Let us show that $u \mapsto Q_\alpha g(u)$ is locally Lipschitz on X. Observe that g is bounded and define $r_o = \alpha^{-1}(2\|g\|_\infty)$. For all $u \in X$, and all $y \in X$ such that $d(y, u) > r_o$, we have

$$g(y) + \alpha(d(u, y)) > -\|g\|_\infty + \alpha(r_o) = \|g\|_\infty.$$

Since $Q_\alpha g \le \|g\|_\infty$, we conclude that $Q_\alpha g(u) = \inf_{d(y,u)\le r_o}\{g(y) + \alpha(d(u, y))\}$. Let $u_o \in X$, and let B_o be the closed ball of center u_o and radius $2r_o$. If $u \in B_o$, then $Q_\alpha g(u) = \inf_{y\in B_o}\{g(y) + \alpha(d(u, y))\}$. Now, if $y \in B_o$, we see that for all $u, v \in B_o$,

$$\begin{aligned} &|\alpha(d(u, y)) - \alpha(d(v, y))| \\ &\quad \le |d(v, y) - d(u, y)| \max_{t\in[0,1]} \alpha'_+(td(u, y) + (1 - t)d(v, y)) \\ &\quad \le L_o d(u, v), \end{aligned}$$

with $L_o = \alpha'_+(4r_o)$. The map $B_o \to \mathbb{R} : u \mapsto Q_\alpha g(u)$ is an infimum of L_o-Lipschitz functions on B_o, so it is L_o-Lipschitz on B_o. This ends the proof. □

LEMMA 3.5. *Let α be a Young function. For any integer n, any $t \geq 0$ and any bounded continuous function $f : X^n \to \mathbb{R}$,*

$$\sum_{i=1}^{n} \alpha^*(t|\nabla_i^+ Q_{\alpha,n} f|) \leq t\xi_\alpha(t)(Q_{\alpha,n} f(x) - f(y^x)),$$

where $y^x \in X^n$ is any point such that $Q_{\alpha,n} f(x) = f(y^x) + \sum_{j=1}^n \alpha(d(x_j, y_j^x))$.

PROOF. Fix n, $t \geq 0$ and a bounded function $f : X^n \to \mathbb{R}^+$. For $x = (x_1, \ldots, x_n) \in X^n$, $i \in \{1, \ldots, n\}$ and $z \in X$, we shall use the following notation:

$$\bar{x}^i z = (x_1, \ldots, x_{i-1}, z, x_{i+1}, \ldots, x_n).$$

Let $x \in X^n$; since f is bounded continuous and closed balls in X are assumed to be compact, it is not difficult to show that there exists $y^x \in X^n$ such that

$$Q_{\alpha,n} f(x) = f(y^x) + \sum_{j=1}^{n} \alpha(d(x_j, y_j^x)).$$

For all $z \in X$ and all $1 \leq i \leq n$, we have also $Q_{\alpha,n} f(\bar{x}^i z) \leq f(y^x) + \sum_{j \neq i} \alpha(d(x_j, y_j^x)) + \alpha(d(z, y_i^x))$. Since the maps $u \mapsto [u]_+$ and α are non-decreasing, it holds

$$\begin{aligned}[Q_{\alpha,n} f(\bar{x}^i z) - Q_{\alpha,n} f(x)]_+ &\leq [\alpha(d(z, y_i^x)) - \alpha(d(x_i, y_i^x))]_+ \\ &\leq [\alpha(d(z, x_i) + d(x_i, y_i^x)) - \alpha(d(x_i, y_i^x))]_+.\end{aligned}$$

Therefore,

$$\begin{aligned}|\nabla_i^+ Q_{\alpha,n} f|(x) &\leq \limsup_{z \to x_i} \frac{[\alpha(d(z, x_i) + d(x_i, y_i^x)) - \alpha(d(x_i, y_i^x))]_+}{d(z, x_i)} \\ &= \alpha'_+(d(x_i, y_i^x)).\end{aligned}$$

Hence, by the very definition of ξ_α,

$$\begin{aligned}\sum_{i=1}^{n} \alpha^*(t|\nabla_i^+ Q_{\alpha,n} f|) &\leq \sum_{i=1}^{n} \alpha^*(t\alpha'_+(d(x_i, y_i^x))) \\ &\leq t\xi_\alpha(t) \sum_{i=1}^{n} \alpha(d(x_i, y_i^x)) \\ &= t\xi_\alpha(t)(Q_{\alpha,n} f(x) - f(y^x)).\end{aligned}$$

□

3.2. *Modified log-Sobolev inequality minus.* In this section we prove the second part of Theorem 1.7, that we restate (in a slightly stronger form) below, namely that the modified log-Sobolev inequality minus $\mathbf{LSI}_\alpha^-$ implies the transport-entropy inequality $\mathbf{T}_\alpha$. Let us define [recall the defintion of ξ_α given in (3.3)]

$$t_\alpha = \sup\{t \in \mathbb{R}^+, \xi_\alpha(t) < +\infty\}.$$

Note that, by Lemma 3.2, if α satisfies the Δ_2-condition, then $t_\alpha \geq 1$.

THEOREM 3.6. *Let μ be a probability measure on X and α a Young function satisfying the Δ_2-condition. If μ satisfies ($\mathbf{LSI}_\alpha^-(A)$) for some constant $A > 0$, then μ satisfies $\mathbf{T}_\alpha(B^-)$ with $B^- = \lim_{t\to t_\alpha} \frac{1}{t} \exp\{\int_0^t \frac{A\xi_\alpha(u)}{u(1+A\xi_\alpha(u))}\, du\}$.*

For more comprehension and to complete the proof of part (ii) of Theorem 1.7, let us prove that $B^- \leq C^-$. If $r_\alpha > 1$, then by Lemma 3.2, $t_\alpha = +\infty$. Moreover, using that $\frac{1}{t} = \exp\{-\int_1^t \frac{1}{u}\, du\}$, $t \geq 1$, one has

$$\begin{aligned}
\log B^- &= \int_0^1 \frac{A\xi_\alpha(u)}{u(1+A\xi_\alpha(u))}\, du - \int_1^{+\infty} \frac{1}{u(1+A\xi_\alpha(u))}\, du \\
&\leq \int_0^1 \frac{A(p_\alpha - 1)u^{1/(p_\alpha - 1)}}{u(1+A(p_\alpha-1)u^{1/(p_\alpha-1)})}\, du \\
&\quad - \int_1^{+\infty} \frac{1}{u(1+A(p_\alpha-1)u^{1/(r_\alpha-1)})}\, du \\
&= \log C^-,
\end{aligned}$$

with

$$C^- = (1 + A(p_\alpha - 1))^{p_\alpha - r_\alpha}(A(p_\alpha - 1))^{r_\alpha - 1}.$$

When $r_\alpha = 1$, since $t_\alpha \geq 1$ and using the fact that the function

$$t \to \frac{1}{t} \exp\left\{\int_0^t \frac{A\xi_\alpha(u)}{u(1+A\xi_\alpha(u))}\, du\right\}$$

is nonincreasing, we get

$$B^- \leq \exp\left\{\int_0^1 \frac{A\xi_\alpha(u)}{u(1+A\xi_\alpha(u))}\, du\right\} \leq (1 + A(p_\alpha - 1))^{p_\alpha - 1} = C^-.$$

PROOF OF THEOREM 3.6. The proof of Theorem 3.6 follows essentially the lines of the proof of Theorem 3.1. Let $n \in \mathbb{N}^*$; thanks to the tensorisation property of ($\mathbf{LSI}_\alpha^-(A)$), it holds

$$\mathrm{Ent}_{\mu^n}(e^g) \leq A \int \sum_{i=1}^n \alpha^*(|\nabla_i^- g|) e^g\, d\mu^n \tag{3.5}$$

for any $g: X^n \to \mathbb{R}$ separately locally Lipschitz and bounded. Take a nonnegative bounded continuous function $f: X^n \to \mathbb{R}^+$. Recall that $P_{\alpha,n} f(x) = \sup_{y \in X^n}\{f(y) - \sum_{i=1}^n \alpha(d(x_i, y_i))\}$. Since $P_{\alpha,n} f = -Q_{\alpha,n}(-f)$, it follows from Lemma 3.4 that $P_{\alpha,n} f$ is separately locally Lipschitz. Applying (3.5) to $g = tP_{\alpha,n} f$, $t > 0$, one gets

$$\mathrm{Ent}_{\mu^n}(e^{tP_{\alpha,n} f}) \le A \int \sum_{i=1}^n \alpha^*(t|\nabla_i^- P_{\alpha,n} f|) e^{tP_{\alpha,n} f}\, d\mu^n.$$

Observe that $P_{\alpha,n} f = -Q_{\alpha,n}(-f)$ and that $|\nabla^-(-h)| = |\nabla^+ h|$, for all $h: X \to \mathbb{R}$. So applying Lemma 3.5, we see that for all $x \in X^n$, there is some $y^x \in X^n$ such that

$$\begin{aligned}\sum_{i=1}^n \alpha^*(t|\nabla_i^- P_{\alpha,n} f|)(x) &= \sum_{i=1}^n \alpha^*(t|\nabla_i^+ Q_{\alpha,n}(-f)|)(x) \\ &\le t\xi_\alpha(t)(Q_{\alpha,n}(-f)(x) + f(y^x)) \\ &\le t\xi_\alpha(t)(\|f\|_\infty - P_{\alpha,n} f(x)).\end{aligned}$$

So we get the following inequality:

$$\mathrm{Ent}_{\mu^n}(e^{tP_{\alpha,n} f}) \le At\xi_\alpha(t) \int (\|f\|_\infty - P_{\alpha,n} f) e^{tP_{\alpha,n} f}\, d\mu^n.$$

As in the proof of Theorem 3.1, we proceed with the Herbst argument. Set $H(t) = \int e^{tP_{\alpha,n} f}\, d\mu^n$, $t \in (0, t_\alpha)$. Since $\mathrm{Ent}_{\mu^n}(e^{tP_{\alpha,n} f}) = tH'(t) - H(t)\log H(t)$, the latter can be rewritten as

$$(t + At\xi_\alpha(t))H'(t) \le H(t)\log H(t) + At\xi_\alpha(t)\|f\|_\infty H(t) \qquad \forall t \in (0, t_\alpha).$$

Set $W(t) = \frac{1}{t}\log H(t)$, $t \in (0, t_\alpha)$, so that the previous differential inequality reduces to

$$W'(t)t(1 + A\xi_\alpha(t)) \le -A\xi_\alpha(t)W(t) + A\xi_\alpha(t)\|f\|_\infty.$$

Set $c(t) = \exp\{-\int_0^t \frac{A\xi_\alpha(u)}{u(1+A\xi_\alpha(u))}\, du\}$ [which belongs to $(0,1)$ thanks to Lemma 3.2]. Since $\lim_{t\to 0} W(t) = \mu^n(P_{\alpha,n} f)$, solving the latter differential inequality, we easily get that for all $t \in (0, t_\alpha)$,

$$H(t) \le e^{tc(t)\mu^n(P_{\alpha,n} f)} e^{t\|f\|_\infty(1-c(t))}.$$

Applying Corollary 2.5 yields that $\mathbf{T}_\alpha(1/(tc(t)))$ holds for all $t \in (0, t_\alpha)$. Observing that the function $t \to tc(t)$ is nondecreasing on $(0, t_\alpha)$, the proof is completed by optimizing in t. $\square$

4. From (τ)-log-Sobolev inequality to transport-entropy inequality. In this section, we prove the second part [(2) $\Rightarrow$ (1)] of Theorem 1.8. Observe that $\mathbf{T}_\alpha(C/\lambda)$ is equivalent to $\mathbf{T}_{\lambda\alpha}(C)$. Hence, changing α into $\lambda\alpha$, we can restate the first part of Theorem 1.8 as follows.

THEOREM 4.1. *Let μ be a probability measure on X and α a Young function satisfying the Δ_2-condition. Let $p_\alpha > 1$ be defined by (1.4). If μ satisfies $(\tau) - \mathbf{LSI}_\alpha(1, A)$ for some $A > 0$, then μ satisfies $\mathbf{T}_\alpha(C)$ with*

$$C = \kappa_{p_\alpha} \max(A, 1)^{p_\alpha - 1},$$

where $\kappa_{p_\alpha} = \frac{p_\alpha^{p_\alpha(p_\alpha-1)}}{(p_\alpha-1)^{(p_\alpha-1)^2}}$.

Two proofs are given below. The first one exactly follows the lines of the proof of $\mathbf{LSI}_\alpha^- \Rightarrow \mathbf{T}_\alpha$, whereas the second one uses the equivalence between transport-entropy inequalities and dimension-free concentration established in [11] together with a change of metric argument.

In each proof, the first step is to tensorise the (τ)-log-Sobolev inequality. Let $n \in \mathbb{N}^*$; using the sub-additivity property (3.1) of the entropy functional, we see that $(\tau) - \mathbf{LSI}_\alpha(1, A)$ implies that

$$\operatorname{Ent}_{\mu^n}(e^h) \le A \int \sum_{i=1}^n (h - Q_\alpha^{(i)} h) e^h \, d\mu^n \qquad \forall h : X^n \to \mathbb{R}, \tag{4.1}$$

where $Q_\alpha^{(i)}$ is the inf-convolution operator with respect to the ith coordinate, namely

$$Q_\alpha^{(i)} h(x) = Q_\alpha(h_{i,x})(x_i) = \inf_{y \in X} \{h(\bar{x}^i y) + \alpha(d(x_i, y))\}$$

(using the notation introduced in Section 3).

As in the proof of Theorem 3.6, applying (4.1) to $h = tP_{\alpha,n} g$, $t \ge 0$ where g belongs to some class of functions, we get

$$\operatorname{Ent}_{\mu^n}(e^{tP_{\alpha,n} g}) \le A \int \sum_{i=1}^n (tP_{\alpha,n} g - Q_\alpha^{(i)}(tP_{\alpha,n} g)) e^{tP_{\alpha,n} g} \, d\mu^n. \tag{4.2}$$

As a main difference, the class of functions g differs in each proof. In the first one, g is any nonnegative bounded separately locally Lipschitz function, whereas in the second proof, g is globally Lipschitz in some sense.

For both proofs, the next step is to bound efficiently the right-hand side of (4.2), in order to use some Herbst argument. This bound will be given by the following lemma.

LEMMA 4.2. *Let α be a Young function satisfying the Δ_2-condition, and let $p_\alpha > 1$ be defined by (1.4). For any bounded continuous function $g : X^n \to \mathbb{R}$, for any $x \in X^n$ and $t \in [0, 1)$,*

$$\sum_{i=1}^n (tP_{\alpha,n} g(x) - Q_\alpha^{(i)}(tP_{\alpha,n} g)(x)) \le t\varepsilon(t) \left(\sum_{i=1}^n \alpha(d(x_i, y_i^x)) \right),$$

where $y^x \in X^n$ *is any point such that* $P_{\alpha,n}g(x) = g(y^x) - \sum_{i=1}^n \alpha(d(x_i, y_i^x))$, *and where*

$$\varepsilon(t) = \frac{1}{(1-t^{1/(p_\alpha-1)})^{p_\alpha-1}} - 1 \qquad \forall t \in [0,1).$$

We postpone the proof of Lemma 4.2 to the end of the section.

4.1. *A first proof.* The first proof of Theorem 4.1 mimics the one of the implication $\mathbf{LSI}_\alpha^- \Rightarrow \mathbf{T}_\alpha$.

Proof of Theorem 4.1. Using (4.2), Lemma 4.2 ensures that for every nonnegative locally Lipschitz bounded function g, for every $t \in [0,1)$,

$$(t + At\varepsilon(t))H'(t) \le H(t)\log H(t) + At\varepsilon(t)\|g\|_\infty H(t),$$

where $H(t) = \int e^{tP_{\alpha,n}g}\, d\mu^n$.

Solving this differential inequality, exactly as in the proof of Theorem 3.6 (we omit details), leads to

$$\int e^{P_{\alpha,n}g}\, d\mu^n \le e^{c\mu^n(P_{\alpha,n}g)} e^{\|g\|_\infty(1-c)},$$

with $c = 1/C$,

$$C = \exp\int_0^1 \frac{A\varepsilon(t)}{t(1+A\varepsilon(t))}\, dt.$$

The inequality $\mathbf{T}_\alpha(C)$ then follows from Corollary 2.5.

Now, let us estimate the constant C. By convexity, one has for every $v \in [0,1]$,

$$(1-v) - (1-v)^{p_\alpha} \le (p_\alpha - 1)v.$$

This inequality easily implies that for all $t \in [0,1)$, $\frac{\varepsilon(t)}{t} \le (p_\alpha - 1)\varepsilon'(t)$. Consequently, we obtain for all $u \in [0,1)$,

$$\begin{aligned}\log C &\le (p_\alpha - 1)\int_0^u \frac{A\varepsilon'(t)}{1+A\varepsilon(t)}\, dt + \int_u^1 \frac{A\varepsilon(t)}{t(1+A\varepsilon(t))}\, dt \\ &\le (p_\alpha - 1)\log(1+A\varepsilon(u)) - \log u.\end{aligned}$$

Optimizing in u, we get

$$\begin{aligned}C &\le \inf_{u\in(0,1)} \frac{(1+A\varepsilon(u))^{p_\alpha-1}}{u} \le \inf_{u\in(0,1)} \frac{(1+\varepsilon(u))^{p_\alpha-1}}{u} \max(A,1)^{p_\alpha-1} \\ &= \kappa_{p_\alpha} \max(A,1)^{p_\alpha-1},\end{aligned}$$

with $\kappa_{p_\alpha} = \frac{p_\alpha^{p_\alpha(p_\alpha-1)}}{(p_\alpha-1)^{(p_\alpha-1)^2}}$. □

4.2. *A second proof.* The idea of this second proof is to prove the theorem in the particular case of the functions $\alpha_p(x) = x^p$ and then to treat the general case by a change of metric argument.

4.2.1. $\mathbf{T}_p$ *inequalities.* Let us introduce some notation and definitions. When $\alpha(x) = \alpha_p(x) = |x|^p$, we will use the notation $\mathbf{T}_p(C)$ and $(\tau) - \mathbf{LSI}_p$ instead of $\mathbf{T}_{\alpha_p}(C)$ and $(\tau) - \mathbf{LSI}_{\alpha_p}$. Let $n \in \mathbb{N}^*$; a function $f : X^n \to \mathbb{R}$ is said to be (L,p)-Lipschitz $L > 0, p > 1$, if

$$|f(x) - f(y)| \le L \left(\sum_{i=1}^n d^p(x_i, y_i) \right)^{1/p} \qquad \forall x, y \in X^n.$$

We recall the following result from [11].

THEOREM 4.3. *The probability μ verifies the transport-entropy inequality $\mathbf{T}_p(C)$, for some $C > 0$ if and only if it enjoys the following dimension free concentration property: for all $n \in \mathbb{N}^*$ and all $f : X^n \to \mathbb{R}$ such that*

$$|f(x) - f(y)| \le L \left(\sum_{i=1}^n d^p(x_i, y_i) \right)^{1/p} \qquad \forall x, y \in X^n$$

for some $L > 0$, it holds

$$\mu^n(f \ge \mu^n(f) + u) \le \exp(-u^p/(L^p C)) \qquad \forall u \ge 0.$$

So to show that a $\mathbf{T}_p$ inequality holds, it is enough to prove the right concentration inequality.

We will use the following result to estimate the right-hand side of (4.2).

LEMMA 4.4. *Let $p > 1$; there exists a constant $\omega_p \ge 1$ such that for all $n \in \mathbb{N}^*$ and all (L,p)-Lipschitz function $f : X^n \to \mathbb{R}$, and all $x \in X^n$, the function*

$$X^n \to \mathbb{R} : y \mapsto f(y) - \sum_{i=1}^n d^p(x_i, y_i)$$

attains its maximum on the closed ball

$$\left\{ y \in X^n; \sum_{i=1}^n d^p(x_i, y_i) \le \left(\frac{L}{\omega_p} \right)^q \right\}, \qquad \text{with } q = \frac{p}{p-1}.$$

When (X,d) is geodesic (see below), then one can take $\omega_p = p$.

Recall that (X,d) is geodesic, if for all $x, y \in X$, there is a path $(z_t)_{t \in [0,1]}$ joining x to y and such that $d(z_s, z_t) = |s - t| d(x,y)$, for all $s, t \in [0,1]$. This notion encompasses the case of Riemannian manifolds.

The proof of the lemma is at the end of the section.

THEOREM 4.5. *Let $p \geq 2$; if μ verifies the $(\tau) - \mathbf{LSI}_p(1, A)$, then it verifies $\mathbf{T}_p(C)$, with $C = (a_p \max(1; A))^{p-1}$, with $a_p = \inf_{t\in(0,1)}\{\frac{1}{t^{q-1}}(1 + \frac{(p/\omega_p)^q}{p-1}\int_0^t \frac{\varepsilon(u)}{u}\,du)\}$.*

REMARK 4.6. Let us compare the constants appearing in Theorems 4.1 and 4.5 for $p = 2$. When $p = 2$, Theorem 4.1 gives $C_1 = 4\max(1; A)$, and Theorem 4.5 gives $C_2 = a_2 \max(1; A)$. A simple calculation shows that when $p = 2$, $a_2 = \inf_{s\in(0,1)}\{\frac{1-(2/\omega_2)^2 \ln(1-s)}{s}\}$. If $\omega_2 = 1$, then $a_2 \simeq 7,5$, and C_1 is smaller than C_2. But if $\omega_2 = 2$ [which is the case, when (X, d) is geodesic], then $a_2 \simeq 3,14$, and C_2 is smaller than C_1.

PROOF OF THEOREM 4.5. Take a (L, p)-Lipschitz function $g : X^n \to \mathbb{R}$. To bound the right-hand side of (4.2), we use Lemmas 4.2 and 4.4.

$$\sum_{i=1}^n (tP_{\alpha_p,n}g(x) - Q_{\alpha_p}^{(i)}(tP_{\alpha_p,n}g)(x)) \leq t\varepsilon(t)(L/\omega_p)^q.$$

So, letting $H(t) = \int e^{tP_{\alpha_p,n}g}\,d\mu^n$, (4.2) provides

$$tH'(t) - H(t)\log H(t) \leq At\varepsilon(t)H(t)(L/\omega_p)^q \qquad \forall t \in [0, 1).$$

Equivalently, the function $K(t) = \frac{1}{t}\log H(t)$ verifies $K'(t) \leq A(L/\omega_p)^q \frac{\varepsilon(t)}{t}$. Since $K(t) \to \mu^n(P_{\alpha_p,n}g)$ when $t \to 0^+$, we conclude that

$$\int e^{tP_{\alpha_p,n}g}\,d\mu^n \leq \exp(t\mu^n(P_{\alpha_p,n}g) + tA(L/\omega_p)^q k(t)) \qquad \forall t \in [0, 1), \tag{4.3}$$

where $k(t) = \int_0^t \frac{\varepsilon(u)}{u}\,du$, $t \in [0, 1)$. Since g is (L, p)-Lipschitz, it holds

$$\begin{aligned} 0 \leq P_{\alpha_p,n}g(x) - g(x) &\leq \sup_{y\in X^n}\left\{L\left(\sum_{i=1}^n d^p(x_i, y_i)\right)^{1/p} - \sum_{i=1}^n d^p(x_i, y_i)\right\} \\ &= \sup_{r\geq 0}\{Lr - r^p\} = (p-1)(L/p)^q. \end{aligned}$$

Plugging the inequalities $g \leq P_{\alpha_p,n}g$ and $\mu^n(P_{\alpha_p,n}g) \leq \mu^n(g) + (p-1)(L/p)^q$ into (4.3), we get

$$\int e^{tg}\,d\mu^n \leq \exp(t\mu^n(g) + t(p-1)(L/p)^q + tA(L/\omega_p)^q k(t)) \qquad \forall t \in [0, 1).$$

Applying this inequality to $g = f/t$ with f a (L, p)-Lipschitz function, we get

$$\int e^{f-\mu^n(f)}\,d\mu^n \leq \exp\left(\frac{(L/p)^q}{t^{q-1}}(p - 1 + A(p/\omega_p)^q k(t))\right) \qquad \forall t \in (0, 1).$$

So, optimizing over $t \in (0,1)$ yields

$$\int e^{f-\mu^n(f)}\, d\mu^n \le \exp((L/p)^q (p-1)\max(1;A) a_p), \tag{4.4}$$

with

$$a_p = \inf_{t\in(0,1)} \left\{ \frac{1}{t^{q-1}} \left(1 + \frac{(p/\omega_p)^q}{p-1} \int_0^t \frac{\varepsilon(u)}{u}\, du \right) \right\}.$$

Using Chebyshev's argument, we derive from (4.4) that

$$\mu^n(f \ge \mu^n(f) + u) \le \exp(-u^p/(L^p C)) \qquad \forall u \ge 0,$$

with $C = (a_p \max(A;1))^{p-1}$. Applying Theorem 4.3, we conclude that μ verifies $\mathbf{T}_p(C)$. □

4.2.2. *Extension via a change of metric.* A change of metric technique, which is explained in the lemma below, enables us to reduce the study of the transport-entropy inequalities $\mathbf{T}_\alpha$ to the study of the inequalities $\mathbf{T}_p$, $p > 1$.

LEMMA 4.7. *Let α be a Young function satisfying the Δ_2-condition, and let $p_\alpha > 1$ be defined by (1.4). The function $x \mapsto \alpha(x)^{1/p_\alpha}$ is subadditive on $\mathbb{R}^+$:*

$$\alpha^{1/p_\alpha}(x+y) \le \alpha^{1/p_\alpha}(x) + \alpha^{1/p_\alpha}(y) \qquad \forall x, y \in \mathbb{R}^+.$$

As a consequence, $d_\alpha(x,y) = \alpha^{1/p_\alpha}(d(x,y))$, $x, y \in X$ is a distance on X.

The proof of Lemma 4.7 is at the end of the section.

PROOF OF THEOREM 4.1. Let α be a Young function and μ a probability on X. According to Lemma 4.7, the function $d_\alpha(x,y) = \alpha^{1/p_\alpha}(d(x,y))$, $x, y \in X$ is a metric on X. Furthermore, it is clear that μ verifies $(\tau) - \mathbf{LSI}_\alpha(1,A)$ [resp., $\mathbf{T}_\alpha(C)$] on (X,d) if and only if μ verifies $(\tau) - \mathbf{LSI}_{p_\alpha}(1,A)$ [resp., $\mathbf{T}_{p_\alpha}(C)$] on (X,d_α). We immediately deduce from Theorem 4.5 that if μ verifies $(\tau) - \mathbf{LSI}_\alpha(1,A)$, then it verifies $\mathbf{T}_\alpha(C)$, with $C = (a_{p_\alpha} \max(1; A))^{p_\alpha - 1}$, where

$$a_{p_\alpha} = \inf_{t\in(0,1)} \left\{ \frac{1}{t^{q_\alpha - 1}} \left(1 + \frac{p_\alpha^{q_\alpha}}{p_\alpha - 1} \int_0^t \frac{\varepsilon(u)}{u}\, du \right) \right\}$$

(since $\omega_{p_\alpha} \ge 1$) and ε defined in Lemma 4.2. Note that the constant C obtained using this approach is in general bigger than the constant obtained in the first proof of Theorem 4.1. □

4.3. *Proofs of the technical lemmas.*

PROOF OF LEMMA 4.2. Fix $t > 0$, $x \in X^n$ and $i \in \{1, \dots, n\}$. Then

$$tP_{\alpha,n}g(x) - Q_\alpha^{(i)}(tP_{\alpha,n}g)(x) = \sup_{z\in X} \{ (tP_{\alpha,n}g(x) - tP_{\alpha,n}g(\bar{x}^i z)) - \alpha(d(x_i,z)) \}.$$

Let y^x be such that $P_{\alpha,n}g(x) = g(y^x) - c(x, y^x)$, where $c(x,y) = \sum_{i=1}^n \alpha(d(x_i, y_i))$. By choosing $w = y^x$ in the expression below, it holds

$$\begin{aligned} P_{\alpha,n}g(x) - P_{\alpha,n}g(\bar{x}^i z) &= g(y^x) - c(x,y^x) - \sup_{w\in X^n}\{g(w) - c(\bar{x}^i z, w)\} \\ &\le c(\bar{x}^i z, y^x) - c(x, y^x) \\ &= \alpha(d(z, y_i^x)) - \alpha(d(x_i, y_i^x)) \\ &= d_\alpha^{p_\alpha}(z, y_i^x) - d_\alpha^{p_\alpha}(x_i, y_i^x) \\ &\le (d_\alpha(x_i, y_i^x) + d_\alpha(x_i, z))^{p_\alpha} - d_\alpha^{p_\alpha}(x_i, y_i^x), \end{aligned}$$

where, in the last line, we used the triangular inequality for the distance d_α defined in Lemma 4.7. Hence, optimizing yields

$$\begin{aligned} &tP_{\alpha,n}g(x) - Q_\alpha^{(i)}(tP_{\alpha,n}g)(x) \\ &\quad\le \sup_{z\in X}\{t[(d_\alpha(x_i, y_i^x) + d_\alpha(x_i, z))^{p_\alpha} - d_\alpha^{p_\alpha}(x_i, y_i^x)] - d_\alpha^{p_\alpha}(x_i, z)\} \\ &\quad= \sup_{r>0}\{t[(d_\alpha(x_i, y_i^x) + r)^{p_\alpha} - d_\alpha^{p_\alpha}(x_i, y_i^x)] - r^{p_\alpha}\} \\ &\quad= t\varepsilon(t)\, d_\alpha^{p_\alpha}(x_i, y_i^x) = t\varepsilon(t)\alpha(d(x_i, y_i^x)). \end{aligned}$$

Taking the sum, we get the result. □

Proof of Lemma 4.4. Let y^x be a point where the function $y \mapsto f(y) - \sum_{i=1}^n d^p(x_i, y)$ reaches its maximum. Then, for all $z \in X^n$, it holds

$$\begin{aligned} \sum_{i=1}^n d^p(x_i, y_i^x) &\le f(y^x) - f(z) + \sum_{i=1}^n d^p(x_i, z_i) \\ &\le L\left(\sum_{i=1}^n d^p(z_i, y_i^x)\right)^{1/p} + \sum_{i=1}^n d^p(x_i, z_i). \end{aligned}$$

Choosing $z = x$, we get $\sum_{i=1}^n d^p(x_i, z_i) \le L^q$.

Now, assume that (X, d) is geodesic. Then the product space $(X^n, d^{(n)})$ with $d^{(n)}(x,y) = (\sum_{i=1}^n d^p(x_i, y_i))^{1/p}$ is geodesic too. In the calculation above, take for z a t-midpoint of x and y^x; that is, choose $z \in X^n$ such that $d^{(n)}(x,z) = td^{(n)}(x, y^x)$ and $d^{(n)}(z, y^x) = (1-t)\, d^{(n)}(x, y^x)$, with $t \in [0,1]$. Then, letting $\ell = d^{(n)}(x, y^x)$, it holds $\ell^p \le L(1-t)\ell + t^p\ell^p$, and so $\frac{1-t^p}{1-t}\ell^{p-1} \le L$. Letting $t \to 1$ gives the result. □

Proof of Lemma 4.7. Let $\varphi(x) = \alpha^{1/p_\alpha}(x)/x$, $x > 0$. Then, by definition of p_α,

$$\varphi'_+(x) = \frac{\alpha^{1/p_\alpha}(x)(x\alpha'_+(x)/(p_\alpha\alpha(x)) - 1)}{x^2} \le 0.$$

So φ is nonincreasing on $(0,+\infty)$. Thus, if $x>y$,

$$\alpha^{1/p_\alpha}(x+y) = (x+y)\varphi(x(1+y/x)) \le (x+y)\varphi(x)$$
$$= \alpha^{1/p_\alpha}(x) + y\frac{\alpha^{1/p_\alpha}(x)}{x} \le \alpha^{1/p_\alpha}(x) + \alpha^{1/p_\alpha}(y). \qquad \square$$

5. Holley–Stroock perturbation lemma: Proof of Theorem 1.9. In this section, we prove Theorem 1.9.

PROOF OF THEOREM 1.9. The proof follows the line of the original proof [18]; see also [26]. Using the following representation of the entropy,

$$\mathrm{Ent}_\mu(g) = \inf_{t>0}\left\{\int\left(g\log\left(\frac{g}{t}\right) - g + t\right)d\mu\right\}$$

with $g = e^f$, we see that [since $g\log(\frac{g}{t}) - g + t \ge 0$]

$$\mathrm{Ent}_{\tilde\mu}(g) \le \frac{e^{\sup\varphi}}{Z}\,\mathrm{Ent}_\mu(g).$$

Since μ verifies $\mathbf{T}_\alpha(C)$, Theorem 2.1 implies that, for all $\lambda\in(0,1/C)$,

$$\mathrm{Ent}_{\tilde\mu}(e^f) \le \frac{e^{\sup\varphi}}{Z}\frac{1}{1-\lambda C}\int (f - Q_\alpha^\lambda f)e^f\,d\mu \le \frac{e^{\mathrm{Osc}(\varphi)}}{1-\lambda C}\int (f-Q_\alpha^\lambda f)e^f\,d\tilde\mu.$$

In other words, $\tilde\mu$ satisfies $(\tau)-\mathbf{LSI}_\alpha(\lambda, \frac{e^{\mathrm{Osc}(\varphi)}}{1-\lambda C})$, for any $\lambda\in(0,1/C)$. Now, applying Theorem 1.8, we conclude that $\tilde\mu$ verifies $\mathbf{T}_\alpha(\widetilde{C})$, with

$$\widetilde{C} = \kappa_{p_\alpha}\inf_{\lambda\in(0,1/C)}\left\{\frac{1}{\lambda}\frac{1}{(1-\lambda C)^{p_\alpha-1}}\right\}e^{(p_\alpha-1)\mathrm{Osc}(\varphi)}$$
$$= \kappa_{p_\alpha}\frac{p_\alpha^{p_\alpha}}{(p_\alpha-1)^{p_\alpha-1}}Ce^{(p_\alpha-1)\mathrm{Osc}(\varphi)} = \widetilde{\kappa}_{p_\alpha}Ce^{(p_\alpha-1)\mathrm{Osc}(\varphi)}. \qquad \square$$

APPENDIX: TECHNICAL RESULTS

In this appendix we prove the technical lemmas on Young functions we used during the paper. First, let us prove Lemma 1.6, that we restate below.

LEMMA A.1. *If α is a Young function satisfying the Δ_2-condition, then*

$$r_\alpha := \inf_{x>0}\frac{x\alpha'_-(x)}{\alpha(x)} \ge 1 \quad \text{and} \quad 1 < p_\alpha := \sup_{x>0}\frac{x\alpha'_+(x)}{\alpha(x)} < +\infty, \tag{A.1}$$

where α'_+ (resp., α'_-) denotes the right (resp., left) derivative of α.

PROOF. Using the convexity of α, we see that $\alpha(x)/x \le \alpha'_-(x)$. This shows that $r_\alpha \ge 1$. On the other hand, the function α is convex, so $\alpha(2x) \ge \alpha(x) + x\alpha'_+(x)$, for all $x>0$. Since α verifies the Δ_2-condition, there is some constant $K\ge 2$ such that $\alpha(2x)\le K\alpha(x)$. So we get $x\alpha'_+(x) \le (K-$

$1)\alpha(x)$, for all $x>0$. This proves that $p_\alpha<+\infty$. Let us show that $p_\alpha>1$. Otherwise we would have $r_\alpha=p_\alpha$ (since $\alpha'_-\le\alpha'_+$) and so $x\alpha'_-(x)/\alpha(x)=x\alpha'_+(x)/\alpha(x)=1$ for all $x>0$. This would imply that α is linear on $[0,\infty)$. This cannot happen, since by assumption Young functions are increasing and such that $\alpha'(0)=0$. So $p_\alpha>1$. □

Now let us prove Lemmas 3.2 and 3.3 whose statements are summarized below. Recall that the function ξ_α is defined by

$$\xi_\alpha(x):=\sup_{u>0}\frac{\alpha^*(x\alpha'_+(u))}{x\alpha(u)},\qquad x>0.$$

LEMMA A.2.

- *Let α be a Young function satisfying the Δ_2-condition, and let $1\le r_\alpha\le p_\alpha$, $p_\alpha>1$ be the numbers defined by (A.1). Then, it holds*

$$\xi_\alpha(x)\le(p_\alpha-1)\max(x^{1/(p_\alpha-1)};x^{1/(r_\alpha-1)})\qquad\forall x>0, \tag{A.2}$$

  with the convention $t^\infty=0$ if $t\le1$ and ∞ otherwise.
- *Let $p_1\ge2$ and $p_2>1$ and let $\alpha=\alpha_{p_1,p_2}$; then $p_\alpha=\max(p_1,p_2)$, and it holds*

$$\xi_\alpha(x)=(p_\alpha-1)x^{1/(p_\alpha-1)}\qquad\forall x\le1.$$

  Moreover, for $x\ge1$, it holds

$$\xi_\alpha(x)=\begin{cases}p_1\left(\dfrac{1}{q_2}x^{1/(p_2-1)}+\left(\dfrac{1}{q_1}-\dfrac{1}{q_2}\right)\dfrac{1}{x}\right), & \text{if } p_1\ge p_2,\\ \max((p_1-1)x^{1/(p_1-1)};(p_2-1)x^{1/(p_2-1)}), & \text{if } p_1\le p_2,\end{cases}$$

  where $q_1=p_1/(p_1-1)$ and $q_2=p_2/(p_2-1)$.

PROOF. Defining $\omega(x)=\sup_{u>0}\frac{\alpha^*(ux)}{\alpha^*(u)}$, for all $x\ge0$, we get

$$\xi_\alpha(x)\le\frac{\omega(x)}{x}\sup_{u>0}\frac{\alpha^*(\alpha'_+(u))}{\alpha(u)}\qquad\forall x>0.$$

From the convexity inequality $\alpha(x)\ge\alpha(u)+(x-u)\alpha'_+(u)$, $x,u\ge0$, we deduce immediately that $\alpha^*(\alpha'_+(u))=u\alpha'_+(u)-\alpha(u)$, for all $u\ge0$. Thus

$$\sup_{u>0}\frac{\alpha^*(\alpha'_+(u))}{\alpha(u)}=p_\alpha-1.$$

So, all we have to show is that $\omega(x)\le\max(x^{p_\alpha/(p_\alpha-1)};x^{r_\alpha/(r_\alpha-1)})$, for all $x\ge0$.

Define $\varphi(u)=\alpha(u)/u^{p_\alpha}$ and $\psi(u)=\alpha(u)/u^{r_\alpha}$, for all $u>0$. As in the proof of Lemma 4.7, a simple calculation shows that φ is nonincreasing, and

ψ is nondecreasing. As a result,

$$\alpha(tu) \le t^{p_\alpha}\alpha(u) \qquad \forall u \ge 0,\ \forall t \ge 1,$$
$$\alpha(tu) \le t^{r_\alpha}\alpha(u) \qquad \forall u \ge 0,\ \forall t \in [0,1].$$

Taking the Fenchel–Legendre transform yields

$$\alpha^*(v/t) \ge t^{p_\alpha}\alpha^*(v/t^{p_\alpha}) \qquad \forall v \ge 0,\ \forall t \ge 1,$$
$$\alpha^*(v/t) \ge t^{r_\alpha}\alpha^*(v/t^{r_\alpha}) \qquad \forall v \ge 0,\ \forall t \in [0,1].$$

Equivalently,

$$\alpha^*(ux) \le x^{p_\alpha/(p_\alpha-1)}\alpha^*(u) \qquad \forall u \ge 0,\ \forall x \in [0,1],$$
$$\alpha^*(ux) \le x^{r_\alpha/(r_\alpha-1)}\alpha^*(u) \qquad \forall u \ge 0,\ \forall x \ge 1.$$

And since $r_\alpha \le p_\alpha$, we conclude that $\omega(x) \le \max(x^{p_\alpha/(p_\alpha-1)}; x^{r_\alpha/(r_\alpha-1)})$, $x \ge 0$.

Now, let us calculate $\xi_{\alpha_{p_1,p_2}}$, for $p_1 \ge 2$, $p_2 > 1$. First observe that $\xi_{\lambda\alpha} = \xi_\alpha$ for all $\lambda > 0$. It will be more convenient to do the calculation with the function $\alpha := \bar\alpha_{p_1,p_2} = \frac{1}{p_1}\alpha_{p_1,p_2}$. Let us denote by $q_1 = \frac{p_1}{p_1-1}$, $q_2 = \frac{p_2}{p_2-1}$, the conjugate exponents of p_1 and p_2. Then the following identity holds: $\alpha^* = \bar\alpha_{q_1,q_2}$. Let us show that

$$\xi_\alpha(x) = (p_\alpha - 1)x^{1/(p_\alpha-1)}$$

for $x \le 1$. The case $x > 1$ is similar and left to the reader. Define

$$\varphi(u) = \frac{\alpha^*(xu)}{\alpha \circ \alpha'^{-1}(u)}, \qquad u > 0.$$

We have to distinguish three cases:

$$\xi_\alpha(x) = \frac{1}{x}\max\Big(\sup_{u\le 1}\varphi(u);\ \sup_{1\le u\le 1/x}\varphi(u);\ \sup_{u\ge 1/x}\varphi(u)\Big).$$

Case 1. $0 < u \le 1$. Then $\varphi(u) = (p_1 - 1)x^{q_1}$.

Case 2. $1 \le u \le 1/x$. Then

$$\varphi(u) = \frac{x^{q_1}}{q_1}\frac{u^{q_1}}{u^{q_2}/p_2 + 1/p_1 - 1/p_2}.$$

If $p_1 \ge p_2$, then the function φ is nonincreasing on $[1, 1/x]$, and so

$$\sup_{1\le u\le 1/x}\varphi(u) = \varphi(1) = (p_1 - 1)x^{q_1}.$$

If $p_1 \le p_2$, then the function φ is nondecreasing on $[1, 1/x]$, and so

$$\sup_{1\le u\le 1/x}\varphi(u) = \varphi(1/x).$$

Case 3. $u \geq 1/x$. Then

$$\varphi(u) = \frac{(xu)^{q_2}/q_2 + 1/q_1 - 1/q_2}{u^{q_2}/p_2 + 1/p_1 - 1/p_2}.$$

If $p_1 \geq p_2$, the function φ is nonincreasing on $[1/x, \infty)$, and so

$$\sup_{u \geq 1/x} \varphi(u) = \varphi(1/x).$$

If $p_1 \leq p_2$, the function φ is nondecreasing on $[1/x, \infty)$, and so

$$\sup_{u \geq 1/x} \varphi(u) = \lim_{u \to \infty} \varphi(u) = (p_2 - 1)x^{q_2}.$$

Observe, in particular, that φ never reaches its supremum at $u = 1/x$. We conclude that

$$\sup_{u>0} \varphi(u) = \max((p_1 - 1)x^{q_1}; (p_2 - 1)x^{q_2}),$$

and so

$$\begin{aligned} \xi_\alpha(x) &= \max((p_1 - 1)x^{1/(p_1-1)}; (p_2 - 1)x^{1/(p_2-1)}) \\ &= (\max(p_1; p_2) - 1)x^{1/(\max(p_1;p_2)-1)}. \end{aligned}$$

Since $p_\alpha = \max(p_1; p_2)$, the proof is complete. □

Acknowledgments. We thank the referees for a careful reading of our paper and for providing useful feedback.

REFERENCES

[1] Ané, C., Blachère, S., Chafaï, D., Fougères, P., Gentil, I., Malrieu, F., Roberto, C. and Scheffer, G. (2000). *Sur les Inégalités de Sobolev Logarithmiques. Panoramas et Synthèses [Panoramas and Syntheses]* **10**. Société Mathématique de France, Paris.

[2] Barthe, F. and Roberto, C. (2008). Modified logarithmic Sobolev inequalities on $\mathbb{R}$. *Potential Anal.* **29** 167–193.

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

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

[5] Bobkov, S. G. and Ledoux, M. (1997). Poincaré's inequalities and Talagrand's concentration phenomenon for the exponential distribution. *Probab. Theory Related Fields* **107** 383–400.

[6] Cattiaux, P. and Guillin, A. (2006). On quadratic transportation cost inequalities. *J. Math. Pures Appl. (9)* **86** 341–361.

[7] Gentil, I. (2008). From the Prékopa–Leindler inequality to modified logarithmic Sobolev inequality. *Ann. Fac. Sci. Toulouse Math. (6)* **17** 291–308.

[8] Gentil, I., Guillin, A. and Miclo, L. (2005). Modified logarithmic Sobolev inequalities and transportation inequalities. *Probab. Theory Related Fields* **133** 409–436.
[9] Gentil, I., Guillin, A. and Miclo, L. (2007). Modified logarithmic Sobolev inequalities in null curvature. *Rev. Mat. Iberoam.* **23** 235–258.
[10] Gozlan, N. (2007). Characterization of Talagrand's like transportation-cost inequalities on the real line. *J. Funct. Anal.* **250** 400–425.
[11] Gozlan, N. (2009). A characterization of dimension free concentration in terms of transportation inequalities. *Ann. Probab.* **37** 2480–2498. MR2573565
[12] Gozlan, N. and Léonard, C. (2007). A large deviation approach to some transportation cost inequalities. *Probab. Theory Related Fields* **139** 235–283.
[13] Gozlan, N. and Léonard, C. (2010). Transport inequalities. A survey. *Markov Process. Related Fields* **16** 635–736.
[14] Gozlan, N., Roberto, C. and Samson, P. M. (2010). From concentration to logarithmic Sobolev and Poincaré inequalities. *J. Funct. Anal.* **260** 1491–1522.
[15] Gozlan, N., Roberto, C. and Samson, P. M. (2011). A new characterization of Talagrand's transport-entropy inequalities and applications. *Ann. Probab.* **39** 857–880.
[16] Gross, L. (1975). Logarithmic Sobolev inequalities. *Amer. J. Math.* **97** 1061–1083.
[17] Helffer, B. (2002). *Semiclassical Analysis, Witten Laplacians, and Statistical Mechanics. Series in Partial Differential Equations and Applications* **1**. World Scientific, River Edge, NJ.
[18] Holley, R. and Stroock, D. (1987). Logarithmic Sobolev inequalities and stochastic Ising models. *J. Stat. Phys.* **46** 1159–1194.
[19] Ledoux, M. (2001). *The Concentration of Measure Phenomenon. Mathematical Surveys and Monographs* **89**. Amer. Math. Soc., Providence, RI.
[20] Lott, J. and Villani, C. (2007). Hamilton–Jacobi semigroup on length spaces and applications. *J. Math. Pures Appl. (9)* **88** 219–229.
[21] Marton, K. (1986). A simple proof of the blowing-up lemma. *IEEE Trans. Inform. Theory* **32** 445–446.
[22] Maurey, B. (1991). Some deviation inequalities. *Geom. Funct. Anal.* **1** 188–197.
[23] Monge, G. (1781). Mémoire sur la théorie des déblais et des remblais. *Histoire de L'Académie Royale des Sciences de Paris* 666–704.
[24] Otto, F. and Villani, C. (2000). Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality. *J. Funct. Anal.* **173** 361–400. MR1760620
[25] Papageorgiou, I. (2011). A note on the modified log-Sobolev inequality. *Potential Anal.* **35** 275–286.
[26] Royer, G. (1999). *Une Initiation aux Inégalités de Sobolev Logarithmiques*. Société Mathématique de France, Paris.
[27] Samson, P. M. (2000). Concentration of measure inequalities for Markov chains and Φ-mixing processes. *Ann. Probab.* **28** 416–461.
[28] Stam, A. J. (1959). Some inequalities satisfied by the quantities of information of Fisher and Shannon. *Information and Control* **2** 101–112.
[29] Talagrand, M. (1991). A new isoperimetric inequality and the concentration of measure phenomenon. In *Geometric Aspects of Functional Analysis (1989–90). Lecture Notes in Math.* **1469** 94–124. Springer, Berlin.
[30] Talagrand, M. (1996). Transportation cost for Gaussian and other product measures. *Geom. Funct. Anal.* **6** 587–600.

[31] Villani, C. (2009). *Optimal Transport: Old and New. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]* **338**. Springer, Berlin.

N. Gozlan
P.-M. Samson
Laboratoire d'Analyse et de Mathématiques Appliquées (UMR CNRS 8050)
Université Paris Est Marne la Vallée
5 bd Descartes
77454 Marne la Vallée Cedex 2
France
E-mail: nathael.gozlan@univ-mlv.fr
paul-marie.samson@univ-mlv.fr

C. Roberto
Laboratoire d'Analyse et de Mathématiques Appliquées (UMR CNRS 8050)
Université Paris Est Marne la Vallée
5 bd Descartes
77454 Marne la Vallée Cedex 2
France
and
Université Paris Ouest Nanterre la Défense
MODAL'X, EA 3454
200 avenue de la République 92000 Nanterre
France
E-mail: croberto@math.cnrs.fr