J. Math. Pures Appl. 80, 7 (2001) 669–696

S0021-7824(01)01208-9/FLA

HYPERCONTRACTIVITY OF HAMILTON–JACOBI EQUATIONS

Sergey G. BOBKOV [a], **Ivan GENTIL** [b], **Michel LEDOUX** [b]

[a] *School of Mathematics, University of Minnesota, Minneapolis, MN 55455, USA*

[b] *Département de Mathématiques, Laboratoire de Statistique et Probabilités associé au C.N.R.S., Université Paul-Sabatier, 31062 Toulouse, France*

Manuscript received 10 June 2000, revised 27 January 2001

ABSTRACT. – Following the equivalence between logarithmic Sobolev inequalities and hypercontractivity showed by L. Gross, we prove that logarithmic Sobolev inequalities are related similarly to hypercontractivity of solutions of Hamilton–Jacobi equations. By the infimum-convolution description of the Hamilton–Jacobi solutions, this approach provides a clear view of the connection between logarithmic Sobolev inequalities and transportation cost inequalities investigated recently by F. Otto and C. Villani. In particular, we recover in this way transportation from Brunn–Minkowski inequalities and for the exponential measure. © 2001 Éditions scientifiques et médicales Elsevier SAS

Keywords: Hypercontractivity, Logarithmic Sobolev inequality, Hamilton–Jacobi equation, Infimum-convolution, Brunn–Minkowski inequality

1. Introduction

The fundamental work by L. Gross [17] put forward the equivalence between logarithmic Sobolev inequalities and hypercontractivity of the associated heat semigroup. Let us consider for example a probability measure μ on the Borel sets of $\mathbb{R}^n$ satisfying the logarithmic Sobolev inequality:

$$\rho\,\mathrm{Ent}_\mu\big(f^2\big) \leqslant 2\int |\nabla f|^2\,\mathrm{d}\mu \tag{1.1}$$

for some $\rho > 0$ and all smooth enough functions f on $\mathbb{R}^n$ where:

$$\mathrm{Ent}_\mu\big(f^2\big) = \int f^2 \log f^2\,\mathrm{d}\mu - \int f^2\,\mathrm{d}\mu \log \int f^2\,\mathrm{d}\mu$$

and where $|\nabla f|$ is the Euclidean length of the gradient ∇f of f. The canonical Gaussian measure with density $(2\pi)^{-n/2}\mathrm{e}^{-|x|^2/2}$ with respect to the Lebesgue measure on $\mathbb{R}^n$ is the basic example of measure μ satisfying (1.1) with $\rho = 1$.

For simplicity, assume furthermore that μ has a strictly positive smooth density which may be written e^{-U} for some smooth function U on $\mathbb{R}^n$. Denote by L the second-order diffusion operator $\mathrm{L} = \Delta - \langle \nabla U, \nabla \rangle$ with invariant measure μ. Integration by parts for L is described by:

E-mail addresses: bobkov@math.umn.edu (S.G. Bobkov), gentil@cict.fr (I. Gentil), ledoux@cict.fr (M. Ledoux).

$$\int f(-\mathrm{L}g)\,\mathrm{d}\mu = \int \langle \nabla f, \nabla g\rangle\,\mathrm{d}\mu$$

for every smooth functions f, g. Under mild growth conditions on U (that will always be satisfied in applications throughout this work), one may consider the time reversible (with respect to μ) semigroup $(P_t)_{t\geqslant 0}$ with generator L. Given f (in the domain of L), $u = u(x,t) = P_t f(x)$ is the fundamental solution of the initial value problem (heat equation with respect to L):

$$\begin{aligned}
&\frac{\partial u}{\partial t} - \mathrm{L}u = 0 \quad \text{in } \mathbb{R}^n \times (0,\infty),\\
&u = f \quad \text{on } \mathbb{R}^n \times \{t = 0\}.
\end{aligned}$$

One of the main results of the contribution [17] by L. Gross is that the logarithmic Sobolev inequality (1.1) for μ holds if and only if the associated heat semigroup $(P_t)_{t\geqslant 0}$ is hypercontractive in the sense that, for every (or some) $1 < p < q < \infty$, and every f (in L^p),

$$\|P_t f\|_q \leqslant \|f\|_p, \tag{1.2}$$

for every $t > 0$ large enough so that

$$\mathrm{e}^{2\rho t} \geqslant \frac{q-1}{p-1}. \tag{1.3}$$

In (1.2), the L^p-norms are understood with respect to the measure μ. The key idea of the proof is to consider a function $q(t)$ of $t \geqslant 0$ such that $q(0) = p$ and to take the derivative in time of $F(t) = \|P_t f\|_{q(t)}$ (for a non-negative smooth function f on $\mathbb{R}^n$). Since the derivative of L^p-norms gives rise to entropy, due to the heat equation $\frac{\partial}{\partial t} P_t f = \mathrm{L} P_t f$ and integration by parts, one gets that:

$$\begin{aligned}
&q(t)^2 F(t)^{q(t)-1} F'(t)\\
&\quad = q'(t)\mathrm{Ent}_\mu\big((P_t f)^{q(t)}\big) + q(t)^2 \int (P_t f)^{q(t)-1} \mathrm{L} P_t f\,\mathrm{d}\mu\\
&\quad = q'(t)\mathrm{Ent}_\mu\big((P_t f)^{q(t)}\big) - 2\big(q(t)-1\big) \int \frac{q(t)^2}{2} |\nabla P_t f|^2 (P_t f)^{q(t)-2}\,\mathrm{d}\mu.
\end{aligned} \tag{1.4}$$

By the logarithmic Sobolev inequality applied to $(P_t f)^{q(t)/2}$, it follows that $F'(t) \leqslant 0$ as soon as $q'(t) = 2\rho(q(t) - 1)$, that is $q(t) = 1 + (p-1)\mathrm{e}^{2\rho t}$, $t \geqslant 0$, which yields the claim. It is classical and easy to see that the same argument also shows that (1.1) is also equivalent to

$$\big\|\mathrm{e}^{P_t f}\big\|_{\mathrm{e}^{2\rho t}} \leqslant \big\|\mathrm{e}^f\big\|_1 \tag{1.5}$$

for every $t \geqslant 0$ and f (cf. [4]). For further comparison, observe that by linearity

$$\big\|\mathrm{e}^{P_t f}\big\|_{a\mathrm{e}^{2\rho t}} \leqslant \ (\text{resp. } \geqslant)\big\|\mathrm{e}^f\big\|_a$$

according as $a \geqslant 0$ (resp. $a \leqslant 0$).

Whenever $-\infty < q < p < 1$ satisfy (1.3), the logarithmic Sobolev inequality is similarly equivalent to the so-called reverse hypercontractivity

$$\|P_t f\|_q \geqslant \|f\|_p \tag{1.6}$$

for every f taking non-negative values.

The main result of this work is to establish a similar relationship for the solutions of Hamilton–Jacobi partial differential equations. Consider the Hamilton–Jacobi initial value problem:

$$
\begin{aligned}
&\frac{\partial v}{\partial t} + \frac{1}{2}|\nabla v|^2 = 0 \quad \text{in } \mathbb{R}^n \times (0,\infty),\\
&v = f \quad \text{on } \mathbb{R}^n \times \{t=0\}.
\end{aligned}
\tag{1.7}
$$

Solutions of (1.7) are described by the Hopf–Lax representation formula as infimum-convolutions. Namely, given a (Lipschitz continuous) function f on $\mathbb{R}^n$, define the infimum-convolution of f with the quadratic cost as:

$$
Q_t f(x) = \inf_{y\in\mathbb{R}^n}\left[f(y) + \frac{1}{2t}|x-y|^2\right], \quad t>0,\, x\in\mathbb{R}^n.
\tag{1.8}
$$

The family $(Q_t)_{t\geqslant 0}$ defines a semigroup with infinitesimal (non-linear) generator $-\frac{1}{2}|\nabla f|^2$. That is, $v = v(x,t) = Q_t f(x)$ is a solution of the Hamilton–Jacobi initial value problem (1.7) (at least almost everywhere). Actually, if in addition f is bounded, the Hopf–Lax formula $Q_t f$ is the pertinent mathematical solution of (1.7), that is its unique viscosity solution (cf., e.g., [3,16]).

Once this has been recognized, it is not difficult to try to follow Gross's idea for the Hamilton–Jacobi equation. Namely, letting now $F(t) = \|\mathrm{e}^{Q_t f}\|_{\lambda(t)}$, $t\geqslant 0$, for some function $\lambda(t)$ with $\lambda(0) = a$, $a\in\mathbb{R}$, the analogue of (1.4) reads as:

$$
\lambda(t)^2 F(t)^{\lambda(t)-1} F'(t) = \lambda'(t)\mathrm{Ent}_\mu\big(\mathrm{e}^{\lambda(t)Q_t f}\big) - \int \frac{\lambda(t)^2}{2}|\nabla Q_t f|^2 \mathrm{e}^{\lambda(t)Q_t f}\,\mathrm{d}\mu.
\tag{1.9}
$$

By the logarithmic Sobolev inequality (1.1) applied to $\mathrm{e}^{\lambda(t)Q_t f}$, $F'(t)\leqslant 0$ as soon as $\lambda'(t) = \rho$, $t\geqslant 0$. As a result (and in complete analogy with (1.5) for example), the logarithmic Sobolev inequality (1.1) shows that, for every $t\geqslant 0$, every $a\in\mathbb{R}$ and every (say bounded) function f,

$$
\big\|\mathrm{e}^{Q_t f}\big\|_{a+\rho t} \leqslant \big\|\mathrm{e}^{f}\big\|_a.
\tag{1.10}
$$

Conversely, if (1.10) holds for every $t\geqslant 0$ and some $a\neq 0$, then the logarithmic Sobolev inequality (1.1) holds. With respect to classical hypercontractivity, it is worthwhile noting that Q_t is defined independently of the underlying measure μ. Actually, hypercontractivity of Hamilton–Jacobi solutions may also be shown to follow from heat kernel hypercontractivity through the so-called vanishing viscosity method. Namely, if u^ε is solution of the heat equation $\partial u^\varepsilon/\partial t = \varepsilon \mathrm{L} u^\varepsilon$ (with initial value $\mathrm{e}^{-f/2\varepsilon}$), then $v^\varepsilon = -2\varepsilon\log u^\varepsilon$ approaches as $\varepsilon\to 0$ the Hopf–Lax solution (1.8). Transferring hypercontractivity of the heat solution u^ε to v^ε yields another approach to our main result. In this Laplace–Varadhan large deviation asymptotic, the second-order term in $\mathrm{L} = \Delta - \langle\nabla U,\nabla\rangle$ is the leading term that gives rise to the Gaussian kernel and the quadratic cost in (1.8) (and an expression for Q_t independent of U and thus of μ).

Due to the homogeneity property $Q_t(sf) = sQ_{st}f$, $s,t>0$, and setting Q for Q_1, (1.10) may be rewritten equivalently as:

$$
\big\|\mathrm{e}^{Qf}\big\|_{r+\rho} \leqslant \big\|\mathrm{e}^{f}\big\|_r
\tag{1.11}
$$

for $r\in\mathbb{R}$. If (1.11) holds for either every $r>0$ (or only large enough) or every $r<0$ (or only large enough), then the logarithmic Sobolev inequality (1.1) holds. The value $r=0$ is however critical.

When $a = 0$ in (1.10), or $r = 0$ in (1.11), these two inequalities actually amount to the infimum-convolution inequality

$$\int \mathrm{e}^{\rho Qf}\,\mathrm{d}\mu \leqslant \mathrm{e}^{\rho\int f\,\mathrm{d}\mu} \tag{1.12}$$

holding for every bounded (or integrable) function f. Inequality (1.12) is known to be the Monge–Kantorovitch–Rubinstein dual version of the transportation cost inequality (see [7] and below):

$$\rho W_2(\mu,\nu)^2 \leqslant H(\nu|\mu) = \mathrm{Ent}_\mu\left(\frac{\mathrm{d}\nu}{\mathrm{d}\mu}\right) \tag{1.13}$$

holding for all probability measures ν absolutely continuous with respect to μ with Radon–Nikodym derivative $\mathrm{d}\nu/\mathrm{d}\mu$. Here W_2 is the Wasserstein distance with quadratic cost:

$$W_2(\mu,\nu)^2 = \inf \iint \frac{1}{2}|x-y|^2\,\mathrm{d}\pi(x,y),$$

where the infimum is running over all probability measures π on $\mathbb{R}^n \times \mathbb{R}^n$ with respective marginals μ and ν and $H(\nu|\mu)$ is the relative entropy, or informational divergence, of ν with respect to μ. (The infimum in W_2 is finite as soon as μ and ν have finite second moment which we shall always assume.) That the transportation cost inequality (1.13) follows from the logarithmic Sobolev inequality (1.1) was established recently by F. Otto and C. Villani [24] and motivated the present work. While the arguments developed in [24] do involve PDE's methods (further inspired by nice geometric interpretations described in [23]), the approach presented here only relies on the basic Hamilton–Jacobi equation (together with the dual formulation (1.12) of the transportation cost inequality (1.13)) and presents a clear view of the connection between logarithmic Sobolev inequalities and transportation cost inequalities. One feature of our approach is the systematic use of the Monge–Kantorovitch dual version of the transportation cost inequality involving infimum-convolution rather than Wasserstein distances.

It is an open problem (although probably with negative answer) to know whether the critical case (1.12) is also equivalent to the logarithmic Sobolev inequality (1.1). When the potential U is convex, it was shown in [24] that the transportation cost inequality (1.13) implies conversely the logarithmic Sobolev inequality (1.1) up to a numerical constant (the precise statement of [24] is somewhat more general and allows small non-convex wells of U). The proof relies on a general HWI inequality involving the entropy H, the Wasserstein distance W_2 and the Fisher information I which may be established using the Brenier–McCann mass transportation by the gradient of a convex function (see [12,24] and the references therein). The hypercontractive tools developed in the present paper do not seem to be of help in providing an alternate description of this converse statement. However, we present in Section 4 a semigroup proof of these results relying on the Bakry–Emery method and Wang's Harnack inequalities [32] by means of a short time parabolic regularization estimate between entropy and Wasserstein distance. In particular, this approach interpolates between the HWI inequality of [24] and the logarithmic Sobolev inequality under exponential integrability of [32]. The subsequent comment note [25] by F. Otto and C. Villani further expands on this theme.

In Section 2 of this work, we give a detailed proof of the main result (1.10). While the general principle outlined above is straightforward, some regularity questions have to be addressed. We also discuss the approach through the vanishing viscosity technique that shows a formal direct equivalence of hypercontractivity for the heat equation and for the Hamilton–Jacobi equation. The principle of proof extends to Riemannian manifolds (with the Riemannian metric as transportation cost). In the next section, we present an alternate deduction of the transportation cost inequalities via the analogue of the Herbst argument. To this task, we first recall the usual Herbst argument, and then adapt it to infimum convolutions. We introduce this section by the Monge–Kantorovitch dual description of transportation cost inequalities. In Section 4, we first mention that quadratic transportation cost inequalities are stronger than the related Poincaré inequalities. We then investigate how to reach HWI and logarithmic Sobolev inequalities for families of log-concave measures following the Bakry–Emery semigroup method. In the fifth section, we show how the Herbst method for infimum convolutions of Section 3 may be used to recover similarly the transportation inequality of M. Talagrand [29] for the exponential measure from the logarithmic Sobolev inequality of [8] (and more generally for measures satisfying a Poincaré inequality). In the final part, we present further applications and discuss possible extensions of the basic principle. In particular, we investigate, following [22] and [9], how Brunn–Minkowski inequalities are related to the infimum-convolution inequalities (1.12) for strictly convex potential. We also discuss the L^1-transportation cost and its relation to some (logarithmic) isoperimetric inequalities.

2. Hamilton–Jacobi equations and logarithmic Sobolev inequalities

This section is devoted to the main result of this work. We first present the direct proof as outlined in the introduction, and then the alternate vanishing viscosity method. We briefly discuss extension to a Riemannian setting.

2.1. Hypercontractivity of Hamilton–Jacobi solutions

In this section, we present our main result connecting logarithmic Sobolev inequalities to hypercontractivity of solutions of Hamilton–Jacobi equations. While the subsequent arguments extend to Riemannian manifolds, we however present, for clarity, the analysis in the more classical Euclidean case. The general principle will apply similarly in the Riemannian setting (Section 2.3).

Let $(Q_t)_{t \geqslant 0}$ be the semigroup of operators:

$$Q_t f(x) = \inf_{y \in \mathbb{R}^n} \left[f(y) + \frac{1}{2t}|x-y|^2 \right], \quad t > 0, x \in \mathbb{R}^n, \tag{2.1}$$

and $Q_0 f(x) = f(x)$. These operators may be applied to arbitrary functions on $\mathbb{R}^n$ with values in $[-\infty, +\infty]$. As is well-known (see, e.g., [3,16]), for any f and $t > 0$, $Q_t f$ is upper semicontinuous. If f is bounded (resp. Lipschitz), $Q_t f$ is bounded and Lipschitz (resp. Lipschitz). Given a bounded function f, $Q_t f(x) \to f(x)$ as $t \to 0$ if and only if f is lower semicontinuous at x.

The infimum convolution Q_t is known as the Hopf–Lax solution of the Hamilton–Jacobi equation:

$$\frac{\partial}{\partial t} Q_t f(x) = -\frac{1}{2}\left|\nabla Q_t f(x)\right|^2, \tag{2.2}$$

with initial value f. More precisely (cf. [16]), given f Lipschitz continuous, the Hopf–Lax f solution is Lipschitz continuous and solves (2.2) almost everywhere in $\mathbb{R}^n \times (0, \infty)$. Standard variants of the classical theory further show that if f is, say bounded, $t \to Q_t f(x)$ is differentiable at every $t \geqslant 0$ for almost every $x \in \mathbb{R}^n$, and (2.2) holds true (at $t > 0$, almost everywhere in x).

Let μ be a probability measure on the Borel sets of $\mathbb{R}^n$. We denote below by $\|\cdot\|_p$, $p \in \mathbb{R}$, the L^p-norms (functionals when $p < 1$) with respect to μ. As is usual, we agree that $\|f\|_0 = \mathrm{e}^{\int \log|f|\,\mathrm{d}\mu}$ whenever $\log|f|$ is μ-integrable. The main result of this work is the following theorem:

THEOREM 2.1. – *Assume that μ is absolutely continuous with respect to Lebesgue measure and that for some $\rho > 0$ and all smooth enough functions f on $\mathbb{R}^n$,*

$$\rho\,\mathrm{Ent}_\mu\big(f^2\big) \leqslant 2\int |\nabla f|^2\,\mathrm{d}\mu. \tag{2.3}$$

Then, for every *bounded measurable function f on $\mathbb{R}^n$, every $t \geqslant 0$ and every $a \in \mathbb{R}$,*

$$\big\|\mathrm{e}^{Q_t f}\big\|_{a+\rho t} \leqslant \big\|\mathrm{e}^f\big\|_a. \tag{2.4}$$

Conversely, if (2.4) *holds for all $t \geqslant 0$ and* some *$a \neq 0$, then the logarithmic Sobolev inequality* (2.3) *holds.*

In Theorem 2.1, inequalities (2.4) are stated for bounded functions for simplicity: they readily extend to larger classes of functions under the proper integrability conditions.

We may define similarly the supremum-convolution semigroup $(\widetilde{Q}_t)_{t>0}$ by:

$$\widetilde{Q}_t f(x) = \sup_{y\in\mathbb{R}^n}\left[f(y) - \frac{1}{2t}|x-y|^2\right], \quad t > 0,\, x \in \mathbb{R}^n,$$

$(\widetilde{Q}_0 f(x) = f(x))$. The operators Q_t and $\widetilde{Q}_t$ are related by the property that for any two functions f and g, $g \geqslant \widetilde{Q}_t f$ if and only if $f \leqslant Q_t g$ so that $\widetilde{Q}_t Q_t f \leqslant f \leqslant Q_t \widetilde{Q}_t f$. We also have that $\widetilde{Q}_t(-f) = -Q_t f$. In particular, the conclusion (2.4) of Theorem 2.1 may be reformulated equivalently on $(\widetilde{Q}_t)_{t\geqslant 0}$ by:

$$\big\|\mathrm{e}^f\big\|_{a+\rho t} \leqslant \big\|\mathrm{e}^{\widetilde{Q}_t f}\big\|_a. \tag{2.5}$$

Note that the families of inequalities (2.4) and (2.5) are stable under the respective semigroups.

If μ is not absolutely continuous, an easy convolution argument leads to (2.4) at least for all bounded continuous functions. Namely, the stability by products of the logarithmic Sobolev inequality shows that if γ_σ is the Gaussian measure on $\mathbb{R}^n$ with covariance $\sigma^2\,\mathrm{Id}$, for every smooth function $\tilde{f}$ on $\mathbb{R}^n \times \mathbb{R}^n$,

$$\min\big(\rho, \sigma^{-1}\big)\mathrm{Ent}_{\mu\otimes\gamma_\sigma}\big(\tilde{f}^2\big) \leqslant 2\int |\nabla \tilde{f}|^2\,\mathrm{d}\mu\otimes\gamma_\sigma.$$

Applied to $\tilde{f}(x,y) = f(x+y)$, $x, y \in \mathbb{R}^n$, for some smooth function f on $\mathbb{R}^n$, we get:

$$\min\big(\rho, \sigma^{-1}\big)\mathrm{Ent}_{\mu*\gamma_\sigma}\big(f^2\big) \leqslant 2\int |\nabla f|^2\,\mathrm{d}\mu*\gamma_\sigma.$$

Theorem 2.1 then applies to $\mu * \gamma_\sigma$. Letting $\sigma \to 0$ yields (2.4) for all bounded continuous functions.

Proof of Theorem 2.1. – In the first part of the argument, we assume that the logarithmic Sobolev inequality (2.3) holds and show that (2.4) is satisfied for any bounded f, and any $t > 0$,

$a \in \mathbb{R}$. By a simple density argument, the logarithmic Sobolev inequality (2.3) holds for all (locally) Lipschitz functions. Let thus f be a bounded function on $\mathbb{R}^n$. (By regularization, it may be assumed that f is compactly supported with bounded derivatives of any orders: however, besides the final step, regularity does not make life easier here.) Let $F(t) = \|e^{Q_t f}\|_{\lambda(t)}$, with $\lambda(t) = a + \rho t$, $t > 0$. For all $t > 0$ and almost every x, the partial derivatives $\frac{\partial}{\partial t} Q_t f(x)$ exist. Thus F is differentiable at every point $t > 0$ where $\lambda(t) \neq 0$. For such points, we get that:

$$\lambda^2(t) F(t)^{\lambda(t)-1} F'(t) = \rho \operatorname{Ent}_\mu\bigl(e^{\lambda(t) Q_t f}\bigr) + \int \lambda^2(t) \frac{\partial}{\partial t} Q_t f e^{\lambda(t) Q_t f} \, d\mu. \tag{2.6}$$

Since

$$\frac{\partial}{\partial t} Q_t f(x) = -\frac{1}{2}\bigl|\nabla Q_t f(x)\bigr|^2$$

almost everywhere in x, and since μ is absolutely continuous,

$$\lambda^2(t) F(t)^{\lambda(t)-1} F'(t) = \rho \operatorname{Ent}_\mu\bigl(e^{\lambda(t) Q_t f}\bigr) - \int \frac{\lambda(t)^2}{2} |\nabla Q_t f|^2 e^{\lambda(t) Q_t f} \, d\mu.$$

Now, since $Q_t f(x)$ is Lipschitz in x for every $t > 0$, we may apply the logarithmic Sobolev inequality (2.3) to $e^{\lambda(t) Q_t f}$ to deduce that $F'(t) \leqslant 0$ for all $t > 0$ except possibly one point (in case $a < 0$). Since F is continuous, it must be non-increasing. Continuity of $Q_t f(x)$ at $t = 0$ however requires f be lower semicontinuous at the point x. Apply then the result to the maximal lower semicontinuous function majorized by f to conclude. (Alternatively, as mentioned previously, we may regularize f to start with and assume f bounded and Lipschitz for example.) The first part of the theorem is established.

Turning to the converse, let f be a bounded C^1 function satisfying (2.4) for every $t > 0$ and some $a \neq 0$. Under (2.4), it thus must be that $F'(0) \leqslant 0$. Since f is differentiable, $\lim_{t \to 0} Q_t f(x) = f(x)$ and

$$\frac{\partial}{\partial t} Q_t f(x)\Big|_{t=0} = \lim_{t \to 0} \frac{1}{t}\bigl[Q_t f(x) - f(x)\bigr] = -\frac{1}{2}\bigl|\nabla f(x)\bigr|^2$$

at every point x so that (2.6) as $t \to 0$ yields

$$\rho \operatorname{Ent}_\mu\bigl(e^{af}\bigr) \leqslant \frac{1}{2} \int |a \nabla f|^2 e^{af} \, d\mu.$$

Since $a \neq 0$, this amounts to (2.3) by setting $g^2 = e^{af}$. The proof of Theorem 2.1 is complete. □

Remark 2.2. – As in the classical case, the proof of Theorem 2.1 similarly shows that a defective logarithmic Sobolev inequality of the type

$$\rho \operatorname{Ent}_\mu\bigl(f^2\bigr) \leqslant 2 \int |\nabla f|^2 \, d\mu + C \int f^2 \, d\mu$$

for some $C > 0$ is equivalent to the hypercontractive bounds ($t \geqslant 0$, $a \in \mathbb{R}$)

$$\bigl\|e^{Q_t f}\bigr\|_{a+\rho t} \leqslant e^{M(t)} \bigl\|e^f\bigr\|_a,$$

where:

$$M(t) = \frac{Ct}{a(a+\rho t)}.$$

2.2. Hypercontractivity and vanishing viscosity

An alternate proof of Theorem 2.1 may be provided by the tool of vanishing viscosity (cf. [16]). We only briefly outline the principle that requires some further technical arguments. The idea is to add a small noise to the Hamilton–Jacobi equation to turn it after an exponential change of functions into the heat equation. Given a smooth function f, and $\varepsilon > 0$, denote namely by $v^\varepsilon = v^\varepsilon(x,t)$ the solution of the initial value partial differential equation:

$$\begin{aligned}
&\frac{\partial v^\varepsilon}{\partial t} + \frac{1}{2}\left|\nabla v^\varepsilon\right|^2 - \varepsilon \mathrm{L} v^\varepsilon = 0 \quad \text{in } \mathbb{R}^n \times (0,\infty),\\
&v^\varepsilon = f \quad \text{on } \mathbb{R}^n \times \{t=0\}.
\end{aligned}$$

As $\varepsilon \to 0$, it is expected that v^ε approaches in a reasonable sense the solution v of (1.7). It is easy to check that $u^\varepsilon = \mathrm{e}^{-v^\varepsilon/2\varepsilon}$ is a solution of the heat equation $\partial u^\varepsilon/\partial t = \varepsilon \mathrm{L} u^\varepsilon$ (with initial value $\mathrm{e}^{-f/2\varepsilon}$). Therefore,

$$u^\varepsilon = P_{\varepsilon t}\big(\mathrm{e}^{-f/2\varepsilon}\big).$$

It must be emphasized that the perturbation argument by a small noise has a clear picture in the probabilistic language of large deviations. Namely, the asymptotic of

$$v^\varepsilon = -2\varepsilon \log P_{\varepsilon t}\big(\mathrm{e}^{-f/2\varepsilon}\big)$$

as $\varepsilon \to 0$ is a Laplace–Varadhan asymptotic with rate described precisely by the infimum convolution of f with the quadratic large deviation rate function for the heat semigroup (cf., e.g., [3]). In this limit, the second order Laplace operator is the leading term in the definition of $\mathrm{L} = \Delta - \langle \nabla, \nabla U\rangle$ so that the limiting solution u given by the infimum-convolution $Q_t f$ is independent of the potential U and thus of μ. In particular, this asymptotic is explicit on the basic Ornstein–Uhlenbeck example.

Apply now classical hypercontractivity to u^ε. More precisely, for $b > a > 0$ fixed, apply the reverse hypercontractivity inequality (1.6) with $0 > p = -2\varepsilon a > q = -2\varepsilon b$ and

$$\mathrm{e}^{2\varepsilon\rho t} = \frac{1+2\varepsilon b}{1+2\varepsilon a}.$$

It follows that:

$$\big\|\mathrm{e}^{v^\varepsilon}\big\|_b \leqslant \big\|\mathrm{e}^{f}\big\|_a.$$

Now, as $\varepsilon \to 0$, $t > 0$ is such that $b = a + \rho t$. We thus recover in this way the main Theorem 2.1. Note however that it was necessary to go through reverse hypercontractivity of the heat semigroup to reach the conclusion.

2.3. Extension to Riemannian manifolds

As announced, Theorem 2.1 and its proof extend to the setting of logarithmic Sobolev inequalities on Riemannian manifolds and infimum-convolutions with the Riemannian metric as in [24]. We briefly outline in this subsection the corresponding result. Let M be a smooth complete Riemannian manifold of dimension n and Riemannian metric d. Let μ be a probability measure absolutely continuous with respect to the standard volume element on M satisfying, for some $\rho > 0$ and all smooth enough functions f on M, the logarithmic Sobolev inequality:

$$\rho \,\mathrm{Ent}_\mu\big(f^2\big) \leqslant 2 \int |\nabla f|^2 \,\mathrm{d}\mu.$$

Here $|\nabla f|$ now stands for the Riemannian length of the gradient of f. Let, for $t > 0$, $x \in M$,

$$Q_t f(x) = \inf_{y \in M}\left[f(y) + \frac{1}{2t} d(x, y)^2 \right].$$

It may be observed that $(Q_t)_{t \geqslant 0}$ forms a semigroup since for the geodesic distance,

$$\inf_{z \in M}\left[\frac{1}{t} d(x, z)^2 + \frac{1}{s} d(z, y)^2 \right] = \frac{1}{s+t} d(x, y)^2$$

for all $x, y \in M$ and $s, t > 0$. Following the argument in the classical Euclidean case (cf. [31]), one shows similarly that $v = v(x, t) = Q_t f(x)$ is again a solution of the initial-value Hamilton–Jacobi problem on M:

$$\begin{aligned} &\frac{\partial v}{\partial t} + \frac{1}{2} |\nabla v|^2 = 0 \quad \text{in } M \times (0, \infty), \\ &v = f \quad \text{on } M \times \{t = 0\}. \end{aligned}$$

Theorem 2.1 and its proof thus readily extend to this case. It might be easier to develop the extension of Hamilton–Jacobi equations to Riemannian manifolds in the compact case first. Regularizing f into a compactly supported function as in the proof of Theorem 2.1 allows us to reduce to this case if necessary.

3. Herbst's argument and transportation inequalities

There is yet another way from logarithmic Sobolev inequalities to infimum-convolution inequalities that goes through the so-called Herbst method (cf. [18]). To introduce it, we first summarize the Monge–Kantorovitch dual versions of the transportation cost inequalities. We then recall the classical Herbst argument and apply it in the infimum-convolution context.

3.1. Monge–Kantorovitch duality

Let us start with the Wassertein distance with linear cost between two probability measures on $\mathbb{R}^n$ defined by:

$$W_1(\mu, \nu) = \inf \iint |x - y| \,\mathrm{d}\pi(x, y),$$

where the infimum is running over all probability measures π on $\mathbb{R}^n \times \mathbb{R}^n$ with respective marginals μ and ν (having a finite first moment). Bythe Monge–Kantorovitch dual characterization (cf. [15,26]),

$$W_1(\mu, \nu) = \sup\left[\int g \,\mathrm{d}\nu - \int f \,\mathrm{d}\mu \right], \tag{3.1}$$

where the supremum is running over all bounded measurable functions f and g such that

$$g(x) \leqslant f(y) + |x - y|$$

for every $x, y \in \mathbb{R}^n$. Perhaps more classicaly, we have equivalently that:

$$W_1(\mu,\nu) = \sup\left[\int g\,\mathrm{d}\mu - \int g\,\mathrm{d}\nu\right], \tag{3.2}$$

where the supremum is running over all Lipschitz functions g with $\|g\|_{\mathrm{Lip}} \leqslant 1$.

The general form of the dual Monge–Kantorovitch representation of some metric space (E,d) for example indicates that (cf. [26]):

$$\inf \iint T(x,y)\,\mathrm{d}\pi(x,y) = \sup\left[\int g\,\mathrm{d}\nu - \int f\,\mathrm{d}\mu\right], \tag{3.3}$$

where the infimum is running over all probability measures π with marginals μ and ν such that T is integrable with respect to π and where the supremum is over all pairs (g,f) of bounded measurable functions (or respectively ν and μ-integrable) such that for all x, y,

$$g(x) \leqslant f(y) + T(x,y).$$

Here T is upper semicontinuous, π-integrable and such that $T(x,y) \leqslant a(x) + b(y)$ for some measurable functions a and b. On $\mathbb{R}^n$, the supremum on the right-hand side of (3.3) may be taken over smaller classes of smooth functions, such as bounded Lipschitz or so on. (This provides an alternate regularization procedure for the arguments developed in the next sections.)

For the quadratic cost in particular, we thus have that:

$$W_2(\mu,\nu)^2 = \sup\left[\int g\,\mathrm{d}\nu - \int f\,\mathrm{d}\mu\right], \tag{3.4}$$

where the supremum is running over all bounded functions f and g such that

$$g(x) \leqslant f(y) + \frac{1}{2}|x-y|^2$$

for every $x, y \in \mathbb{R}^n$. In the infimum-convolution notation,

$$g(x) = \inf_{y\in\mathbb{R}^n}\left[f(y) + \frac{1}{2}|x-y|^2\right] = Qf(x)$$

achieves the optimal choice.

3.2. Linear transportation cost

In this section, we recall the Herbst argument and its interpretation as a transportation result with linear cost. Assume the logarithmic Sobolev inequality

$$\rho\,\mathrm{Ent}_\mu\big(f^2\big) \leqslant 2\int |\nabla f|^2\,\mathrm{d}\mu \tag{3.5}$$

holds for some $\rho > 0$ and all smooth enough (locally Lipschitz) functions f on $\mathbb{R}^n$. For simplicity, assume below that μ is absolutely continuous with respect to Lebesgue measure.

Now, let g be a (bounded) Lipschitz function on $\mathbb{R}^n$ with Lipschitz coefficient $\|g\|_{\mathrm{Lip}}$. Let us then apply (3.5) to $f^2 = \mathrm{e}^{\lambda g - \lambda^2 \|g\|_{\mathrm{Lip}}^2/2\rho}$ where $\lambda \in \mathbb{R}$. Set $G(\lambda) = \int \mathrm{e}^{\lambda g - \lambda^2 \|g\|_{\mathrm{Lip}}^2/2\rho}\,\mathrm{d}\mu$. Since $|\nabla g| \leqslant \|g\|_{\mathrm{Lip}}$ almost everywhere, we get from (3.5) that, for every $\lambda \in \mathbb{R}$:

$$\int \left[\lambda g - \frac{1}{2\rho}\lambda^2 \|g\|_{\mathrm{Lip}}^2\right] \mathrm{e}^{\lambda g - \lambda^2 \|g\|_{\mathrm{Lip}}^2/2\rho}\,\mathrm{d}\mu - G(\lambda)\log G(\lambda) \leqslant \frac{1}{2\rho}\lambda^2 \|g\|_{\mathrm{Lip}}^2 G(\lambda).$$

In other words,

$$\lambda G'(\lambda) \leqslant G(\lambda)\log G(\lambda), \quad \lambda \in \mathbb{R}. \tag{3.6}$$

This differential inequality is easily integrated to yield, since $G'(0) = \int g\,\mathrm{d}\mu$, that for every Lipschitz (integrable) function g on $\mathbb{R}^n$:

$$\int \mathrm{e}^g\,\mathrm{d}\mu \leqslant \mathrm{e}^{\int g\,\mathrm{d}\mu + \|g\|_{\mathrm{Lip}}^2/2\rho}. \tag{3.7}$$

By Chebychev's inequality, this inequality describes the concentration properties of a measure μ satisfying a logarithmic Sobolev inequality (cf. [18]).

Inequality (3.7) has been recognized in [7] as a transportation inequality for the W_1 Wasserstein distance in the form of:

$$\rho W_1^2(\mu,\nu) \leqslant 2H(\nu|\mu) = 2\,\mathrm{Ent}_\mu\left(\frac{\mathrm{d}\nu}{\mathrm{d}\mu}\right) \tag{3.8}$$

holding for all probability measures ν absolutely continuous with respect to μ with Radon–Nikodym derivative $\mathrm{d}\nu/\mathrm{d}\mu$. Namely by (3.8) and (3.1) (one could use completely similarly (3.2)), for every bounded measurable functions f and g such that $g(x) \leqslant f(y) + |x-y|$ for all $x, y \in \mathbb{R}^n$,

$$\int g\,\mathrm{d}\nu - \int f\,\mathrm{d}\mu \leqslant \sqrt{\frac{2}{\rho}\mathrm{Ent}_\mu\left(\frac{\mathrm{d}\nu}{\mathrm{d}\mu}\right)},$$

or, equivalently, for every $\lambda > 0$,

$$\int g\,\mathrm{d}\nu - \int f\,\mathrm{d}\mu \leqslant \frac{\lambda}{2\rho} + \frac{1}{\lambda}\mathrm{Ent}_\mu\left(\frac{\mathrm{d}\nu}{\mathrm{d}\mu}\right).$$

Set $\varphi = \mathrm{d}\nu/\mathrm{d}\mu$. The preceding indicates that

$$\int \psi\varphi\,\mathrm{d}\mu \leqslant \mathrm{Ent}_\mu(\varphi),$$

where $\psi = \lambda g - \lambda^2/2\rho - \lambda\int f\,\mathrm{d}\mu$. Since this inequality holds for every choice of φ (i.e. ν), applying it to $\varphi = \mathrm{e}^\psi/\int \mathrm{e}^\psi\,\mathrm{d}\mu$ yields $\log\int \mathrm{e}^\psi\,\mathrm{d}\mu \leqslant 0$. In other words,

$$\int \mathrm{e}^{\lambda g}\,\mathrm{d}\mu \leqslant \mathrm{e}^{\lambda\int f\,\mathrm{d}\mu + \lambda^2/2\rho}.$$

When f is Lipschitz with $\|f\|_{\mathrm{Lip}} \leqslant 1$, one may choose $g = f$ so that the latter exactly amounts to (3.7). Since

$$\mathrm{Ent}_\mu(\varphi) = \sup \int \varphi\psi\,\mathrm{d}\mu,$$

where the supremum is running over all ψ's such that $\int e^{\psi}\,d\mu \leqslant 1$, the preceding argument clearly indicates that (3.7) is actually equivalent to (3.8). This result easily extends to arbitraty metric spaces.

3.3. Quadratic transportation cost

The aim of this section is to describe how the preceding Herbst argument may be applied completely similarly to infimum-convolutions. In particular, we recover in this case the conclusion of Theorem 2.1 at the critical value $a = 0$.

Given a (bounded Lipschitz) function g on $\mathbb{R}^n$, apply now the logarithmic Sobolev inequality (3.5) to $f^2 = e^{\rho Q(\lambda g)}$ (where we recall that $Q = Q_1$). Since $Q(\lambda g) = \lambda Q_\lambda g$, $\lambda > 0$, we see from the Hamilton–Jacobi equation that, almost everywhere in space:

$$Q(\lambda g) = \lambda \frac{\partial}{\partial \lambda} Q(\lambda g) + \frac{1}{2}\left|\nabla Q(\lambda g)\right|^2.$$

We thus immediately deduce from the logarithmic Sobolev inequality (3.5) the differential inequality (3.6) on $G(\lambda) = \int e^{\rho Q(\lambda g)}\,d\mu$. Since $G'(0) = \rho \int g\,d\mu$, it follows similarly that:

$$\int e^{\rho Q g}\,d\mu \leqslant e^{\rho \int g\,d\mu}, \tag{3.9}$$

that is the infimum-convolution inequality (1.12).

Inequality (3.9) amounts, as announced in the introduction, to the transportation cost inequality for the quadratic cost

$$\rho W_2(\mu,\nu)^2 \leqslant H(\nu|\mu) = \mathrm{Ent}_\mu\left(\frac{d\nu}{d\mu}\right) \tag{3.10}$$

for every ν absolutely continuous with respect to μ. Exactly as for the equivalence between (3.7) and (3.8), by the dual description of W_2:

$$\int g\,d\nu - \int f\,d\mu \leqslant \frac{1}{\rho}\mathrm{Ent}_\mu\left(\frac{d\nu}{d\mu}\right)$$

for all bounded functions f and g such that

$$g(x) \leqslant f(y) + \frac{1}{2}|x-y|^2$$

for every $x, y \in \mathbb{R}^n$. Since $g = Qf$ achieves the optimal choice, setting $\varphi = d\nu/d\mu$, the preceding amounts to

$$\int \psi\varphi\,d\mu \leqslant \mathrm{Ent}_\mu(\varphi),$$

where $\psi = Qf - \int f\,d\mu$. Since the inequality holds for every choice of φ, it is equivalent to say that $\int e^{\rho\psi} \leqslant 1$, that is exactly (3.9).

As a consequence of either Theorem 2.1 or the preceding, we may state the following corollary first established in [24].

COROLLARY 3.1. – *Assume that* μ *is absolutely continuous and that for some* $\rho > 0$ *and all smooth enough functions* f *on* $\mathbb{R}^n$:

$$\rho \operatorname{Ent}_\mu\bigl(f^2\bigr) \leqslant 2 \int |\nabla f|^2 \,\mathrm{d}\mu.$$

Then, for every probability measure ν *absolutely continuous with respect to* μ,

$$\rho W_2(\mu,\nu)^2 \leqslant H(\nu|\mu).$$

Replacing $|x-y|$ by the Riemannian distance $d(x,y)$ yields the same conclusion on a smooth manifold M.

It might be worthwhile mentioning that whenever g is Lipschitz,

$$Qg \geqslant g - \frac{1}{2}\|g\|_{\mathrm{Lip}}^2.$$

So clearly, (3.9) represents an improvement upon (3.7) (replacing g by g/ρ). Actually, Theorem 2.1 (cf. (1.11)) then indicates that for every $r \in \mathbb{R}$:

$$\bigl\|\mathrm{e}^g\bigr\|_{\rho+r} \leqslant \bigl\|\mathrm{e}^g\bigr\|_r \mathrm{e}^{\|g\|_{\mathrm{Lip}}^2/2},$$

a much stronger property.

4. Semigroup tools and HWI inequalities

In this section, we examine some converse results from transportation cost inequalities to logarithmic Sobolev inequalities. We first describe how quadratic transportation cost inequalities imply spectral inequalities. Then, under appropriate log-concavity assumptions on the underlying measure, we review the Bakry–Emery criterion and put in parallel the HWI inequalities of [24] and the results of [32].

4.1. Transportation cost inequalities and spectral gap

Using again the dual Monge–Kantorovitch description (1.12) of the quadratic transportation inequality (1.13), it is not difficult to see that (1.12) implies the spectral gap, or Poincaré inequality, for μ, in the sense that for all smooth functions f on $\mathbb{R}^n$:

$$\rho \operatorname{Var}_\mu(f) \leqslant \int |\nabla f|^2 \,\mathrm{d}\mu, \tag{4.1}$$

where $\operatorname{Var}_\mu(f) = \int f^2\,\mathrm{d}\mu - (\int f\,\mathrm{d}\mu)^2$. Indeed, homogeneity in (1.12) yields

$$\int \mathrm{e}^{\rho t Q_t f}\,\mathrm{d}\mu \leqslant \mathrm{e}^{\rho t \int f\,\mathrm{d}\mu}.$$

As $t \to 0$, $Q_t f \sim f - \frac{t}{2}|\nabla f|^2$ so that:

$$1 + \rho t \int f\,\mathrm{d}\mu - \frac{\rho t^2}{2}\int |\nabla f|^2\,\mathrm{d}\mu + \frac{\rho^2 t^2}{2}\int f^2\,\mathrm{d}\mu \leqslant 1 + \rho t \int f\,\mathrm{d}\mu + \frac{\rho^2 t^2}{2}\biggl(\int f\,\mathrm{d}\mu\biggr)^2 + \mathrm{o}\bigl(t^2\bigr)$$

and thus (4.1). A different derivation of this result is given in [24].

It is well-known and classical that, applying the logarithmic Sobolev inequality (1.1) to $1 + tf$ and letting $t \to 0$ also yields (4.1). Furthermore, both the logarithmic Sobolev inequality (1.1) and the transportation cost inequality (1.12) (or (1.13)) entail concentration properties. In particular, logarithmic Sobolev inequality and the transportation inequality for the quadratic cost are stable by products and therefore lead to dimension free concentration inequalities (cf. [7,18, 21,29] etc.).

4.2. The Bakry–Emery criterion

Before turning to our main question in the next subsection, it is worthwhile to briefly review the Bakry–Emery criterion [2,4,18], for logarithmic Sobolev inequalities under strict log-concavity of the measure.

Let thus $\mathrm{d}\mu = \mathrm{e}^{-U}\,\mathrm{d}x$ be a probability measure on the Borel sets of $\mathbb{R}^n$ where U is a smooth potential.

THEOREM 4.1. – *Assume that for some $c > 0$, $\mathrm{Hess}(U)(x) \geqslant c\,\mathrm{Id}$ in the sense of symmetric matrices uniformly in $x \in \mathbb{R}^n$. Then μ satisfies the logarithmic Sobolev inequality*:

$$\mathrm{Ent}_\mu\big(f^2\big) \leqslant \frac{2}{c}\int |\nabla f|^2\,\mathrm{d}\mu$$

for every smooth function f on $\mathbb{R}^n$.

The proof by D. Bakry and M. Emery of this result relies on the commutation properties of the gradient with the semigroup $(P_t)_{t\geqslant 0}$ with generator $\mathrm{L} = \Delta - \langle \nabla U, \nabla\rangle$. Namely, the condition $\mathrm{Hess}(U)(x) \geqslant c\,\mathrm{Id}$ uniformly in $x \in \mathbb{R}^n$ for some $c \in \mathbb{R}$ (non necessarily strictly positive) is actually equivalent to saying that for every smooth function f:

$$|\nabla P_t f| \leqslant \mathrm{e}^{-ct} P_t\big(|\nabla f|\big) \tag{4.2}$$

(cf. [2,19]). Then, given a smooth strictly positive bounded function f, we may write:

$$\mathrm{Ent}_\mu(f) = -\int_0^\infty \frac{\mathrm{d}}{\mathrm{d}t}\int P_t f \log P_t f\,\mathrm{d}\mu\,\mathrm{d}t = \int_0^\infty I(P_t f)\,\mathrm{d}t,$$

where

$$I(P_t f) = \int \frac{|\nabla P_t f|^2}{P_t f}\,\mathrm{d}\mu$$

is the Fisher information of $P_t f$. By (4.2) and the Cauchy–Schwarz inequality:

$$|\nabla P_t f|^2 \leqslant \mathrm{e}^{-2ct} P_t\left(\frac{|\nabla f|^2}{f}\right) P_t f,$$

so that, by invariance of P_t,

$$I(P_t f) \leqslant \mathrm{e}^{-2ct} I(f). \tag{4.3}$$

When $c > 0$, it immediately follows that:

$$\mathrm{Ent}_\mu(f) \leqslant \frac{1}{2c} I(f)$$

which amounts to Theorem 4.1 by changing f into f^2.

4.3. HWI inequalities

We examine here what happens to the Bakry–Emery argument when the lower bound c on the Hessian of U is not strictly positive. While the argument clearly breaks down, it may efficiently be complemented by transportation cost inequalities. We reach in this way the HWI inequalities of [24]. For simplicity, we again work in the Euclidean case, all the results and methods however going through in a Riemannian setting.

Namely, for any $T > 0$, we may still apply the Bakry–Emery criterion up to time T. That is, for any smooth positive and bounded function f on $\mathbb{R}^n$ such that $\int f\,\mathrm{d}\mu = 1$, we may write:

$$\mathrm{Ent}_\mu(f) = \int_0^T I(P_t f)\,\mathrm{d}t + \mathrm{Ent}_\mu(P_T f).$$

Assuming that $\mathrm{Hess}(U) \geqslant c\,\mathrm{Id}$ for some $c \in \mathbb{R}$, and using (4.3) shows that:

$$\mathrm{Ent}_\mu(f) \leqslant \alpha(T) I(f) + \mathrm{Ent}_\mu(P_T f), \tag{4.4}$$

where

$$\alpha(T) = \frac{1 - \mathrm{e}^{-2cT}}{2c} \quad (= T \text{ if } c = 0).$$

The idea is now to control $\mathrm{Ent}_\mu(P_T f)$, $T > 0$, by some transportation bound. We will prove the following lemma that describes a kind of reverse transportation cost inequality for $P_T f$ in the form of a short time parabolic regularization estimate. In the subsequent comment note [25], F. Otto and C. Villani mention that Lemma 4.2 below may actually be shown to follow from their proof of Theorem 4.3 using the Brenier–McCann transference plan theorem. They establish in the same way a stronger regularization estimate showing that both entropy and Fisher information become finite in arbitrarily short time (like $\mathrm{O}(t^{-1})$ and $\mathrm{O}(t^{-2})$ respectively) as a variant of an estimate for gradient flows of a convex function on a Hilbert space going back to H. Brézis [11, Theorem 3.7].

LEMMA 4.2. – *Assume* $\mathrm{Hess}(U) \geqslant c\,\mathrm{Id}$, $c \in \mathbb{R}$, *and denote by* $(P_t)_{t \geqslant 0}$ *the semigroup with generator* $\mathrm{L} = \Delta - \langle \nabla U, \nabla \rangle$. *Let* f *on* $\mathbb{R}^n$ *be non-negative and such that* $\int f\,\mathrm{d}\mu = 1$. *Then, for any* $T > 0$:

$$\mathrm{Ent}_\mu(P_T f) \leqslant \left(\frac{1}{2\alpha(T)} - c\right) W_2(\mu, \nu)^2,$$

where $\mathrm{d}\nu = f\,\mathrm{d}\mu$.

Optimizing in $T > 0$ in (4.4) together with Lemma 4.2, we obtain the following result that describes the so-called HWI inequalities connecting entropy H, Wassertein distance W_2 and Fisher information I.

THEOREM 4.3. – *Let* $\mathrm{d}\mu = \mathrm{e}^{-U}\,\mathrm{d}x$ *and assume that* $\mathrm{Hess}(U) \geqslant c\,\mathrm{Id}$ *for some* $c \in \mathbb{R}$. *Then, for every smooth non-negative function* f *such that* $\int f\,\mathrm{d}\mu = 1$:

$$H(\nu|\mu) = \mathrm{Ent}_\mu(f) \leqslant \sqrt{2I(f)}\,W_2(\mu, \nu) - c W_2(\mu, \nu)^2,$$

(*where we recall that* $\mathrm{d}\nu = f\,\mathrm{d}\mu$).

Theorem 4.3 has been obtained by F. Otto and C. Villani [24] using the Brenier–McCann mass transport [10,20] together with further PDE arguments. A simple proof, relying on the same tool, was recently given by D. Cordero-Erausquin [12]. Theorem 4.3 admits the following corollary that complements Theorem 4.1.

COROLLARY 4.4. – *Let* $\mathrm{d}\mu = \mathrm{e}^{-U}\,\mathrm{d}x$ *and assume that* $\mathrm{Hess}(U) \geqslant c\,\mathrm{Id}$ *for some* $c \leqslant 0$. *Assume that for some* $\rho > 0$ *and every* ν,

$$\rho W_2(\mu,\nu)^2 \leqslant H(\nu|\mu).$$

Then, provided that $1 + c/\rho > 0$, μ *satisfies the logarithmic Sobolev inequality*

$$\rho' \,\mathrm{Ent}_\mu\big(f^2\big) \leqslant 2\int |\nabla f|^2\,\mathrm{d}\mu$$

for every smooth f *with*

$$\rho' = \frac{\rho}{4}\bigg(1 + \frac{c}{\rho}\bigg)^2.$$

To complete our proof of Theorem 4.3, we have to establish Lemma 4.2. To this task, we make use of a Harnack type result of F.-Y. Wang in [32], that actually bridges the result of [24] with the logarithmic Sobolev inequalities under exponential integrability of [32] (see also [1,18]).

Proof of Lemma 4.2. – Rewrite first $\mathrm{Ent}_\mu(P_T f)$ by time reversibility as:

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

We bound $P_T(\log P_T f)$ by the method of [32]. Fix x, y in $\mathbb{R}^n$. Let $x(t) = \frac{1-t}{T}x + \frac{t}{T}y$, $0 \leqslant t \leqslant T$. Let further $h:[0,T] \to [0,T]$ be a C^1 speed function such that $h(0) = 0$ and $h(T) = T$. Set $\gamma(t) = x \circ h(t)$ and

$$\Psi\big(t,\gamma(t)\big) = P_t(\log P_{2T-t} f)\big(\gamma(t)\big), \quad 0 \leqslant t \leqslant T.$$

We have

$$\begin{aligned}\frac{\mathrm{d}\Psi}{\mathrm{d}t} &= -P_t\bigg(\frac{|\nabla P_{2T-t} f|^2}{(P_{2T-t} f)^2}\bigg)\big(\gamma(t)\big) + \frac{h'(t)}{T}\big\langle \nabla P_t(\log P_{2T-t} f), y - x\big\rangle \\ &\leqslant -P_t\bigg(\frac{|\nabla P_{2T-t} f|^2}{(P_{2T-t} f)^2}\bigg)\big(\gamma(t)\big) + \frac{|h'(t)|}{T}|x-y|\big|\nabla P_t(\log P_{2T-t} f)\big|.\end{aligned}$$

Using (4.2),

$$\big|\nabla P_t(\log P_{2T-t} f)\big| \leqslant \mathrm{e}^{-ct} P_t\bigg(\frac{|\nabla P_{2T-t} f|}{P_{2T-t} f}\bigg).$$

Hence, with

$$X = \frac{|\nabla P_{2T-t} f|^2}{(P_{2T-t} f)^2} \quad \text{and} \quad Y = \frac{|h'(t)|}{2T}|x-y|\mathrm{e}^{-ct},$$

we have that

$$\frac{\mathrm{d}\Psi}{\mathrm{d}t} \leqslant P_t\big(-X^2 + 2XY\big)$$

and thus

$$\frac{\mathrm{d}\Psi}{\mathrm{d}t} \leqslant P_t\big(Y^2\big) = \frac{|h'(t)|^2}{4T^2}|x-y|^2\mathrm{e}^{-2ct}.$$

It follows that:

$$P_T(\log P_T f)(x) - \log P_{2T} f(y) \leqslant \frac{|x-y|^2}{4T^2}\int_0^T \big|h'(t)\big|^2\mathrm{e}^{-2ct}\,\mathrm{d}t.$$

For the optimal choice of the speed h, that is

$$h(t) = T\big(\mathrm{e}^{2cT}-1\big)^{-1}\big(\mathrm{e}^{2ct}-1\big), \quad 0 \leqslant t \leqslant T,$$

this leads to

$$P_T(\log P_T f)(x) \leqslant \log P_{2T} f(y) + \frac{1}{2S}|x-y|^2, \tag{4.5}$$

where

$$\frac{1}{S} = \frac{1}{2\alpha(T)} - c.$$

Inequality (4.5) is the analogue, adapted to our purposes, of the Harnack inequality of [32]. For x fixed, take then the infimum in y in (4.5) to get:

$$P_T(\log P_T f)(x) \leqslant Q_S\varphi(x),$$

where $\varphi = \log P_{2T} f$. Since by Jensen's inequality

$$\int \varphi\,\mathrm{d}\mu = \int \log P_{2T} f\,\mathrm{d}\mu \leqslant \log\bigg(\int P_{2T} f\,\mathrm{d}\mu\bigg) = 0,$$

we actually have that:

$$P_T(\log P_T f) \leqslant Q_S\varphi - \int \varphi\,\mathrm{d}\mu.$$

Therefore,

$$\mathrm{Ent}_\mu(P_T f) = \int f P_T(\log P_T f)\,\mathrm{d}\mu \leqslant \sup\bigg[\int Q_S\varphi\,\mathrm{d}\nu - \int \varphi\,\mathrm{d}\mu\bigg],$$

where the supremum is over all bounded measurable functions φ. By the dual Monge–Kantorovitch description (3.4) of W_2 together with the scaling property of infimum-convolutions, the lemma is established. □

As mentioned before, Theorem 4.3 and its proof, in particular Lemma 4.2, hold similarly in a Riemannian context.

Remark 4.4. – Lemma 4.2 provides a bridge between the logarithmic Sobolev inequalities of Theorem 4.3 under the quadratic transportation cost and the result of F.-Y. Wang [32] under exponential integrability of the square distance, immediate consequence of linear transportation cost. Indeed, if we integrate inequality (4.5) in $\mathrm{d}\mu(y)$ rather than to take the infimum in y, we get that:

$$\begin{aligned}\mathrm{Ent}_\mu(P_T f) &= \int P_T f \log P_T f \,\mathrm{d}\mu \\ &\leqslant \iint f(x) \log P_{2T} f(y) \,\mathrm{d}\mu(x)\,\mathrm{d}\mu(y) + \iint f(x) \frac{|x-y|^2}{4\alpha(T)\mathrm{e}^{2cT}} \,\mathrm{d}\mu(x)\,\mathrm{d}\mu(y) \\ &\leqslant \iint f(x) \frac{|x-y|^2}{2S} \,\mathrm{d}\mu(x)\,\mathrm{d}\mu(y),\end{aligned}$$

where we used Jensen's inequality. By Young's inequality $ab \leqslant a \log a + \mathrm{e}^b$, $a \geqslant 0$, $b \in \mathbb{R}$,

$$\iint f(x) \frac{|x-y|^2}{2S} \,\mathrm{d}\mu(x)\,\mathrm{d}\mu(y) \leqslant \frac{1}{2}\mathrm{Ent}_\mu(f) + \iint \mathrm{e}^{\frac{|x-y|^2}{S}} \,\mathrm{d}\mu(x)\,\mathrm{d}\mu(y).$$

Together with (4.4), we thus get:

$$\mathrm{Ent}_\mu(f) \leqslant 2\alpha(T) I(f) + 2 \iint \mathrm{e}^{\frac{|x-y|^2}{S}} \,\mathrm{d}\mu(x)\,\mathrm{d}\mu(y). \tag{4.6}$$

Assume now that for some $\varepsilon > 0$,

$$\iint \mathrm{e}^{(-\tilde{c}+\varepsilon)|x-y|^2} \,\mathrm{d}\mu(x)\,\mathrm{d}\mu(y) < \infty, \tag{4.7}$$

where $\tilde{c} = \min(c, 0)$. We may then choose $T > 0$ so that the integral in (4.6) is finite. We thus conclude that for some $C > 0$ (depending on the value of the latter),

$$\mathrm{Ent}_\mu(f) \leqslant C\big(I(f) + 1\big).$$

By homogeneity, for every smooth enough f on $\mathbb{R}^n$:

$$\mathrm{Ent}_\mu\big(f^2\big) \leqslant C\bigg(\int |\nabla f|^2 \,\mathrm{d}\mu + \int f^2 \,\mathrm{d}\mu\bigg). \tag{4.8}$$

This is a defective logarithmic Sobolev inequality. One classical way to switch it into a true logarithmic Sobolev inequality (cf., e.g., [2]) is to establish first the Poincaré inequality for μ under the same condition (4.7). This can be achieved similarly on the basis the Harnack type inequality of [32] (cf. [1,18]). (With respect to Corollary 4.4, it should be emphasized for applications that the constant in (4.8), that depends on the value of the integral in (4.7), is highly dimensional.)

5. Transportation cost for the exponential measure

In this section, we apply the method of Section 3 to investigate the transportation cost inequality for the exponential measure first explored in [29]. To this task, we need to work with non-quadratic Hamilton–Jacobi equations.

5.1. Non-quadratic Hamilton–Jacobi equations

The general principle based on Hamilton–Jacobi equations can be extended to other cost functions than the square function. Let namely H be smooth and convex on $\mathbb{R}^n$ with $\lim_{|x|\to\infty} H(x)/|x| = +\infty$. For a smooth (Lipschitz e.g.) function f, the (unique viscosity) solution $u = u(x,t)$ of the minimization problem (cf. [3,16]):

$$\begin{aligned} &\frac{\partial u}{\partial t} + H(\nabla u) = 0 \quad \text{in } \mathbb{R}^n \times (0,\infty), \\ &u = f \quad \text{on } \mathbb{R}^n \times \{t = 0\}, \end{aligned} \tag{5.1}$$

is given by the Hopf–Lax formula

$$u(x,t) = \begin{cases} Q_t^L f(x) = \inf_{y\in\mathbb{R}^n}\left[f(y) + tL\left(\dfrac{x-y}{t}\right)\right], & t > 0,\ x \in \mathbb{R}^n, \\ f(x), & t = 0,\ x \in \mathbb{R}^n, \end{cases} \tag{5.2}$$

where L is the convex conjugate of H defined by:

$$L(y) = \sup_{x\in\mathbb{R}^n}\big[\langle x, y\rangle - H(x)\big].$$

For arbitrary cost, $Q_t^L f$ is not continuous in general at $t = 0$ even for smooth f.

Following the proof of Theorem 2.1, the derivative of $F(t) = \|e^{Q_t^L f}\|_{\lambda(t)}$ then leads to:

$$\lambda^2(t) F(t)^{\lambda(t)-1} F'(t) = \rho \operatorname{Ent}_\mu\big(e^{\lambda(t) Q_t^L f}\big) - \lambda^2(t) \int H\big(\nabla Q_t^L f\big) e^{\lambda(t) Q_t^L f} \, d\mu.$$

Useful applications of this principle however seem to require some homogeneity properties of H.

A first set of applications is obtained by replacing the Euclidean norm by arbitrary norms $\|\cdot\|$ on $\mathbb{R}^n$. Setting namely $L(y) = \frac{1}{2}\|y\|^2$, $y \in \mathbb{R}^n$, then, since H and L are self-dual, $H(x) = \frac{1}{2}\|x\|_*^2$, $x \in \mathbb{R}^n$, where $\|\cdot\|_*$ is the dual norm of $\|\cdot\|$. Therefore, under the logarithmic Sobolev inequality

$$\rho \operatorname{Ent}_\mu\big(f^2\big) \leqslant 2 \int \|\nabla f\|_*^2 \, d\mu \tag{5.3}$$

holding for some $\rho > 0$ and all smooth enough functions f on $\mathbb{R}^n$, we may conclude as in Theorem 2.1 to the hypercontractive estimates

$$\big\|e^{Q_t^L f}\big\|_{a+\rho t} \leqslant \big\|e^f\big\|_a$$

for every, say bounded f, $t \geqslant 0$ and $a \in \mathbb{R}$. In particular,

$$\int e^{\rho Q^L f} \, d\mu \leqslant e^{\rho \int f \, d\mu}$$

and, in its equivalent transportation cost form:

$$\rho W_L^2(\mu, \nu) \leqslant H(\nu|\mu).$$

Here

$$W_L^2(\mu,\nu) = \inf \iint \frac{1}{2}\|x-y\|^2 \, \mathrm{d}\pi(x,y),$$

where the infimum is running over all probability measures π on the product space $\mathbb{R}^n \times \mathbb{R}^n$ with marginals μ and ν. One may also consider more generally p-convex, $p \geqslant 2$, potentials (cf. [9]).

5.2. Modified logarithmic Sobolev inequalities

Another important example in the setting of Section 5.1 is the logarithmic Sobolev inequality for the exponential measure [8] that will lead, via this principle, to the transportation cost inequality of M. Talagrand [29] for the exponential measure. Recall from [8] that whenever μ is the measure on the real line with density $\frac{1}{2}\mathrm{e}^{-|x|}$ with respect to Lebesgue measure, for every Lipschitz function f on $\mathbb{R}$ such that $|f'| \leqslant c < 1$ almost everywhere:

$$\mathrm{Ent}_\mu\left(\mathrm{e}^f\right) \leqslant \frac{2}{1-c} \int f'^2 \mathrm{e}^f \, \mathrm{d}\mu. \tag{5.4}$$

Fix for simplicity $c = 1/2$. Set now

$$H(x) = \begin{cases} 4x^2 & \text{if } |x| \leqslant \dfrac{1}{2}, \\ +\infty & \text{if } |x| > \dfrac{1}{2}. \end{cases}$$

Its dual function is given by:

$$L(y) = \begin{cases} \dfrac{y^2}{16} & \text{if } |y| \leqslant 4, \\ \dfrac{|y|}{2} - 1 & \text{if } |y| > 4. \end{cases}$$

One may rewrite (5.4) as

$$\mathrm{Ent}_\mu\left(\mathrm{e}^f\right) \leqslant \int H(f')\mathrm{e}^f \, \mathrm{d}\mu. \tag{5.5}$$

Note that $H(\lambda x) \leqslant \lambda^2 H(x)$ whenever $|\lambda| \leqslant 1$.

Although H does not exactly fit all the hypotheses of the classical Hamilton–Jacobi theory, one may however check that $(Q_t^L f)'$ is (almost everywhere) in the domain of H (i.e. $|x| \leqslant 1/2$). We may then argue as in Section 2. Since we cannot expect however for a characterization through some kind of hypercontractivity (due to the lack of homogeneity of H), it is actually more simple to adapt the Herbst argument of Section 3. Namely, given a bounded (Lipschitz) function f, one first shows that $Q_t^L f$ is differentiable in $t > 0$ and almost every $x \in \mathbb{R}^n$ and that:

$$\frac{\partial}{\partial t} Q_t^L f + H\left(\left(Q_t^L f\right)'\right) = 0.$$

Set $F(t) = \int \mathrm{e}^{t Q_t^L f} \, \mathrm{d}\mu$ which is differentiable in $t > 0$. By (5.5),

$$tF'(t) \leqslant F(t) \log F(t), \quad 0 < t \leqslant 1.$$

While $Q_t^L f$ is not continuous at $t = 0$, it is easy to check however that $tQ_t^L f \to 0$ as $t \to 0$. Therefore $F'(0) \leqslant \int f \,\mathrm{d}\mu$, and integrating the preceding differential inequality as in the previous section, one concludes that:

$$\int \mathrm{e}^{Q^L f} \,\mathrm{d}\mu \leqslant \mathrm{e}^{\int f \,\mathrm{d}\mu}, \tag{5.6}$$

where $Q^L = Q_1^L$. The latter inequality (5.6) actually corresponds exactly to the transportation cost inequality for the exponential measure put forward in [29]. Namely, by the dual Monge–Kantorovitch principle (cf. [26]), (5.6) is equivalent to saying that, for every probability measure ν on the real line absolutely continuous with respect to μ

$$W_L(\mu, \nu) \leqslant H(\nu|\mu) \tag{5.7}$$

with

$$W_L(\mu, \nu) = \inf \iint L(x - y) \,\mathrm{d}\pi(x, y),$$

where the infimum is running over all probability measures on $\mathbb{R} \times \mathbb{R}$ with respective (integrable) marginals μ and ν. It is then easy to check that the cost L is equivalent, up to numerical constants, to the cost used in [29].

The preceding extends to products of the exponential distribution by considering the functions on $\mathbb{R}^n$ given by $\sum_{i=1}^n H(x_i)$ and $\sum_{i=1}^n L(x_i)$ for a vector $(x_1, \ldots, x_n) \in \mathbb{R}^n$. To this task, one may either tensorize the logarithmic Sobolev inequality (5.4) or the transportation inequality (5.7). As in [29], the main difficulty arises in dimension one.

5.3. Poincaré inequalities and exponential transportation cost

As a main result of the work [8], it was actually shown that every measure μ (absolutely continuous say) satisfying the Poincaré inequality

$$\lambda \operatorname{Var}_\mu(f) \leqslant \int |\nabla f|^2 \,\mathrm{d}\mu \tag{5.8}$$

for some $\lambda > 0$ and all smooth functions f actually satisfies a modified logarithmic Sobolev inequality such as (5.4):

$$\operatorname{Ent}_\mu\big(\mathrm{e}^f\big) \leqslant K(c) \int |\nabla f|^2 \mathrm{e}^f \,\mathrm{d}\mu \tag{5.9}$$

for every bounded Lipschitz function f such that $|\nabla f| \leqslant c < 2\sqrt{\lambda}$ almost everywhere, where $K(c) > 0$ only depends on c and λ. Setting:

$$H(x) = H_c(x) = \begin{cases} K(c)|x|^2 & \text{if } |x| \leqslant c, \\ +\infty & \text{if } |x| > c, \end{cases}$$

with dual function

$$L(y) = L_c(y) = \begin{cases} \dfrac{|y|^2}{4K(c)} & \text{if } |y| \leqslant 2cK(c), \\ c|y| - c^2 K(c) & \text{if } |y| > 2cK(c), \end{cases} \tag{5.10}$$

and arguing exactly as before, we may state the following corollary.

COROLLARY 5.1. – *Let μ be a measure on the Borel sets of $\mathbb{R}^n$ satisfying the Poincaré inequality*

$$\lambda\,\mathrm{Var}_\mu(f) \leqslant \int |\nabla f|^2\,\mathrm{d}\mu$$

for some $\lambda > 0$ and all smooth functions f. Then, for every $c < 2\sqrt{\lambda}$, μ satisfy the transportation cost inequality

$$W_L(\mu,\nu) \leqslant H(\nu|\mu) \tag{5.11}$$

for every probability measure $\nu \ll \mu$ where $L = L_c$ is the cost function (5.10). *In addition, all the inequalities* (5.8), (5.9) *and* (5.11) *are equivalent* (*up to constants*).

The last assertion of Corollary 5.1 simply follows from the fact that the transportation inequality (5.11) implies back the Poincaré inequality. Namely, for f smooth with compact support (say) and $t \to 0$, it is easy to see that the infimum $\inf_{y\in\mathbb{R}^n}[tf(y) + L(x-y)]$ is attained at some $y_0 = y_0(t) \to x$ as $t \to 0$. It follows that:

$$Q^L(tf)(x) \sim tf(x) - K(c)t^2\left|\nabla f(x)\right|^2.$$

Applying the transportation inequality (5.6) to tf and letting $t \to 0$ then shows, as in the introduction for the quadratic cost, that the Poincaré inequality (5.8) holds with $\lambda = \frac{1}{2K(c)}$. It should be pointed out that sharp constants carry over this procedure. Namely, it is shown in [8] that $K(c)$ may be chosen to satisfy:

$$K(c) = \frac{1}{2\lambda}\left(\frac{2\sqrt{\lambda}+c}{2\sqrt{\lambda}-c}\right)^2 \mathrm{e}^{c\sqrt{5/\lambda}}.$$

As $c \to 0$, $K(c) \to \frac{1}{2\lambda}$.

See also [5] for an approach based on optimal transportation and the Brenier–McCann theorem extending Talagrand's method for the Gaussian and exponential measures [29]. Applications to concentration properties are lengthly discussed in [8] and [18].

6. Brunn–Minkowski inequalities and logarithmic isoperimetry

In this final section, we present some further applications of the preceding results. We first describe exponential integrability of convex functions under a logarithmic Sobolev inequality. We then present another approach to the Bakry–Emery criterion through Brunn–Minkowski inequalities and our hypercontractivity result in Theorem 2.1. We finally discuss some analogues for L^1 logarithmic inequalities.

6.1. Exponential integrability of convex functionals

We start by elementary consequences of the transportation inequality:

$$\int \mathrm{e}^{\rho Qf}\,\mathrm{d}\mu \leqslant \mathrm{e}^{\rho\int f\,\mathrm{d}\mu} \tag{6.1}$$

for every bounded measurable f (where we write Q for Q_1) that corresponds to the critical value $a = 0$ in Theorem 2.1. Equivalently:

$$\int e^{\rho f}\,d\mu \leqslant e^{\rho \int \widetilde{Q} f\,d\mu} \tag{6.2}$$

(where we write $\widetilde{Q}$ for $\widetilde{Q}_1$). These inequalities can easily be extended from the class of all bounded measurable functions to the class of all μ-integrable functions f in (6.1) and the class of all measurable functions f in (6.2) with μ-integrable sup-convolution.

The operator Q_t represents a bijection from the class of all concave functions on $\mathbb{R}^n$ with values in $[-\infty, +\infty)$ onto itself. Respectively, $\widetilde{Q}_t$ is a bijection on the class of all convex functions on $\mathbb{R}^n$ with values in $(-\infty, +\infty]$. In particular, if we start with a homogeneous convex function

$$f(x) = \sup_{\theta\in T}\langle\theta, x\rangle, \quad x \in \mathbb{R}^n, T \subset \mathbb{R}^n,$$

then

$$\widetilde{Q}^{-1} f(x) = \sup_{\theta\in T}\left[\langle\theta, x\rangle - \frac{1}{2}|\theta|^2\right].$$

The supremum-convolution inequality (6.2) then yields (after a simple approximation argument)

$$\int e^{\rho \sup_\theta[\langle\theta,x\rangle - |\theta|^2/2]}\,d\mu \leqslant e^{\rho\int \sup_\theta\langle\theta,x\rangle\,d\mu}. \tag{6.3}$$

For the canonical Gaussian measure on $\mathbb{R}^n$, this inequality was discovered by B.S. Tsirel'son [30] in connection with Gaussian mixed volumes. In the general setting of logarithmic Sobolev inequalities and non-homogeneous convex functions it may be formulated in the following way.

COROLLARY 6.1. – *Under the logarithmic Sobolev inequality* (2.3) *of Theorem* 2.1*, for any convex μ-integrable function f on $\mathbb{R}^n$*:

$$\int e^{\rho(f - \frac{1}{2}|\nabla f|^2)}\,d\mu \leqslant e^{\rho\int f\,d\mu}.$$

For the proof, since f is differentiable almost everywhere, for every point $x \in \mathbb{R}^n$ at which f is differentiable, and all $z \in \mathbb{R}^n$, $f(x+z) \geqslant f(x) + \langle\nabla f(x), z\rangle$. Therefore:

$$Qf(x) \geqslant \inf_{z\in\mathbb{R}^n}\left[f(x) + \big\langle\nabla f(x), z\big\rangle - \frac{1}{2}|z|^2\right] = f(x) - \frac{1}{2}\big|\nabla f(x)\big|^2.$$

6.2. Brunn–Minskowski inequalities and hypercontractivity

Brunn–Minkowski inequalities may be used to prove the hypercontractive inequalities of Theorem 2.1 for some classes of measures with log-concave densities. Assume that $d\mu = e^{-U}\,dx$ where $U : \mathbb{R}^n \to \mathbb{R}$ is smooth and such that for some $c > 0$, uniformly in $x \in \mathbb{R}^n$:

$$\mathrm{Hess}(U)(x) \geqslant c\,\mathrm{Id}$$

in the sense of symmetric matrices. This condition is thus the Bakry–Emery criterion [4] (cf. [2,18]) under which the logarithmic Sobolev inequality for μ holds with $\rho = c$ as we have seen in Theorem 4.1. The classical Brunn–Minkowski inequality, in its functional form (see [14] for the historical developments of this result), may be used to provide a simple proof of the hypercontractive estimates of Theorem 2.1 (with $a = 1$), and thus of the logarithmic Sobolev inequality. Recall that, in its functional formulation, the Brunn–Minkowski theorem indicates that whenever $\alpha, \beta > 0$, $\alpha + \beta = 1$, and u, v, w are non-negative measurable functions on $\mathbb{R}^n$ such that for all $x, y \in \mathbb{R}^n$:

$$w(\alpha x + \beta y) \geqslant u(x)^\alpha v(y)^\beta, \tag{6.3}$$

then

$$\int w \,\mathrm{d}x \geqslant \left(\int u \,\mathrm{d}x\right)^\alpha \left(\int v \,\mathrm{d}x\right)^\beta. \tag{6.4}$$

Given a (bounded) function f on $\mathbb{R}^n$, apply then (6.4) to the functions:

$$u(x) = \mathrm{e}^{\frac{1}{\alpha} Q_{\beta/c\alpha} f(x) - U(x)}, \qquad v(y) = \mathrm{e}^{-U(y)}, \qquad w(z) = \mathrm{e}^{f(z) - U(z)}.$$

Due to the convexity condition $\mathrm{Hess}(U) \geqslant c\,\mathrm{Id}$, for every $\alpha, \beta > 0$, $\alpha + \beta = 1$ and $x, y \in \mathbb{R}^n$,

$$\alpha U(x) + \beta U(y) - U(\alpha x + \beta y) \geqslant \frac{c\alpha\beta}{2}|x - y|^2 \tag{6.5}$$

so that condition (6.3) will be satisfied by the very definition of the infimum-convolution $Q_{\beta/c\alpha} f$. Therefore,

$$\int \mathrm{e}^f \,\mathrm{d}\mu \geqslant \left(\int \mathrm{e}^{\frac{1}{\alpha} Q_{\beta/c\alpha} f} \,\mathrm{d}\mu\right)^\alpha.$$

Setting $1/\alpha = 1 + ct$, $t \geqslant 0$, immediately yields (2.4) with $\rho = c$ and $a = 1$. In particular the logarithmic Sobolev inequality for μ holds with $\rho = c$. The same arguments holds when considering an arbitrary norm in (6.5) to yield the logarithmic Sobolev inequality (6.3). We thus recover with the Hamilton–Jacobi approach the Bakry–Emery result (Theorem 4.1) as well as some of the main results of [9].

It was shown similarly in [7] and [9] how Brunn–Minkowski inequalities may be used to deduce directly the transportation cost inequalities of Section 3. See also [5] for further results. The recent Riemannian version of the functional Brunn–Minkowski inequality of [13] may be used to extend the preceding to a Riemannian setting and to recover in this way the logarithmic Sobolev inequality of D. Bakry and M. Emery [4] in manifolds with a strictly positive lower bound on the Ricci curvature.

It might be worthwhile mentioning that the alternate choice (used in particular in [7,9,22]) in the functional Brunn–Minkowski inequality of

$$u(x) = \mathrm{e}^{-\beta f(x) - U(x)}, \qquad v(y) = \mathrm{e}^{\alpha Q_{1/c} f(y) - U(y)}, \qquad w(z) = \mathrm{e}^{-U(z)},$$

leads to

$$\left(\int \mathrm{e}^{\alpha Q_{1/c} f} \,\mathrm{d}\mu\right)^{1/\alpha} \left(\int \mathrm{e}^{-\beta f} \,\mathrm{d}\mu\right)^{1/\beta} \leqslant 1. \tag{6.6}$$

As $\beta \to 0$, (6.6) only yields (2.4) with $a = 0$, that is the infimum convolution inequality (6.1) (with $\rho = c$). In the notation (1.11), (6.6) corresponds to the range $-1 \leqslant r \leqslant 0$. While to reach the logarithmic Sobolev inequality itself would require all r (negative) large enough, it is already interesting to point out that the value $r = 0$ (the infimum-convolution inequality (6.1)) is actually equivalent to the whole interval $-1 \leqslant r \leqslant 0$ (the inequalities (6.6)). To prove this claim, rewrite (6.6) as:

$$\frac{1}{\alpha}\log\int \mathrm{e}^{\alpha Q_{1/c}f}\,\mathrm{d}\mu + \frac{1}{\beta}\log\int \mathrm{e}^{-\beta f}\,\mathrm{d}\mu \leqslant 0 \tag{6.7}$$

for every $\alpha, \beta > 0$, $\alpha + \beta = 1$. Now,

$$\log\int \mathrm{e}^{g}\,\mathrm{d}\mu = \sup\left[\int gh\,\mathrm{d}\mu - \mathrm{Ent}_\mu(h)\right],$$

where the supremum is running over all bounded measurable functions $h \geqslant 0$ such that $\int h\,\mathrm{d}\mu = 1$. Thus we may further rewrite (6.7) as:

$$\int Q_{1/c}fh_1\,\mathrm{d}\mu - \int fh_2\,\mathrm{d}\mu \leqslant \frac{1}{\alpha}\mathrm{Ent}_\mu(h_1) + \frac{1}{\beta}\mathrm{Ent}_\mu(h_2),$$

$\alpha, \beta > 0$, $\alpha + \beta = 1$, that should therefore hold for all $h_1, h_2 \geqslant 0$ with $\int h_1\,\mathrm{d}\mu = \int h_2\,\mathrm{d}\mu = 1$. Optimizing over α and β we get:

$$\int Q_{1/c}fh_1\,\mathrm{d}\mu - \int fh_2\,\mathrm{d}\mu \leqslant \left(\sqrt{\mathrm{Ent}_\mu(h_1)} + \sqrt{\mathrm{Ent}_\mu(h_2)}\right)^2,$$

that is

$$\int Q_{1/c}f\,\mathrm{d}\nu_1 - \int f\,\mathrm{d}\nu_2 \leqslant \left(\sqrt{H(\nu_1|\mu)} + \sqrt{H(\nu_2|\mu)}\right)^2, \tag{6.8}$$

where $\mathrm{d}\nu_1 = h_1\,\mathrm{d}\mu$, $\mathrm{d}\nu_2 = h_2\,\mathrm{d}\mu$ are arbitrary probability measures on $\mathbb{R}^n$ absolutely continuous with respect to μ. These measures may also be assumed to have finite second moment. Now the supremum over all f's on the left-hand side of (6.8) is equal to $\frac{c}{2}W_2(\nu_1, \nu_2)^2$ so that (6.8) becomes

$$\sqrt{c}\,W_2(\nu_1, \nu_2) \leqslant \sqrt{H(\nu_1|\mu)} + \sqrt{H(\nu_2|\mu)}. \tag{6.9}$$

We thus reduced (6.6) to (6.9). But now the latter follows from (3.7) (with $\rho = c$) by the triangle inequality for the metric W_2. This proves the claim.

6.3. Logarithmic isoperimetry

In this last part, we turn some to L^1-versions of our hypercontractivity results. Let μ be a probability measure on the Borel sets of a metric space (E, d) and assume it satisfies the (logarithmic) isoperimetric inequality:

$$\mu^+(A) \geqslant c\bigl(1 - \mu(A)\bigr)\log\left(\frac{1}{1 - \mu(A)}\right) \tag{6.10}$$

for every Borel set A in E and some $c > 0$. Recall that in general the μ-perimeter $\mu^+(A)$ of a Borel set $A \subset E$ is defined by:

$$\mu^+(A) = \liminf_{t \to 0}\frac{1}{t}\bigl[\mu(A_t) - \mu(A)\bigr],$$

where A_t, $t > 0$, is the open t-neighborhood of A in the metric d on E.

The isoperimetric inequality (6.10) is connected with hypercontractivity of the convolution operators

$$Q_t f(x) = \inf_{y \in E; d(x,y) < t} f(y), \quad t > 0, x \in E.$$

As we will see indeed, (6.10) holds if and only if

$$\|Q_t f\|_q \leqslant \|f\|_p \tag{6.11}$$

for every non-negative measurable function f and all $0 < p < q < \infty$ and $t > 0$ such that $\mathrm{e}^{ct} \geqslant q/p$. To hint this connection, apply (6.11) to $f = \mathbf{1}_{E \setminus A}$. Since $Q_t f = \mathbf{1}_{E \setminus A_t}$, (6.11) turns into

$$\log\bigl(1 - \mu(A_t)\bigr) \leqslant \mathrm{e}^{ct} \log\bigl(1 - \mu(A)\bigr). \tag{6.12}$$

As $t \to 0$, this amounts to (6.10).

It should be noted that in "regular" situations one has $\mu^+(A) = \mu^+(M \setminus A)$. This is certainly the case for μ absolutely continuous on $E = \mathbb{R}^n$, as well as in a more general Riemannian manifold setting. In the latter cases, it was shown by O. Rothaus [27] that the isoperimetric inequality (6.10) is equivalent to the logarithmic Sobolev inequality

$$c\,\mathrm{Ent}_\mu(f) \leqslant \int |\nabla f| \,\mathrm{d}\mu \tag{6.13}$$

which should hold in the class of all non-negative locally Lipschitz function f on $\mathbb{R}^n$ (or on a manifold). Furthermore, the standard theory shows that (given a locally Lipschitz) function f on $\mathbb{R}^n$, the function $v = v(x,t) = Q_t f(x)$ provides a solution of the initial-value partial differential equation:

$$\begin{aligned} &\frac{\partial v}{\partial t} + |\nabla v| = 0 \quad \text{in } \mathbb{R}^n \times (0, \infty), \\ &v = f \quad \text{on } \mathbb{R}^n \times \{t = 0\}. \end{aligned} \tag{6.14}$$

The equivalence between (6.10) and (6.11) may then be proved on the basis of the partial differential equation (6.14) arguing as in the proof of our main result in Section 2. The particular structure of the L^1 case makes it however more general than equation (6.14) and the result actually holds in the setting of abstract metric spaces, with a purely "metric" proof.

THEOREM 6.2. – *Let μ be a probability measure on the Borel sets of a metric space (E, d). The probability measure μ satisfies the isoperimetric inequality*:

$$\mu^+(A) \geqslant c\bigl(1 - \mu(A)\bigr) \log\left(\frac{1}{1 - \mu(A)}\right)$$

for some $c > 0$ in the class of all Borel sets A in E if and only if

$$\|Q_t f\|_q \leqslant \|f\|_p$$

for every non-negative measurable function f on E and all $0 < p < q < \infty$ and $t > 0$ such that

$$\mathrm{e}^{ct} \geqslant \frac{q}{p}.$$

Proof. – We only need to show the sufficiency part. Since $(Q_t f)^p = Q_t f^p$, it is enough to deal with the case $p = 1$, and thus $q = \mathrm{e}^{ct} \geqslant 1$. The isoperimetric inequality (6.10) can be iterated in $t > 0$ so to yield (6.12) for every Borel A. Given a measurable function $f \geqslant 0$ on E, and $\lambda > 0$, set $A = \{f < \lambda\}$. By definition of Q_t, for every $t > 0$:

$$\{Q_t f < \lambda\} = A_t,$$

so that by (6.13), we get

$$\mu(Q_t f \geqslant \lambda) \leqslant \mu(f \geqslant \lambda)^q.$$

Hence

$$\|Q_t f\|_q^q = \int_0^\infty \mu(Q_t f \geqslant \lambda)\,\mathrm{d}\lambda^q \leqslant \int_0^\infty \mu(f \geqslant \lambda)^q\,\mathrm{d}\lambda^q.$$

Now it is know that the right-hand side of the latter inequality defines the so-called $\|f\|_{1,q}$ Lorentz norm of f, and that $\|f\|_{1,q} \leqslant \|f\|_1$ (cf. [28]). This stronger conclusion implies the result. □

A dual statement to Theorem 6.3 can be formulated with:

$$\widetilde{Q}_t f(x) = \sup_{y \in E;\, d(x,y) < t} f(y), \quad t > 0, x \in E.$$

Both inequalities (6.10) and (6.11) imply the logarithmic Sobolev inequality

$$\rho\,\mathrm{Ent}_\mu\big(f^2\big) \leqslant 2 \int |\nabla f|^2\,\mathrm{d}\mu$$

for some $\rho = \rho(c) > 0$ (cf. [27]).

It was shown in [6] that every log-concave measure μ on $\mathbb{R}^n$ supported by a ball of radius r satisfies the isoperimetric inequality (6.10) with $c = 1/2r$. In particular, the uniform distribution on a convex compact body $K \subset \mathbb{R}^n$ satisfies (6.10) with some $c > 0$. It would be of interest to estimate this constant in some special situations. For example, when K is the unit ball, the extremal sets in the isoperimetric problem are known. Another important case is the unit cube $K = [0,1]^n$. One may also consider the case of the sphere.

Acknowledgement

We sincerely thank C. Villani for explaining to us his results with F. Otto and for numerous helpful comments and remarks on the present work. We also thank F. Otto and C. Villani for their welcome addition [25] to this paper.

REFERENCES

[1] S. AIDA, Uniform positivity improving property, Sobolev inequalities and spectral gaps, *J. Funct. Anal.* 158 (1998) 152–185.

[2] D. BAKRY, L'hypercontractivité et son utilisation en théorie des semigroupes, in: *Ecole d'Eté de Probabilités de St-Flour*, Lecture Notes in Math., Vol. 1581, Springer, 1994, pp. 1–114.

[3] G. BARLES, *Solutions de Viscosité des Équations de Hamilton–Jacobi*, Springer, 1994.

[4] D. BAKRY, M. EMERY, Diffusions hypercontractives, in: *Séminaire de Probabilités XIX*, Lecture Notes in Math., Vol. 1123, Springer, 1985, pp. 177–206.

[5] G. BLOWER, The Gaussian isoperimetric inequality and transportation, Positivity (1999), to appear.

[6] S. BOBKOV, Isoperimetric and analytic inequalities for log-concave probability measures, *Ann. Probability* 27 (1999) 1903–1921.

[7] S. BOBKOV, F. GÖTZE, Exponential integrability and transportation cost related to logarithmic Sobolev inequalities, *J. Funct. Anal.* 163 (1999) 1–28.

[8] S. BOBKOV, M. LEDOUX, Poincaré's inequalities and Talagrand's concentration phenomenon for the exponential measure, *Probab. Theory Related Fields* 107 (1997) 383–400.

[9] S. BOBKOV, M. LEDOUX, From Brunn–Minkowski to Brascamp–Lieb and to logarithmic Sobolev inequalities, *Geom. Funct. Anal.* 10 (2000) 1028–1052.

[10] Y. BRENIER, Polar factorization and monotone rearrangement of vector-valued functions, *Comm. Pure Appl. Math.* 44 (1991) 375–417.

[11] H. BRÉZIS, *Opérateurs Maximaux Monotones*, North-Holland, 1973.

[12] D. CORDERO-ERAUSQUIN, Some applications of mass transport to Gaussian type inequalities, *Arch. Rational Mech. Anal.* (2000), to appear.

[13] D. CORDERO-ERAUSQUIN, R. MCCANN, M. SCHMUCKENSCHLÄGER, *A Riemannian Interpolation Inequality à la Borell*, Brascamp and Lieb, 2000, Invent. Math., to appear.

[14] S. DAS GUPTA, Brunn–Minkowski inequality and its aftermath, *J. Multivariate Anal.* 10 (1980) 296–318.

[15] R.M. DUDLEY, *Real Analysis and Probability*, Chapman & Hall, 1989.

[16] L.C. EVANS, *Partial Differential Equations*, Graduate Studies in Math., Vol. 19, Amer. Math. Soc., 1997.

[17] L. GROSS, Logarithmic Sobolev inequalities, *Amer. J. Math.* 97 (1975) 1061–1083.

[18] M. LEDOUX, Concentration of measure and logarithmic Sobolev inequalities, in: *Séminaire de Probabilités XXXIII*, Lecture Notes in Math., Vol. 1709, Springer, 1999, pp. 120–216.

[19] M. LEDOUX, The geometry of Markov diffusion generators, *Ann. Fac. Sci. Toulouse* (2000), to appear.

[20] R.J. MCCANN, Existence and uniqueness of monotone measure-preserving maps, *Duke Math J.* 80 (1995) 309–323.

[21] K. MARTON, A measure concentration inequality for contracting Markov chains, *Geom. Funct. Anal.* 6 (1997) 556–571.

[22] B. MAUREY, Some deviations inequalities, *Geom. Funct. Anal.* 1 (1991) 188–197.

[23] F. OTTO, The geometry of dissipative evolution equations: the porus medium equation, *Comm. Partial Differential Equations* (1999), to appear.

[24] F. OTTO, C. VILLANI, Generalization of an inequality by Talagrand, and links with the logarithmic Sobolev inequality, *J. Funct. Anal.* 173 (2000) 361–400.

[25] F. OTTO, C. VILLANI, J. Math. Pures Appl., this issue.

[26] S.T. RACHEV, *Probability Metrics and the Stability of Stochastic Models*, Wiley, 1991.

[27] O. ROTHAUS, Analytic inequalities, isoperimetric inequalities and logarithmic Sobolev inequalities, *J. Funct. Anal.* 64 (1985) 296–313.

[28] E. STEIN, G. WEISS, *Introduction to Fourier Analysis on Euclidean Spaces*, Princeton University Press, 1971.

[29] M. TALAGRAND, Transportation cost for Gaussian and other product measures, *Geom. Funct. Anal.* 6 (1996) 587–600.

[30] B.S. TSIREL'SON, A geometric approach to a maximum likelihood estimation for infinite-dimensional Gaussian location II, *Theor. Probab. Appl.* 30 (1985) 820–828.

[31] C. VILLANI, Topics in mass transport, Lecture Notes (2000).

[32] F.-Y. WANG, Logarithmic Sobolev inequalities on noncompact Riemannian manifolds, *Probab. Theory Related Fields* 109 (1997) 417–424.