On quadratic transportation cost inequalities

Patrick Cattiaux \textsuperscript{a,\ast}, Arnaud Guillin \textsuperscript{b}

\textsuperscript{a} Ecole Polytechnique, CMAP, F-91128 Palaiseau cedex, and Université Paris X Nanterre, équipe MODAL'X, UFR SEGMI, 200 avenue de la République, F-92001 Nanterre, Cedex, France

\textsuperscript{b} CEREMADE, Université Paris IX Dauphine, F-75775 Paris cedex, CNRS 7534, France

Received 10 November 2005
Available online 10 August 2006

Abstract
In this paper we study quadratic transportation cost inequalities. To this end we introduce new families of inequalities (for quadratic transportation cost and for relative entropy) that are shown to be equivalent to the Poincaré inequality. This allows us to give some examples of measures satisfying $T_2$ but not the logarithmic-Sobolev inequality.

© 2006 Elsevier SAS. All rights reserved.

MSC: 47D07; 60E15; 60G10

Keywords: Transportation inequalities; Spectral gap; Gaussian concentration

1. Introduction, framework and main results

Transportation inequalities recently deserved a lot of interest, especially in connection with the concentration of measure phenomenon (see [15,16]). Links with other renowned functional inequalities, in particular logarithmic-Sobolev inequalities, were also particularly studied (see [5,18,4,16],...), as no direct or tractable criteria were available for this kind of inequalities.

Given a metric space $(E, d)$ equipped with its Borel $\sigma$ field, the $L^p$ Wasserstein distance between two probability measures $\mu$ and $\nu$ on $E$ is defined as

$$W_p(\mu, \nu) = \left( \inf_{\pi} \int_{E \times E} d^p(x, y) \pi(dx, dy) \right)^{1/p},$$

(1.1)
where \( \pi \) describes the set of all coupling of \((\mu, v)\), i.e., the set of all probability measures on the product space with marginal distributions \(\mu\) and \(v\).

A probability measure \(\mu\) is said to satisfy the \(T_p(C)\) transportation cost inequality if for all probability measure \(v\),
\[
W_p(\mu, v) \leq \sqrt{2CH(\nu, \mu)},
\]
where \(H(\nu, \mu)\) stands for the Kullback–Leibler information (or relative entropy), i.e.,
\[
H(\nu, \mu) = \begin{cases} 
\int \log \left( \frac{d\nu}{d\mu} \right) d\nu & \text{if } \nu \ll \mu; \\
+\infty & \text{otherwise.}
\end{cases}
\]

As shown by K. Marton [17], \(T_1\) implies a Gaussian type concentration for \(\mu\).

Let us briefly recall the general argument, we shall use later.

For any Borel set \(A\) with measure \(\mu(A) \geq \frac{1}{2}\) introduce \(A_r^c = \{x, d(x, A) \geq r\}\) and \(d_{\mu A} = \frac{1}{\mu(A)} d_{\mu}\). Set \(B\) for \(A_r^c\) and assume that \(W_1(\nu, \mu) \leq \phi(H(\nu, \mu))\) for all \(\nu\). Then,
\[
r \leq W_1(\mu_B, \mu_A) \leq W_1(\mu_B, \mu) + W_1(\mu, \mu_A) \leq \phi(H(\mu_A, \mu)) + \phi(H(\mu_B, \mu)) \leq \phi\left( H(\mu_A, \mu) \right) + \phi\left( H(\mu_B, \mu) \right) = \phi\left( \log \frac{1}{\mu(A)} \right) + \phi\left( \log \frac{1}{\mu(A_r^c)} \right).
\]

When \(\phi(u) = \sqrt{2Cu}\) we immediately obtain:
\[
\mu(A_r^c) \leq \exp\left(-1/2C\left(r - \sqrt{2C \log \left( \frac{1}{\mu(A)} \right)} \right)^2\right).
\]

Hence criteria for \(T_1\) to hold are very useful. Such a criterion was first obtained by Bobkov and Götze [5, Theorem 3.1] and recently discussed by Djellout, Guillin and Wu [13, Theorem 2.3] where the following is proved.

**Theorem 1.4.** [13] \(\mu\) satisfies \(T_1\) if and only if there exist \(\varepsilon > 0\) and \(x_0 \in E\) such that
\[
\int_E e^{\varepsilon d^2(x, x_0)} \mu(dx) < +\infty. \tag{E I_\varepsilon(2)}
\]

Unfortunately \(T_1\) is not well adapted to dimension free bounds, while \(T_2\) is, as shown by Talagrand [21]. The first example of measure satisfying \(T_2\) is the standard Gaussian measure \([21]\), for which \(C = 1\). When \(E\) is a complete smooth Riemannian manifold of finite dimension, with \(d\) the geodesic distance and \(d_{\mu}\) the volume measure, Otto and Villani [18] have studied the \(T_2\) property for absolutely continuous probability measures (Boltzmann measures), \(\mu(dx) = e^{-V(x)} dx\), (B.M.)

for \(V \in C^2(E)\) in connection with the logarithmic-Sobolev inequality. Their method was recently improved by Wang [26] in order to skip the curvature assumption made in [18].

In the sequel we shall assume that \(\mu\) is a Boltzmann measure with \(V \in C^3\), and that the diffusion process built on \(E\) with generator \(L = 1/2 \text{div}(\nabla) - 1/2 \nabla V. \nabla\) is nonexplosive. We denote \((P_t)_{t \geq 0}\) the associated semigroup.

This is assumption (A) in [26]. Conditions for nonexplosion are known. Here are two different among others when \(E = \mathbb{R}^d\):

- there exists some \(\psi\) such that \(\psi(x) \to +\infty\) as \(|x| \to +\infty\) and \(\Delta \psi - \nabla V. \nabla \psi\) is bounded from above,
- \(\int |\nabla V|^2 d\mu < +\infty\).

For the first two see, e.g., [20, p. 26] (replacing \(V\) therein by \(\psi\)), for the second one see, e.g., [10].

Then we have:

**Theorem 1.5.** [18,4,26] (see also [12]) If \(\mu\) satisfies the logarithmic-Sobolev inequality (L.S.I.),
\[
\int g^2 \log(g^2) d\mu - \left( \int g^2 d\mu \right) \log\left( \int g^2 d\mu \right) \leq 2C \int |\nabla g|^2 d\mu.
\]
for all smooth \( g \), then \( \mu \) satisfies \( T_2(C) \).

A partial converse of Theorem 1.5 is also shown in [18, Corollary 3.1], namely:

**Theorem 1.6.** [18,4] Let \( E = \mathbb{R}^n \). If \( \mu \) satisfies \( T_2(C) \) and the curvature assumption:

\[
\text{Hess}(V) \geq K \text{Id}_n,
\]

for some \( K \in \mathbb{R} \), then \( \mu \) satisfies a logarithmic-Sobolev inequality (with some new constant \( \tilde{C} \)), provided

\[
1 + KC > 0.
\]

The latter restriction is very important and has to be compared with Wang’s results ([24] and [25]) telling that a logarithmic-Sobolev inequality holds provided the curvature assumption above and the integrability condition \((EI_\varepsilon(2))\) in Theorem 1.4 hold with

\[
\varepsilon + K > 0.
\]

In other words, according to Theorems 1.4 and 1.6, under the curvature assumption, log-Sobolev, \( T_1(C_1) \), \( T_2(C_2) \) are all equivalent for appropriate constants \( C_1 \) and \( C_2 \).

Whether this equivalence holds without restrictions on the constants or not was left open by these authors. One aim of this paper is to show that this equivalence does not hold. Before stating a more precise result, let us complete the picture.

On one hand, as shown by Otto and Villani (see [4, Section 4.1] for another approach).

**Theorem 1.7.** If \( \mu \) satisfies \( T_2(C) \) then \( \mu \) satisfies the Poincaré (or spectral gap) inequality (S.G.I.), i.e., for all smooth \( f \),

\[
\text{Var}_\mu(f) \leq C \int |\nabla f|^2 d\mu.
\]

This result gives us a first hint on what should be the difference between \( T_1 \) and \( T_2 \) as \( T_1 \) is well known to hold when (S.G.I.) fails (see [13, Remark 2.4]).

On the other hand, the difference between \( T_2 \) and \( T_1 \) is only concerned with small entropies due to the following elementary:

**Lemma 1.8.** Assume that \( \mu \) satisfies \((EI_\varepsilon(p))\) for some \( \varepsilon > 0 \). Then there exists a constant \( C(\varepsilon) \) such that for all \( v \) satisfying \( H(v, \mu) \geq 1 \), \( W_p^p(v, \mu) \leq C(\varepsilon) H(v, \mu) \).

Here \((EI_\varepsilon(p))\) is defined as in Theorem 1.4 with \( d^p \) instead of \( d^2 \).

Hence the transportation inequalities \( T_2 \) and \( T_1 \) are “equivalent” for large entropy. Since Marton’s method is essentially concerned with large entropy, \( T_2 \) cannot furnish a better concentration result than \( T_1 \). The main interest of \( T_2 \) for the concentration of measure phenomenon is thus that \( T_2 \) can be tensorized.

At this point we shall mention that the proof of Lemma 1.8 is using the trivial independent coupling. We learned from F. Bolley and C. Villani [7] that, using a less trivial coupling in [22], this statement can be greatly improved, in particular:

**Proposition 1.9.** (Bolley and Villani)

\[
(EI_\varepsilon(p)) \implies W_p^p(v, \mu) \leq C(\varepsilon)(H(v, \mu) + H^{1/2}(v, \mu)).
\]

Bolley and Villani are then able to get back Theorem 1.4, i.e., \((EI_\varepsilon(2))\) is equivalent to the transportation inequality \( T_1 \), but with some better constant than in [13].

Let us come to the contents of the present paper where we shall mainly focus on \( W_2 \).
In Section 2 we shall show that (S.G.I.) implies some quadratic transportation inequality for measures \( \nu \) with a bounded density. Actually we prove an interpolation result between (S.G.I.) and (L.S.I.) through a family of inequalities \( I(\alpha) \) introduced by Latala and Oleszkiewicz (see [14]) for \( 0 \leq \alpha \leq 1 \),

\[
I(\alpha) \sup_{p \in [1,2]} \int f^2 \, d\mu - \left( \int |f|^p \, d\mu \right)^{2/p} \leq C(\alpha) \int |\nabla f|^2 \, d\mu. \tag{1.10}
\]

Note that \( I(0) \) is the Poincaré inequality and \( I(1) \) reduces to the logarithmic-Sobolev inequality. Our first result is the following:

**Theorem 1.11.** Let \( \mu \) be as above. If \( I(\alpha) \) holds then for all \( \nu \) such that \( \| d\nu / d\mu \|_{\infty} \leq K \) the following modified transportation inequality holds:

\[
W_2(\nu, \mu) \leq D(\alpha, K) \sqrt{C(\alpha)H(\nu, \mu)},
\]

where

\[
D(\alpha, K) = 8 \exp \left( \frac{1 - \alpha}{2} (1 - \log(1 - \alpha)) \right) (\log K)^{(1-\alpha)/2}, \quad \text{for } K \geq e^{1-\alpha},
\]

and

\[
D(\alpha, K) = 8 \sqrt{K} \quad \text{for } K \leq e^{1-\alpha}.
\]

Remark that the previous theorem and Marton’s trick allow to recover the concentration property shown in [14]. Indeed, recall (1.3) and remark that the interesting \( K \) is given by \( K = 1/\mu(A^c_r) \). We immediately see that if \( I(\alpha) \) holds, \( \mu(A^c_r) \) behaves like \( \exp(-Cr^2/(2-\alpha)) \). Also note that for \( \alpha = 1 \) we recover Otto–Villani result, since \( K \) can be arbitrarily chosen.

We refer to [26,23,3,11] and [2] for more refined results in connection with \( I(\alpha) \).

If the meaning of a transportation inequality reduced to bounded densities is not clear, the previous theorem nevertheless allows us to obtain as a first consequence the following (weak) transportation inequality proved in Section 3.

**Corollary 1.12.** Let \( \mu \) be as above. If \( (EI_\varepsilon(2)) \) and \( I(\alpha) \) are satisfied, there exists some constant \( C \) such that

\[
W_2^2(\nu, \mu) \leq C \left( 1 + (1 - \alpha)^{-1/(2-\alpha)} \left( \log^+ \left( 1/H(\nu, \mu) \right) \right)^{(1-\alpha)/(2-\alpha)} \right) H(\nu, \mu).
\]

Remark that this corollary presents an interpolation between Poincaré and log-Sobolev inequality, which is expected to tensorize, but it is not dimension free, as a \( \left( \log(n) \right)^{(1-\alpha)/(2-\alpha)} \) appears in the tensorization procedure, which is however better than the factor \( n \) obtained with the sole \( T_1 \). We will see that we have to impose conditions slightly stronger than Poincaré’s inequality (at least in the real line case) to get rid of the extra \( \sqrt{\log} \) term leading thus to the true \( T_2 \) inequality.

Actually one can get some equivalences between various inequalities for bounded functions. This will be done in Section 2. The result is the following:

**Theorem 1.13.** Let \( \mu \) be as above. Then (up to the constants) the following statements are equivalent:

1. \( I(0) \) (i.e., the Poincaré inequality) holds,
2. the (modified) transportation inequality in Theorem 1.11 holds for \( \alpha = 0 \),
3. the following (restricted) log-Sobolev inequality holds: for all nonnegative \( h \) such that \( \int h \, d\mu = 1 \),

\[
\int h \log h \, d\mu \leq C \left( 1 + \log(\|h\|_{\infty}) \right) \int \frac{|\nabla h|^2}{h} \, d\mu,
\]

4. there exist some \( C \) and some \( K > 1 \) such that the previous log-Sobolev inequality holds for all \( h \) as above and bounded by \( K \).
Corollary 1.15. Assume that $h$ is such that $h \geq a$ for some constant $a$. If one wants to improve this result, one has to truncate with some function $C(M)$ provided $\text{Hess}(V) \geq R \text{Id}$ for some $R \in \mathbb{R}$ and $E = \mathbb{R}^N$.

We shall give a proof of Corollary 1.16 in Section 4.3. In both corollaries one may reduce a little bit the set of allowed densities assuming in addition that $h \geq a$ for some $a > 0$.

To get rid of the curvature assumption one has to call upon the methods in [4], namely Herbst’s argument and the beautiful characterization of $T_2$ obtained by Bobkov and Götze [5].

Theorem 1.17. Assume that $(EI_s(2))$ holds. If the restricted logarithmic-Sobolev inequality,

$$\int f^2 \log f^2 \, d\mu - \left( \int f^2 \, d\mu \right) \log \left( \int f^2 \, d\mu \right) \leq C \int |\nabla f|^2 \, d\mu,$$

holds for all

$$f^2 \leq \left( \int f^2 \, d\mu \right) e^{\eta(d^2(x,x_0) + f^2(y,x_0)) \mu(dy)} \quad (1.18)$$

for some $\eta < \varepsilon/2$, then $T_2$ holds.

We have chosen to write the hypotheses in a slightly different form but this result is of course the same as Corollary 1.16 without the curvature assumption. Note that it generalizes slightly the principal result in [4]: a full logarithmic-Sobolev inequality is too strong to get $T_2$.

The proof of this theorem will be given at the end of Section 4.4. It is an almost immediate adaptation of Section 3.3 in [4]. Since it is the most general result of the section, the reader should ask about the interest of the remainder of Section 4. As we said some of the results therein have their own interest, but the comparison between both approaches (Otto and Villani coupling on one hand, Infimum convolution on the other hand) is more interesting. Indeed both approaches are qualitatively very different: Otto–Villani’s coupling yield local results (if one wants to get some estimate for the Wasserstein distance for a single $h$, one only needs to look at $P_t h$) while the infimum convolution approach is global (since variational) in nature. In particular, for the latter approach we did not succeed in reducing the problem to small entropies and/or bounded below densities.

In the final Section 5, we give some Hardy’s like conditions implying a $T_2$ inequality for measure on the real line. It finally enables us to build explicitly a potential $V$ such that $\mu$ satisfies $T_2$ but does not satisfy a logarithmic-Sobolev inequality, however with unbounded below curvature. These examples show that $T_2$ is strictly weaker than...
(L.S.I.), which was in fact the primary goal of the authors, the second one being still an open question: an explicit characterization of the $T_2$ inequality.

2. Modified transportation and functional inequalities

In this section we shall discuss several functional inequalities involving bounded functions. We start with the proof of Theorem 1.11.

Proof of Theorem 1.11. Let $\nu$ be a probability measure such that $h = \frac{d\nu}{d\mu}$ satisfies $0 < \beta \leq h(x) \leq K$. We assume first that $h \in D$, i.e., is the sum of a constant and a $C^\infty$ function with compact support.

Let $P_t$ denote the $\mu$ symmetric semigroup with generator $L = \frac{1}{2} \text{div}(\nabla) - \frac{1}{2} \nabla \nu \cdot \nabla$, and define $\nu_t = (P_t h) \mu$.

Our method relies on Otto–Villani’s coupling [18], refined by Wang [26], whose idea is the following: to provide a coupling between $\nu_t$ and $\nu_{t+s}$ as $\pi_s(dx, dy) = \nu_t(dx) \delta_{\phi_s(y)}(dy)$ where $\phi_s$ is the well defined unique (under our assumptions) solution of the p.d.e.

$$
\frac{d}{ds} \phi_s = -\xi_{t+s} \circ \phi_s, \quad \phi_0 = \text{Id}, \ S \geq 0,
$$

with $\xi_{t+s}(x) = \nabla \log P_{t+s} h(x)$.

Then, according to Otto and Villani [18, Lemma 2 (or more exactly its proof)], or Wang [26, Section 3],

$$
A = \frac{d^+}{dt} (-W_2(\nu_t, \mu)) \leq \lim_{s \to 0^+} \frac{1}{s} W_2(\nu_t, \nu_{t+s}) \leq 2 \left( \int |\nabla \sqrt{P_t h}|^2 d\mu \right)^{1/2}.
$$

(2.1)

Using $I(\alpha)$ we obtain for all $1 < p < 2$,

$$
A \leq \frac{2 \sqrt{C(\alpha)(2-p)}}{\sqrt{1 - (\int (P_t h)^{p/2} d\mu)^{2/p}}}
$$

(2.2)

Now using a similar argument as in Lemma 3.1 in [26] or simply the fact that $D$ is a nice core for the diffusion semigroup, the following computation is rigorous:

$$
\frac{d}{dt} \left( 1 - \left( \int (P_t h)^{p/2} d\mu \right)^{1/p} \right) = -\frac{1}{2} \left( \int (P_t h)^{p/2} d\mu \right)^{1/p-1} \int (P_t h)^{p/2-1} L P_t h \, d\mu
$$

$$
= \frac{1}{4} \left( \int (P_t h)^{p/2} d\mu \right)^{1/p-1} \int \left( \frac{p}{2} - 1 \right) (P_t h)^{p/2-2} |\nabla P_t h|^2 d\mu
$$

$$
= \frac{1}{2} \left( \int (P_t h)^{p/2} d\mu \right)^{1/p-1} \int (p-2)(P_t h)^{p/2-1} |\nabla P_t h|^2 d\mu \leq 0.
$$

(2.3)

But since $h \leq K$, $P_t h \leq K$ hence according to (2.2) and (2.3),

$$
A \leq \frac{2 \sqrt{C(\alpha)(2-p)}}{\sqrt{1 - (\int (P_t h)^{p/2} d\mu)^{1/p}} \sqrt{1 + (\int (P_t h)^{p/2} d\mu)^{1/p}}}.
$$

(2.4)

For the latter inequality we have used $\int (P_t h)^{p/2} d\mu \leq 1$.

It remains to integrate in $t$. Since $I(\alpha)$ implies (S.G.I.), we know that $P_t h$ goes to $1$ in $L^2(\mu)$ as $t$ goes to infinity. Arguing as in [26, p. 10], one can show that $W_2(\nu_t, \mu)$ goes to $0$ as $t$ goes to $\infty$, so that we have obtained:
\[
W_2(v, \mu) \leq 8\sqrt{C(\alpha)}(2 - p)^{\alpha/2 - 1}K^{1 - p/2} \sqrt{1 - \left( \int h^{p/2} \, d\mu \right)^{1/p}}
\leq 8\sqrt{C(\alpha)}(2 - p)^{\alpha/2 - 1}K^{1 - p/2} \sqrt{1 - \left( \int h^{p/2} \, d\mu \right)^{2/p}}.
\]

(2.5)

Now we shall use the two following elementary inequalities for \( p \in [1, 2) \):

- \( 1 - u^{2/p} \leq \frac{2}{p}(1 - u) \) for \( u \in [0, 1] \),

- \( \xi \log \xi + 1 - \xi \geq 0 \) for \( \xi > 0 \).

The latter yields \( \log \xi^k \geq 1 - \xi - k \), hence \( \xi \log \xi^k \geq \xi - \xi^{1-k} \) and finally for \( k = 1 - p/2, (1 - p/2)\xi \log \xi \geq \xi - \xi^{p/2} \). We apply this with \( h(x) = \xi \), integrate with respect to \( \mu \) and use the former inequality in order to get:

\[
1 - \left( \int h^{p/2} \, d\mu \right)^{2/p} \leq \frac{2}{p} \left( 1 - \frac{p}{2} \right) H(v, \mu).
\]

(2.6)

Plugging (2.6) into (2.5) furnishes (using \( p \geq 1 \)):

\[
W_2(v, \mu) \leq 8\sqrt{C(\alpha)}(2 - p)^{\alpha/2 - 1}K^{1 - p/2} \sqrt{H(v, \mu)}.
\]

(2.7)

It is now enough to optimize in \( p \), just taking care that \( p \geq 1 \). The optimal value is obtained for \( 2 - p = 1 - \alpha \log K \) if \( K \geq e^{1-\alpha} \) and for \( p = 1 \) otherwise. A simple calculation yields the exact bound in Theorem 1.11.

It remains to extend the result to densities \( h \) that are no more bounded away from 0, by using standard tools.

This modified transportation inequality does not seem tensorizable. Now we come to the proof of Theorem 1.13.

**Proof of Theorem 1.13.** The first implication is given by the previous theorem. The equivalence between (1) and (2) follows from Otto–Villani’s way of proof of \( T_2 \Rightarrow (S.G.I.) \). Namely choose some smooth \( f \) with compact support (hence bounded) such that \( \int f \, d\mu = 0 \), and set for \( \varepsilon \) small enough \( \mu_\varepsilon = (1 + \varepsilon f)\mu \). Recall that

\[
H(\mu_\varepsilon, \mu)/\varepsilon^2 \to \int f^2 \, d\mu.
\]

By Taylor formula at order 2, as \( f \) is smooth and compactly supported, one may find a constant \( C \) such that for all \( x, y \),

\[
f(x) - f(y) \leq |\nabla f(y)||x - y| + C|x - y|^2.
\]

Denote by \( \pi_\varepsilon \) an “optimal coupling” for the Wasserstein distance between \( \mu \) and \( \mu_\varepsilon \), then for \( \varepsilon \) small enough:

\[
\int f \, d \left( \frac{\mu_\varepsilon - \mu}{\varepsilon} \right) = \frac{1}{\varepsilon} \int (f(x) - f(y)) \, d\pi_\varepsilon
\leq \frac{1}{\varepsilon} \int |\nabla f(y)||x - y| \, d\pi_\varepsilon + \frac{C}{\varepsilon} \int |x - y|^2 \, d\pi_\varepsilon
\leq \frac{1}{\varepsilon} \int |\nabla f|^2 \, d\mu W_2(\mu_\varepsilon, \mu) + \frac{C}{\varepsilon} W_2^2(\mu_\varepsilon, \mu)
\leq 8\sqrt{1 + \varepsilon \|f\|_\infty} \sqrt{\int |\nabla f|^2 \, d\mu \sqrt{C(0)H(\mu_\varepsilon, \mu)/\varepsilon^2 + \frac{64C}{\varepsilon} (1 + \varepsilon \|f\|_\infty)C(0)H(\mu_\varepsilon, \mu)}}.
\]

Let \( \varepsilon \) tend to 0. In the limit we obtain that for all those \( f \):

\[
\int f^2 \, d\mu \leq 8\sqrt{C(0)} \sqrt{\int |\nabla f|^2 \, d\mu} \sqrt{\int f^2 \, d\mu}
\]

which gives the (S.G.I.) (but with a worse constant) for all those \( f \), and then extend by density.
Let us come to the restricted log-Sobolev inequality. That (4) implies (1) is standard. To prove (2) implies (3), one can get a precise result by using the robust version of the logarithmic Sobolev inequality proved in [8] (formula (2.6)) namely:

$$\int f^2 \log f^2 \, d\mu \leq \frac{t}{\beta} \int |\nabla f|^2 \, d\mu + \frac{2}{\beta} \log \left( \int f^{1+\beta} P_t \, f \, d\mu \right),$$

(2.8)

that holds for any $P_t$ (satisfying the assumptions in the introduction), any nonnegative $\beta$ and any nonnegative $f$ such that $\int f^2 \, d\mu = 1$.

Indeed, for $\beta \leq 1$, $\int f \, d\mu$ and $\int f^{1+\beta} \, d\mu$ are less or equal to 1, hence

$$ \log \left( \int f^{1+\beta} P_t \, f \, d\mu \right) \leq \log \left( 1 + \int f^{1+\beta} P_t \left( f - \int f \, d\mu \right) \, d\mu \right) \leq \int f^{1+\beta} P_t \left( f - \int f \, d\mu \right) \, d\mu \leq \text{Var}_{\mu}^{1/2} (f^{1+\beta}) \text{Var}_{\mu}^{1/2} (P_t f).$$

If Poincaré holds with constant $C_P$, we thus obtain:

$$\int f^2 \log f^2 \, d\mu \leq \frac{t}{\beta} \int |\nabla f|^2 \, d\mu + \frac{2(1+\beta)}{\beta} C_P e^{-t/C_P} \left( \int |\nabla f|^2 \, d\mu \right)^{1/2} \left( \int |\nabla f|^{2\beta} \, d\mu \right)^{1/2},$$

and finally

$$\int f^2 \log f^2 \, d\mu \leq \left( \frac{t}{\beta} + \frac{2(1+\beta)}{\beta} C_P e^{-t/C_P} \| f^{2\beta} \|_{\infty} \right) \int |\nabla f|^2 \, d\mu.$$  

An easy optimization in $t$ shows that the best choice of $\beta$ is $\beta = 1$ and yields, for $h = f^2,$

$$\int h \log h \, d\mu \leq C_P \left( 2 \log 2 + \frac{1}{2} \log \| h \|_{\infty} \right) \int \frac{|\nabla h|^2}{h} \, d\mu. \quad (2.9)$$

Theorem 1.13 gives another characterization of the Spectral Gap property in terms of transportation inequalities. It has to be compared with Corollary 5.1 in [4], where (S.G.I.I) is shown to be equivalent to some $W_L$ transportation inequality (see Section 3).

**Remark 2.10.** One may prove (2) implies (3) in an elementary way using once again a truncation argument and careful calculus but with less precise constants.

**Remark 2.11.** One can easily get a similar but weaker statement without any effort. Indeed recall that $u \log u - u + 1 \geq 0$. Writing this inequality with $v = 1/u$ and then multiplying by $v^2$ yields $v \log v \leq v^2 - v$. Applying this with $v = h(x)$ and integrating with respect to $\mu$ yields:

$$\int h \log h \, d\mu \leq \text{Var}_{\mu} (h),$$

if $h$ is a density of probability.

Hence if Poincaré holds:

$$\int h \log h \, d\mu \leq C_P \int |\nabla h|^2 \, d\mu \leq C_P \| h \|_{\infty} \int \frac{|\nabla h|^2}{h} \, d\mu.$$  

**3. Application to transportation inequalities**

In this section we shall see how to use the functional inequalities of the previous section in order to obtain transportation inequalities (in particular Corollary 1.12). We shall also compare this results with other similar results in the literature.

We start this section by the proof of the elementary Lemma 1.8 showing that the obstruction for $T_2$ to hold is in a neighborhood of $\mu$. Notice that many results in this section are available in a general metric space.
Proof of Lemma 1.8. Introduce the Young function,

$$\tau(u) = u \log^+(u), \quad (3.1)$$

and its Legendre conjugate function $\tau^*(v) = v 1_{v<1} + e^{v-1} 1_{v \geq 1}$.

Among all possible coupling of $(\mu, \nu)$, the simplest one is the independent one, i.e., if we denote $h = \frac{d\nu}{d\mu}$,

$$\pi(dx, dy) = h(x)\mu(dx)\mu(dy).$$

Accordingly

$$W^p_p(\nu, \mu) \leq \int d^p(x, y)h(x)\mu(dx)\mu(dy) \leq 2N_T(h)N_T^*(d^p),$$

where $N_T$ and $N_T^*$ are the gauge norms in the corresponding Orlicz spaces, the second inequality being the classical Hölder–Orlicz inequality (see, e.g., [19] for all concerned with Orlicz spaces). Recall that the gauge norm for a general Young function $\psi$ is defined by:

$$N_\psi(g) = \inf \left\{ \lambda > 0, \int \psi(g/\lambda)(x, y)\mu(dx)\mu(dy) \leq 1 \right\},$$

such that an easy convexity argument yields:

$$N_\psi(g) \leq \max \left\{ 1, \int \psi(g) d\mu \otimes d\mu \right\}. \quad (3.2)$$

In addition remark that

$$\int h \log^+(h) = \int h \log(h) - \int h \log(h) \leq \int h \log(h) + 1/e.$$

Hence if $H(v, \mu) \geq 1$,

$$1 \leq \int h \log^+(h) \leq (1 + 1/e)H(v, \mu),$$

and according to (3.2) and what precedes:

$$W^p_p(\nu, \mu) \leq 2(1 + 1/e)N_T^*(d^p)H(v, \mu).$$

Finally, thanks to $I_\varepsilon(p)$, $N_T^*(d^p) < +\infty$ and the result follows.

One can improve the preceding result by showing that (up to the constant) it holds for $H(v, \mu)$ bounded away from 0. But as quoted in Proposition 1.9 one can also get a precise bound for the behavior of the Wasserstein distances when entropy goes to 0.

Theorem 1.13 suggests that working with bounded density is natural with regard to transportation cost inequalities, starting with Poincaré inequality. We are so tempted to use some truncation for $\nu$, i.e., if $a > 0$ we define:

$$\nu_a = \left(1/\nu(h \leq a)\right)h 1_{h \leq a}\mu, \quad (3.3)$$

and look at what happens. According to Lemma 1.8 and (3.2) we may and will assume that $H(v, \mu)$ is small enough. We start with two elementary lemmata.

Lemma 3.4. Let $\nu = h\mu$ be a probability measure. If $a > e$, then

1. $H(v, \mu) \geq (1 - 1/\log a)\int_{h>a} h \log h \, d\mu$,
2. $v(h > a) \leq (1/(\log a - 1))H(v, \mu)$.

Proof. Again we start with $u \log u + 1 - u \geq 0$ which yields:

$$\int_{h \leq a} h \log h \, d\mu + 1 - \int_{h \leq a} h \, d\mu \geq 0,$$
hence
\[ H(v, \mu) \geq \int_{h > a} h \log h \, d\mu - v(h > a). \]

(2) follows immediately since \( \log h > \log a \) on \( \{h > a\} \). For (1) we have:
\[ v(h > a) \leq \int_{h > a} \frac{\log h}{\log a} \, h \, d\mu = (1/\log a) \int_{h > a} h \log h \, d\mu. \]

Lemma 3.5. Let \( v = h\mu \) be a probability measure such that \( H(v, \mu) \leq 1/2 \). If \( a > e^{3/2} \) and \( v_a \) is given by (3.3), then
\[ H(v_a, \mu) \leq \left(1 + \frac{1}{2(\log a - 3/2)} + \frac{2}{\log a - 1}\right) H(v, \mu). \]

Proof.
\[
H(v_a, \mu) = \int h 1_{h \leq a} \frac{h}{v(h \leq a)} \log \left( \frac{h}{v(h \leq a)} \right) \, d\mu \\
\leq H(v, \mu) + ((1/v(h \leq a)) - 1) \int_h h \log h \, d\mu - \log(v(h \leq a)) - \int_h h \log h \, d\mu \\
\leq H(v, \mu) + \frac{v(h > a)}{v(h \leq a)} H(v, \mu) - \log(1 - v(h > a)).
\]

But if \( 0 \leq x \leq 1/2, \, -\log(1 - x) \leq 2x \), hence according to Lemma 3.4(2), if \( H(v, \mu) \leq 1/2, \, -\log(1 - v(h > a)) \leq (2/(\log a - 1)) H(v, \mu) \), and
\[
\frac{v(h > a)}{v(h \leq a)} \leq \frac{H(v, \mu)}{\log a - 1 - H(v, \mu)}
\]
and we get the desired result. \( \Box \)

We turn now to the proof of Corollary 1.12.

Proof of Corollary 1.12. We assume that \( I(\alpha) \) and the exponential integrability condition \( (EI_\varepsilon(2)) \) are satisfied. Let consider a positive constant \( C \) that may change line to line, but which does not depend neither on \( a \) nor \( \alpha \).

Since \( W_2 \) is a distance, it holds:
\[ W_2(v, \mu) \leq W_2(v_a, \mu) + W_2(v_a, v). \]

But, on one hand, since \( I(\alpha) \) holds, according to Theorem 1.11, for \( a \) large enough,
\[
W_2^2(v_a, \mu) \leq C \log^{1-\alpha}(a/v(h \leq a)) H(v_a, \mu) \\
\leq 2C \log^{1-\alpha}(a/v(h \leq a)) H(v, \mu),
\]
provided \( H(v, \mu) \leq 1/2 \) thanks to Lemma 3.5.

On the other hand, a classical result in mass transportation theory (see [22, Proposition 7.10]) tells that for any \( x_0 \),
\[ W_2^2(v', v) \leq 2 \int d^2(x, x_0)|h' - h| \, d\mu. \]

Applying (3.7) with \( v' = v_a \) yields, assuming again that \( H(v, \mu) \leq 1/2, \)
\[
W_2^2(v_a, v) \leq 2 \frac{v(h > a)}{v(h \leq a)} \int_{h \leq a} d^2(x, x_0) h \, d\mu + 2 \int_{h > a} d^2(x, x_0) \, d\nu \leq C(H(v, \mu) + N_\tau(h 1_{h > a})),
\]
according to Lemma 3.4(2) and Orlicz–Hölder inequality (see the proof of Lemma 1.8 at the beginning of the section), since \( EI_\varepsilon(2) \) is satisfied for the latter.
Plugging all this into (3.6), we get that there exists a constant $C$ such that
\[
W_2^2(v, \mu) \leq C(\log^{1-a}(a)H(v, \mu) + N_\tau(h1_{h>a})).
\] (3.8)

Now we choose $a = 1/H^q(\mu, v)$ for some $q > 0$ (recall that we may assume that $H(v, \mu)$ is small enough). Lemma 3.4(2) furnishes:
\[
v(h > a) \leq C H(\nu, \mu) \log \left(\frac{1}{H(\nu, \mu)}\right),
\]
so that it is easily seen that
\[
N_\tau(h1_{h>a}) = \inf \left\{ \lambda > 0; \int \tau(h1_{h>a}) d\mu \right\} \leq \int h1_{h>a} \log^+ h d\mu \leq C q^{-1} H(\nu, \mu).
\]
We have thus obtained
\[
W_2^2(\nu, \mu) \leq C \left( (q \log(1/H(\nu, \mu)))^{1-a} + 1/q \right) H(\nu, \mu),
\]
so that optimizing in $q (q^{2-a} = 1/(1-a) \log^{1-a}(1/H))$ so that $a$ is big for small entropy $H$) we complete the proof of Corollary 1.12. \qed

**Remark 3.9.** It is worthwhile noticing that $N_\tau(h1_{h>a})$ behaves like
\[
H(\nu, \mu) \frac{\log(1/H(\nu, \mu))}{\log a}
\]
for small entropies. This is why some extra logarithm appears in Corollary 1.12.

Also notice that a similar result can be directly obtained using the $\mathcal{W}_L$ transportation inequality in [4], when $\alpha = 0$, i.e., when Poincaré holds. We briefly indicate how to do below (some of the bounds are clearly nonsharp).

Indeed taking an optimal coupling $\Pi$ (or an almost optimal coupling, and then taking limits) for the $L$ cost introduced in [4, Section 5.2] (recall that $L$ is, up to constants, the square of the distance for small distances and the distance for large ones), it is immediate that
\[
W_2^2(v, \mu) \leq C \sqrt{\log(1/H)} \int 1_{d^2(x, x_0) \leq q \log(1/H)} 1_{d^2(y, x_0) \leq q \log(1/H)} L(x, y) d\Pi
\]
\[
+ \int 1_{d^2(x, x_0) \geq q \log(1/H)} 1_{d^2(y, x_0) \geq q \log(1/H)} d^2(x, y) d\Pi
\]
\[
\leq C \sqrt{\log(1/H)} W_L + 2 \int 1_{d^2(x, x_0) \geq q \log(1/H)} (d^2(x, x_0) + C')(d\nu + d\mu),
\]
where $C' = \int d^2(x, x_0)(d\mu + d\nu)$. Hence if Poincaré holds, according to [4] the first term in the right-hand side is less than $CH \sqrt{\log(1/H)}$. For the second term first write:
\[
\int_{e^{d^2(x, x_0)} \geq 1/H^q} d^2(x, x_0) d\mu \leq \left( \int d^4 d\mu \right)^{1/2} \left( \mu(d^2 \geq q \log(1/H)) \right)^{1/2} \leq C \left( \int d^4 d\mu \right)^{1/2} H^{q/2},
\]
where $\eta$ is the constant of the Gaussian concentration of $\mu$, namely for which $(EI_\eta(2))$ holds. It is now enough to choose $q$ large enough for this term to be less than $CH$. It remains to study
\[
\int_{e^{d^2(x, x_0)} \geq 1/H^q} d^2(x, x_0) d\nu.
\]
First remark that we can also assume that $h > K$, for $K$ large enough, in this integral, since on the complement set $h \leq K$ we may use the previous inequality.

But according to Young’s inequality,
\[
d^2(x, x_0)h(x) \leq (1/e)\left( h(x) \log(h(x)) + e^{c d^2(x, x_0)} \right),
\]
so that the previous quantity can be controlled by

$$\int_{h > K} h \log h \, d\mu,$$

and

$$\int_{e^{d^2(x,x_0)} > 1/Hq} e^{d^2(x,x_0)} \, d\mu.$$  

Choosing $\varepsilon$ small enough, and using $(EI\varepsilon(2))$ we may argue as before (replacing $d^4$ by $e^{2d^2}$) to bound the latter term by (constant times) $H$ again, while for the former we can use Lemma 3.4(1).

This remark shows that the optimal coupling (if it exists) for $W_L$ achieves (up to the constants) the bound obtained in Corollary 1.12. One can thus think that this bound is not optimal.

4. Towards a criterion for Talagrand inequality

For simplicity from now on we assume that $E = \mathbb{R}^n$.

First recall (2.1):

$$A = \frac{d}{dt} (-W_2(\nu_t, \mu)) \leq \limsup_{s \to 0^+} \frac{1}{s} W_2(\nu_t + s \nu, \nu_t + s \mu) \leq 2 \left( \int |\nabla \sqrt{P_t h}|^2 \, d\mu \right)^{1/2} = \left( -\frac{d}{dt} H(\nu_t, \mu) \right)^{1/2}.$$  

Hence using the restricted log-Sobolev inequality given by Theorem 1.13(3) for bounded functions we should directly recover the modified transportation inequality in Theorem 1.11 for $\alpha = 0$ (up to the constants). More generally the proof of Theorem 1.11 works for any $P_t$ stable subset of functions for which a (restricted) logarithmic Sobolev inequality holds.

Since truncating by constants is not sufficient (in view of the preceding section), we shall first explain what kind of truncation is useful.

4.1. A first reduction

**Lemma 4.1.** If,

$$W_2 \left( \frac{v + \mu}{2}, \mu \right) \leq \sqrt{C} \left( \frac{v + \mu}{2}, \mu \right),$$

then

$$W_2(v, \mu) \leq \frac{\sqrt{C}}{\sqrt{2} - 1} \sqrt{H(v, \mu)}.$$  

**Proof.** Since $W_2$ is a distance, then

$$W_2(v, \mu) \leq W_2 \left( \frac{v + \mu}{2}, \mu \right) + W_2 \left( \frac{v + \mu}{2}, v \right).$$

But $W_2^2$ is convex in each argument, hence

$$W_2 \left( \frac{v + \mu}{2}, v \right) \leq \sqrt{1/2} W_2(v, \mu),$$

and

$$W_2(v, \mu) \leq \frac{\sqrt{2}}{\sqrt{2} - 1} W_2 \left( \frac{v + \mu}{2}, \mu \right).$$
In addition since relative entropy is also convex, \( H(\frac{\nu + \mu}{2}, \mu) \leq \frac{1}{2} H(\nu, \mu) \) and we get the result. \( \square \)

The meaning of this lemma is clear: it is enough to show \( T_2 \) for densities \( h \) such that

\[
\text{for all } x, \quad h(x) \geq \frac{1}{2}. \tag{4.2}
\]

Note that this set is \( P_t \) stable.

More useful is the following:

**Lemma 4.3.** Let \( \gamma_\eta(x) = e^{\eta d^2(x, x_0)} \) for some nonnegative \( \eta \), and define:

\[
dv^K_{\gamma_\eta} = h_{K \gamma_\eta} d\mu = z^{-1} \min(h, K \gamma_\eta) d\mu,
\]

where \( z \) is a normalizing constant, assuming that \( K \geq e^2 \).

Assume that \( (EI_\varepsilon(2)) \) is satisfied.

If for some \( \eta < \varepsilon \),

\[
W_2^2(\nu_{K \gamma_\eta}, \mu) \leq \sqrt{CH(\nu_{K \gamma_\eta}, \mu)},
\]

then provided \( H(\nu, \mu) \leq 1/2 \) it holds:

\[
W_2(\nu, \mu) \leq \sqrt{C(\eta, K, C) H(\nu, \mu)}.
\]

**Proof.** Note that \( 1 \geq z \geq \nu(h \leq K) \). So on one hand,

\[
H(\nu_{K \gamma_\eta}, \mu) = z^{-1} H(\nu, \mu) - z^{-1} \log(z) - z^{-1} \int_{h \geq K \gamma_\eta} (h - K \gamma_\eta) \log(h/z) d\mu + z^{-1} \int_{h \geq K \gamma_\eta} K \gamma_\eta \log(K \gamma_\eta/h) d\mu
\]

\[
\leq z^{-1} H(\nu, \mu) - z^{-1} \log(z) \leq C \left( 1 + \frac{2}{\log K - 1} \right) H(\nu, \mu), \tag{4.4}
\]

as soon as \( H(\nu, \mu) \leq 1/2 \), hence

\[
H(\nu_{K \gamma_\eta}, \mu) \leq C(K) H(\nu, \mu).
\]

On the other hand, (3.7) with \( \nu' = \nu^K_{\gamma_\eta} \) furnishes,

\[
W_2^2(\nu^K_{\gamma_\eta}, \nu) \leq 2 \int d^2(x, x_0)|h - h_{K \gamma_\eta}| d\mu, \tag{4.5}
\]

and thus we get:

\[
W_2^2(\nu_{K \gamma_\eta}, \nu) \leq 2(z^{-1} - 1) \int_{h \leq K \gamma_\eta} d^2 h d\mu + 2 \int_{h > K \gamma_\eta} d^2 \left| 1 - z^{-1} \frac{K \gamma_\eta}{h} \right| h d\mu
\]

\[
\leq \frac{2}{\nu(h \leq K)} \int d^2 h d\mu + \frac{4}{\nu(h \leq K)} \int_{h > K \gamma_\eta} \frac{1}{\eta} \log(h/K) h d\mu
\]

\[
\leq M(K, \eta) H(\nu, \mu), \tag{4.6}
\]

where we used Lemma 3.4 and the smallness of \( H(\nu, \mu) \) (in particular \( N_t(h) \) is bounded, hence \( \int d^2 h d\mu \) is bounded by some constant only depending of \( \eta \)).

Putting all this together, we thus have shown for \( H(\nu, \mu) \leq 1/2 \),

\[
W_2(\nu, \mu) \leq W_2(\nu_{K \gamma_\eta}, \nu) + W_2(\nu_{K \gamma_\eta}, \mu) \leq \sqrt{C(\eta, K, C) H(\nu, \mu)}. \quad \square
\]

Once again this lemma shows that we may assume that \( h \leq K \gamma_\eta \) for some \( \eta < \varepsilon \). But unfortunately this set does not seem to be \( P_t \) stable in general.
However it is included into the subset of densities \( h \) such that \( \int h^q \, d\mu \leq M \) for some \( q > 1 \) and \( M < +\infty \). Hence as a consequence we obtain Corollary 1.15.

Unfortunately we do not know whether the restricted log-Sobolev inequality stated in Corollary 1.15 is strictly weaker than the (full) log-Sobolev inequality or not. Notice that the method used in Remark 2.10 cannot be extended to the case \( q < +\infty \). Indeed in this case one has to call upon Hölder inequality, thus introduce some power of the Dirichlet form.

### 4.2. Decay of entropy

In Section 2 we bound the right-hand side in (2.1) by some derivative, using Poincaré inequality, and then we integrated in time. One can also first integrate in time and then use inequalities. The result presented here means that this methodology is promised to failure, at least in the bounded curvature case.

Choose some weight function \( \xi \) such that \( \int_0^{+\infty} \xi^{-1}(t) \, dt = 1 \). Integrating (2.1) with respect to time yields (recall that \( W_2(\nu_t, \mu) \) goes to 0 as \( t \) tends to infinity):

\[
W_2(\nu, \mu) \leq \frac{1}{2} \int_0^{+\infty} \left( -\frac{d}{dt} H(\nu_t, \mu) \right)^{1/2} \, dt = \frac{1}{2} \int_0^{+\infty} \left( -\xi^2(t) \frac{d}{dt} H(\nu_t, \mu) \right)^{1/2} \xi^{-1}(t) \, dt \\
\leq \left( \int_0^{+\infty} -\xi(t) \frac{d}{dt} H(\nu_t, \mu) \, dt \right)^{1/2},
\]

where we have used Cauchy–Schwarz inequality for the probability measure \( \xi^{-1}(t) \, dt \) to get the latter inequality. Hence provided

\[ \xi(t) H(\nu_t, \mu) \] goes to 0 as \( t \) goes to \( +\infty \),

we have obtained

\[
W_2^2(\nu, \mu) \leq \xi(0) H(\nu, \mu) + \int_0^{+\infty} \xi'(t) H(\nu_t, \mu) \, dt,
\]

where the right-hand side is finite provided the relative entropy goes to 0 quickly enough.

**Remark 4.9.** Note that if we choose \( \xi^{-1}(t) = (1/T)1_{t \leq T} \), the derivation in (4.7) furnishes:

\[
W_2(\nu, \mu) \leq T^{1/2} \left( H(\nu, \mu) - H(\nu_T, \mu) \right)^{1/2} + W_2(\nu_T, \mu).
\]

Hence a uniform decay of the Wasserstein distance implies \( T_2 \).

But this result also shows that \( T_2 \) holds as soon as it holds for the probability densities of the form \( h = P_T g \) for some \( T > 0 \).

The natural question is thus to know whether one can find other uniform decays than the exponential one for relative entropy. Of course the exponential decay of the relative entropy,

\[ H(\nu_t, \mu) \leq e^{-Ct} H(\nu, \mu), \]

is known to be equivalent to a logarithmic Sobolev inequality, in which case it is enough to take \( \xi(t) = e^{\theta t} \) with \( \theta \) smaller than the inverse of the log-Sobolev constant.

More generally if for some \( s > 0 \) and some \( \lambda > 0 \),

\[ H(\nu_t, \mu) \leq e^{-\lambda t} H(\nu, \mu), \]

for all \( \nu \), using the semi-group property and the fact that \( t \rightarrow H(\nu_t, \mu) \) is nonincreasing, it is easy to see that

\[ H(\nu_t, \mu) \leq e^{-\lambda (t-1)} H(\nu, \mu) \]
for all $t > s$, hence the relative entropy is exponentially decaying, but there is some constant $\epsilon^\lambda$ in front of the $e^{-Ct}$. This kind of exponential decay is no more immediately equivalent to a log-Sobolev inequality, and we may ask whether it is strictly weaker or not.

Unfortunately the following lemma shows that in many cases both are equivalent

**Lemma 4.10.** Assume that the potential $V$ satisfies the curvature condition $\text{Hess}(V) \geq R \text{Id}$ for some $R \in \mathbb{R}$. Then if for some $s > 0$ and some $\lambda > 0$,

$$H(v_s, \mu) \leq e^{-\lambda} H(v, \mu),$$

for all $v, \mu$ satisfies a logarithmic Sobolev inequality.

**Proof.** Recall the classical commutation properties (see [1, Theorem 5.4.7 and Remark 5.4.8])

$$P_t(h \log h) - P_t h \log(P_t h) \leq \frac{1 - e^{-Rt}}{2R} P_t \left( \left| \nabla h \right|^2 h \right), \quad (4.11)$$

(just being careful since the semi-group therein is our $P_{2t}$) the constants being $t/2$ for $R = 0$. Integrating the right-hand inequality with respect to $\mu$ we obtain for $h$ a density of probability and $v = h\mu$ (see also [4, Eq. (4.4)],

$$H(v, \mu) \leq H(v_t, \mu) + \frac{1 - e^{-Rt}}{2R} \int \frac{|\nabla h|^2}{h} \, d\mu.$$

Applying this at time $s$ we finally obtain:

$$H(v, \mu) \leq \frac{1 - e^{-Rs}}{2R(1 - e^{-\lambda})} \int \frac{|\nabla h|^2}{h} \, d\mu. \quad \square$$

A similar result is true for the Spectral Gap Inequality, without any restriction according to the well known robust inequality,

$$\text{Var}_\mu(g) \leq t \int |\nabla g|^2 \, d\mu + \text{Var}_\mu(P_t g).$$

**4.3. A natural $P_t$ almost stable subset**

The preceding subsection has shown that there is no hope to get some uniform decay of relative entropy without log-Sobolev. Hence we really have to find an appropriate $P_t$ stable subset for which a restricted log-Sobolev inequality is available. As we saw in Section 4.1 the natural one is the subset of densities smaller than constant times some Gaussian density, but these sets do not seem to be $P_t$ stable. Fortunately we can combine the ideas of both previous subsections in order to build an appropriate almost $P_t$ stable subset. The result is the following:

**Lemma 4.12.** Assume that the potential $V$ satisfies the curvature condition $\text{Hess}(V) \geq R \text{Id}$ for some $R \leq 0$ and that $EI_{\varepsilon}(2)$ is satisfied. Then if $h \leq K\gamma_\eta$ for $\eta < \varepsilon/2$, $P_t h \leq M(K, R)\gamma_\beta$ with $\beta = (2\eta R)/(\varepsilon(e^{Rt} - 1))$.

In particular, for any $\theta > 0$, there exist $T > 0$ and $\eta > 0$ such that for all $t \geq T$, $P_t h \leq M\gamma_\theta$.

**Proof.** Recall the beautiful Harnack–Wang inequality (see [24] and [25, (2.1)]),

$$|P_t h(x)| \leq (P_t(|h|^q))^{1/q}(y) \exp \left( \frac{Rd^2(x, y)}{2(q - 1)(e^{Rt} - 1)} \right), \quad (4.13)$$

that holds for any $(x, y)$, any $q > 1$ and any continuous and bounded $h$. Again we shall integrate with respect to $\mu(dy)$, use the elementary $d^2(x, y) \leq 2d^2(x, x_0) + 2d^2(x_0, y)$ and apply Cauchy–Schwarz in order to get:

$$|P_t h(x)| \leq M \exp \left( \frac{Rd^2(x_0, x)}{(q - 1)(e^{Rt} - 1)} \right), \quad (4.14)$$
with
\[ M = \left( \int (P_t(|h|^q))^{2/q} (y) \mu(dy) \right)^{1/2} \left( \int \exp\left( \frac{2Rd^2(x_0, y)}{(q-1)(e^{Rt} - 1)} \right) \mu(dy) \right)^{1/2}. \]

Inequality (4.14) is interesting provided \( M \) is finite, i.e., provided:
\[ q > 2, \quad \frac{2R}{(q-1)(e^{Rt} - 1)} \leq \varepsilon \quad \text{and} \quad h \in L^q, \tag{4.15} \]

since \( u^{2/q} \leq 1 + u \) for a nonnegative \( u \) if \( q > 2 \). Note that, in this case,
\[ \int (P_t(|h|^q))^{2/q} (y) \mu(dy) \leq K^2 \left( 1 + \int |h/K|^q \mu \right) \leq K^2 \left( 1 + \int e^{\eta q d^2} \mu \right), \]
so that for \( \eta < \varepsilon \) the latter does not depend on \( \eta \) but only depends on \( R \) and \( \varepsilon \).

Hence if \( h \leq K \gamma \eta \) for \( \eta < \varepsilon/2 \), we may take \( q = \varepsilon/\eta \) and obtain that \( P_t h \leq M(K, R) \gamma \beta \) with \( \beta = (2\eta R)/(\varepsilon(e^{Rt} - 1)). \)

How we shall use this result is now clear. According to Lemma 4.3 we may assume that \( h \leq K \gamma \eta \) with \( \eta \) as small as we want, in order to ensure that \( P_t h \leq M \gamma \theta \) for \( \theta \) small enough, and all \( t \) large enough. But thanks to Remark 4.9 it is enough to consider such densities and the required \( P_t \) stability is now ensured. So we have shown Corollary 1.16 at least with the additional curvature assumption.

It should be very interesting to know whether the statement of Lemma 4.12 is still true without the curvature assumption or not. This would complete the picture of what can be done using Otto and Villani coupling.

4.4. The infimum convolution approach

Let us as an introduction of this method present a refinement of Lemma 4.3. Actually one can obtain a more precise result if instead of Villani’s coupling used in (3.7) one uses the inf-convolution method in [4].

Indeed recall that
\[ W^2_2(v, \mu) = \sup \left( \int g \, dv - \int f \, d\mu \right), \]

where the supremum is running over all pairs \((f, g)\) of measurable and bounded functions satisfying \( g(x) \leq f(y) + d^2(x, y) \) for all \((x, y)\). Adding a constant to both \( f \) and \( g \) we may assume that \( \int f \, d\mu = 0 \). Denote by:
\[ Qf(x) = \inf_{y \in E} (f(y) + d^2(x, y)), \]

the function achieving the optimal choice. Integrating with respect to \( \mu \) it holds:
\[ Qf(x) \leq \int d^2(x, y) \mu(dy) \leq 2d^2(x, x_0) + 2 \int d^2(y, x_0) \mu(dy), \]
i.e., \( Qf(x) \leq 2d^2(x, x_0) + C(x_0) \).

Recall that by (see [4]), the condition \( \int e^{\eta Qf} \, d\mu \leq 1 \) is equivalent to \( T_2 \). Remark now that for \( 2\eta < \varepsilon \) the density \( h = e^{\eta Qf} / \int e^{\eta Qf} \, d\mu \) is such that either \( \int e^{\eta Qf} \, d\mu \geq 1 \) and \( h \leq e^{\eta C(x_0)} \gamma_2 \eta \) or \( \int e^{\eta Qf} \, d\mu \leq 1 \). We may thus focus on the first condition.

If \( \beta W^2_2(v, \mu) \leq H(v, \mu) \) for all \( v \) such that \( dv/d\mu \leq K \gamma \theta \), then for \( 2\eta < \theta \) and such that \( e^{\eta C(x_0)} \leq K \), \( v = h \mu \) satisfies the previous condition so that
\[ \beta \int Qf \, dv \leq H(v, \mu). \]

If in addition \( \eta < \beta \) we may replace \( \beta \) by \( \eta \) in the left-hand side of the previous inequality (even if this left-hand side is nonpositive, since the right-hand side is nonnegative), and thus obtain:
\[ \left( \int e^{\eta Qf} \, d\mu \right) \log \left( \int e^{\eta Qf} \, d\mu \right) \leq 0, \]
i.e., \( \int e^{Qf} \, d\mu \leq 1 \) for all \( f \), which leads then to a refined version of Lemma 4.3.

We may now go further in this infimum convolution approach and mimic arguments of [4, Section 3.3]. This approach based on Herbst argument is apparently not well suited for restricted logarithmic Sobolev inequalities. Indeed it requires the use of nonnormalized functions for which the hypothesis in Corollary 1.16 reads as

\[
f^2 \leq \left( \int f^2 \, d\mu \right) Ke^{\tilde{d}^2(x,x_0)}.
\]

However, and surprisingly enough, a very slight improvement of the argument yields the result. Before starting the proof, remark that we have used a slightly different form for the definition of the in-convolution than in [4], but all calculus presented in there work is only modified by constants.

**Proof of Theorem 1.17.** In fact the theorem will be established under some more general hypothesis, namely we do not need that the restricted logarithmic Sobolev inequality be verified for all the functions satisfying (1.18) but only a subclass. It is however more convenient to write the theorem with this larger class and easier to derive conditions on the real line for such a restricted logarithmic Sobolev inequality.

Remember that \( Qf(x) \leq 2d^2(x,x_0) + C(x_0) \) for all \( f \) such that \( \int f \, d\mu = 0 \). For \( 2\eta \leq \varepsilon \) and all \( \lambda \) introduce \( f_\lambda^2 = e^{Q(\lambda f)} \) and \( G(\lambda) = \int f_\lambda^2 \, d\mu \). Then either \( G(\lambda) \leq 1 \) or \( G(\lambda) > 1 \) and in this case \( f_\lambda^2 \) satisfies (4.16) (for some well chosen \( \tilde{\eta} \) depending on \( \eta \) and \( \lambda \)).

Assume that \( G(1) > 1 \) and introduce:

\[
\lambda_0 = \inf\{ \lambda \in [0, 1], \ G(u) > 1 \ \text{for all}\ u \geq \lambda \}.
\]

Then \( \lambda_0 < 1, G(\lambda_0) = 1 \) (remark that \( G(0) = 1 \) and \( G(\lambda) > 1 \) on \( ]\lambda_0, 1[ \)). Hence the restricted log-Sobolev holds for all \( \lambda \in ]\lambda_0, 1[ \). An easy computation using the Hamilton–Jacobi semigroup described in [4] (see Section 2.1 formula (2.6) and (3.3) first formula p. 380) yields:

\[
\lambda G'(\lambda) = \int f_\lambda^2 \log f_\lambda^2 \, d\mu - \frac{4}{\eta} \int |\nabla f_\lambda|^2 \, d\mu.
\]

We may always assume that \( C\eta \leq 4 \) (decreasing \( \eta \) if necessary), so that the latter yields:

\[
\lambda G'(\lambda) \leq G(\lambda) \log G(\lambda),
\]

(4.17) on \( ]\lambda_0, 1[ \). This differential inequality can be rewritten:

\[
\frac{d}{d\lambda} \left( \frac{\log G(\lambda)}{\lambda} \right) \leq 0,
\]

so that \( \log G(\lambda)/\lambda \) is nonincreasing, hence

\[
\lambda_0 \log G(1) \leq \log G(\lambda_0) = 0.
\]

If \( \lambda_0 > 0 \), we get \( G(1) \leq 1 \) in contradiction with our assumption \( G(1) > 1 \). If \( \lambda_0 = 0 \), \( \lim_{\lambda \to 0} \frac{\log G(\lambda)}{\lambda} = \frac{G'(0)}{G(0)} = \eta \int f \, d\mu = 0 \) and the same conclusion holds.

Hence \( G(1) \leq 1 \) for all \( f \) as above, which is known to be equivalent to \( T_2 \). \( \square \)

5. The case of the real line

As for many functional inequalities the one-dimensional case is much simpler thanks to Hardy inequalities. We thus consider a probability measure \( \mu \) on the real line such that \( \int e^{x^2} \, d\mu < +\infty \), and denote by \( v \) the second moment \( v = \int x^2 \, d\mu \). We also denote by \( M \) the quantity \( M = e^{2v} \). In the sequel \( \eta \) will be a positive number smaller than \( 1 \wedge \varepsilon/2 \) so that \( e^{2\eta v} \leq M \). Recall that \( \mu(dx) = e^{-V(x)} \, dx \).

Let \( h \) be such that \( \int h \, d\mu = 1 \) and \( h \leq Me^{\eta x^2} \). Define \( v = h\mu \). For \( K \) large enough to be chosen later, we get:
\[
\int h \log h \, d\mu = \int_{h \leq K} h \log h \, d\mu + \int_{h > K} h \log h \, d\mu \\
\leq \int (h \land K) \log(h \land K) \, d\mu + \int_{h > K} h (\log M + \eta x^2) \, d\mu \\
\leq \int (h \land K) \log(h \land K) \, d\mu + \log M v(h > K) + \int \psi^2(\sqrt{h}) x^2 \, d\mu,
\]
(5.1)

where \(\psi(u) = 0\) if \(0 \leq u \leq \sqrt{K/2}\), \(\psi(u) = (\sqrt{2}/(\sqrt{2} - 1))(u - \sqrt{K/2})\) if \(\sqrt{K/2} \leq u \leq \sqrt{K}\), and \(\psi(u) = u\) if \(u \geq \sqrt{K}\). If we choose \(K > 2M\) then \(\psi(\sqrt{h})(0) = 0\).

We shall now bound the three terms in the right-hand side of (5.1). To this end we first observe that

\[
H(\nu, \mu) \leq \int M e^{\eta x^2} (\log M + \eta x^2) \, d\mu = C(\eta, \mu) < +\infty.
\]

Hence, according to Lemma 3.4(2), as soon as \(K > e\),

\[
v(h > K) \leq \frac{1}{\log K - 1} H(v, \mu) \leq \frac{1}{\log K - 1} C(\eta, \mu).
\]

It follows that \(\log M v(h > K) \leq 1/2 H(v, \mu)\) as soon as \(2 \log M \leq \log K - 1\). Furthermore

\[
1 \geq z_K = \int (h \land K) \, d\mu \geq 1 - \frac{C(\eta, M)}{\log K - 1},
\]

so that for \(\log K - 1 \geq 2C(\eta, M), z_K \geq 1/2\). Thus

\[
\int (h \land K) \log(h \land K) \, d\mu \leq \int (h \land K) \log\left(\frac{h \land K}{z_K}\right) \, d\mu,
\]

and \(h \land K \leq 2Kz_K\). Hence if Poincaré holds, we may apply (2.9) and get

\[
\int (h \land K) \log(h \land K) \, d\mu \leq C_P \left(2 \log 2 + (1/2) \log(2K)\right) \int \frac{(h')^2}{h} \, d\mu.
\]
(5.2)

Plugging these two estimates into (5.1) we arrive at

\[
(1/2) H(v, \mu) \leq C(C_P, K, \mu) \int \frac{(h')^2}{h} \, d\mu + n \int \psi^2(\sqrt{h}) x^2 \, d\mu.
\]
(5.3)

In order to obtain the desired restricted logarithmic Sobolev inequality, it remains to bound the second term in the right-hand side of (5.3). To this end we shall use Hardy’s inequality on the positive and on the negative half line. We only write things on the positive half line. Since \(\psi(\sqrt{h})(0) = 0\), Hardy’s inequality (see, e.g., [1, Theorem 6.2.1]) gives:

\[
\int_0^\infty \psi^2(\sqrt{h}) x^2 \, d\mu \leq A^+ \int_0^\infty (\psi')^2(\sqrt{h}) \frac{(h')^2}{h} \, d\mu,
\]
(5.4)

where

\[
A^+ = \sup_{x \geq 0} \left(\int_0^\infty t^2 e^{-V(t)} \, dt \int_0^x e^{V(t)} \, dt\right).
\]

Since \(\psi'\) is bounded we have obtained:

**Proposition 5.5.** Let \(E = \mathbb{R}\), \(d\mu = e^{-V} \, dx\). Assume that \((E, I_\varepsilon(2))\) is satisfied. Then the restricted logarithmic Sobolev inequality in Theorem 1.17 holds as soon as

\[
A^+ = \sup_{x \geq 0} \left(\int_0^\infty t^2 e^{-V(t)} \, dt \int_0^x e^{V(t)} \, dt\right)
\]
and

\[
A^- = \sup_{x \leq 0} \left( \int_{-\infty}^{x} t^2 e^{-V(t)} \, dt \int_{x}^{0} e^{V(t)} \, dt \right)
\]

are finite. Hence in this case \( \mu \) satisfies \( T_2 \).

Note first that the boundedness of \( A^+ \) and \( A^- \) are sufficient for the Poincaré inequality to hold (see, e.g., [1, Chapter 6], also see [9] for \( d \)-dimensional general results).

It remains to find sufficient conditions for all these hypotheses to hold. Here we shall follow Section 6.4 in [1] to describe some understandable sufficient conditions. To this end we shall assume that

\[
\lim_{x \to \infty} V'(x) > 0 \quad \text{and} \quad V''(x)/(V'(x))^2 \to 0 \quad \text{when} \ x \ \text{goes to} \ \infty.
\]  

(5.6)

Note that \( V - 2 \log |x| \) also fulfills (5.6). In this case it is known that \( \mu \) satisfies Poincaré inequality (see Theorem 6.4.3(1) in [1]). Furthermore \( A^+ \) is finite as soon as

\[
\lim_{x \to \infty} x^2/(V'(x))^2 < +\infty,
\]  

(5.7)

thanks to the estimates in Corollaire 6.4.2 in [1].

Note in addition that a logarithmic Sobolev inequality holds if in only if we have (in addition to (5.6)),

\[
\lim_{x \to \infty} V(x)/(V'(x))^2 < +\infty.
\]  

(5.8)

In particular if \( V(x) \leq \alpha x^2 \) at infinity, (5.7) implies (5.8). According to Wang’s argument, under the curvature assumption, if \( V(x) \geq \alpha x^h \) at infinity for some \( p > 2 \) then \( \mu \) satisfies a log-Sobolev inequality too. Hence it is not easy with our rough estimate in Proposition 5.5 to build an example of measure with bounded below curvature, satisfying \( T_2 \) but not the log-Sobolev inequality. If we relax the curvature assumption then the construction is simpler.

**Example 5.9.** We only describe the behavior of \( V \) near \( +\infty \). Thus choose:

\[ V(x) = x^3 + 3x^2 \sin^2 x + x^\beta, \]

then

\[ V'(x) = 3x^2(1 + 2 \sin 2x) + 6x \sin^2 x + \beta x^{\beta - 1}, \]

and

\[ V''(x) = 6x^2 \cos 2x + 6x(1 + 2 \sin 2x) + 6 \sin^2 x + \beta(\beta - 1)x^{\beta - 2}, \]

so that (5.6) and (5.7) are satisfied as soon as \( \beta > 2 \), but (5.8) is not satisfied if \( \beta < 5/2 \).

This furnishes an example of a measure satisfying \( T_2 \) but not the log-Sobolev inequality.

**Acknowledgements**

We wish to thank François Bolley and Cedric Villani for numerous and fruitful exchanges. Franck Barthe, Ivan Gentil, Michel Ledoux and Liming Wu are also gratefully acknowledged for their kind interest in this work and their remarks. We want to particularly acknowledge Michel Ledoux for a careful reading of our first draft, and for pointing out some mistakes therein. Also a special thank for the final comments of Franck Barthe.

**References**
