Remarks on some transportation cost inequalities

Michel Ledoux

University of Toulouse, France

Abstract

We discuss some variations and improvements on certain families of transportation cost inequalities. Relying on two distinct interpolation schemes along the Kantorovich dual description of the Monge-Kantorovich metrics, we emphasize in particular an improved form of the Otto-Villani theorem, from a logarithmic Sobolev inequality to a quadratic transportation cost inequality, under only a polynomial decay of entropy along the flow. An analogous statement under a Poincaré inequality is also simplified. One illustration is addressed in the context of Coulomb transport inequalities.

1 Introduction

Given $p \geq 1$, the Monge-Kantorovich distance (cf. [20] e.g.) between two probability measures ν and μ on the Borel sets of a metric space (M, d) with a finite p-th moment is defined by

$$\mathrm{W}_p(\nu, \mu) \,=\, \bigg(\int_{M \times M} d(x,y)^p d\pi(x,y) \bigg)^{1/p}$$

where the infimum is taken over all couplings π on $M \times M$ with respective marginals ν and μ.

On the other hand, consider the Rényi-Tsallis entropies or divergences between two probability measures ν and μ on M with $\nu << \mu$ defined by

$$\mathrm{T}_\alpha\big(\nu \,|\, \mu\big) \,=\, \frac{1}{\alpha - 1} \bigg(\int_M f^\alpha d\mu - 1 \bigg)$$

where $f = \frac{d\nu}{d\mu}$ is the Radon-Nikodym derivative and $\alpha > 0$. The most important indices are $\alpha = \frac{1}{2}$ (Hellinger distance), $\alpha = 1$ (Kullback-Leibler distance or relative entropy)

$$\mathrm{T}_1\big(\nu \,|\, \mu\big) \,=\, \mathrm{H}\big(\nu \,|\, \mu\big) \,=\, \int_M f \log f \, d\mu,$$

and $\alpha = 2$ (quadratic Rényi-Tsallis divergence)

$$\mathrm{T}_2\big(\nu \,|\, \mu\big) \,=\, \mathrm{Var}_\mu(f).$$

The functional T_α is non-decreasing in α, so, for growing indices the distances are strengthening. In the range $0 < \alpha < 1$, all T_α are comparable to each other and are metrically equivalent to the total variation distance.

The transportation cost inequalities discussed in this note are inequalities comparing $\mathrm{W}_p(\nu, \mu)$ and $\mathrm{T}_\alpha(\nu \,|\, \mu)$ in the form of

$$\mathrm{W}_p^p(\nu, \mu) \,\leq\, C\, \mathrm{T}_\alpha\big(\nu \,|\, \mu\big)$$

for some constant $C > 0$ and all $\nu << \mu$.

We review a few known results, mostly refereing to [20, 2] as general references. In order to present them, it will be convenient to deal with a metric space arising from a (smooth, complete) Riemannian manifold (M, g), denoting by d the Riemannian distance and by dx the Riemannian volume element. Actually, to both deal with compact Riemannian manifolds, in which case the Riemannian volume element will be assumed to be normalized to a probability measure, and probability measures on $\mathbb{R}^n$, we will consider weighted probability measures $d\mu = \mathrm{e}^{-V} dx$ on (M, g), where $V : M \to \mathbb{R}$ is some smooth potential. A typical example is of course simply the standard Gaussian measure $d\mu(x) = \mathrm{e}^{-|x|^2/2} \frac{dx}{(2\pi)^{n/2}}$ on $\mathbb{R}^n$. If (M, g) is compact, we then still denote by μ the normalized Riemannian measure. An abstract framework covering these examples is the setting of Markov triples (E, μ, Γ) of [2] in which most of the conclusions here may be transferred.

On $\mathbb{R}^n$, for the standard Gaussian measure $d\mu(x) = \mathrm{e}^{-|x|^2/2} \frac{dx}{(2\pi)^{n/2}}$, the quadractic transportation cost inequality, known as Talagrand's inequality,

$$\mathrm{W}_2^2(\nu, \mu) \,\leq\, 2C\, \mathrm{T}_1\big(\nu \,|\, \mu\big) \tag{1}$$

holds true for every $\nu << \mu$, with $C = 1$. The inequality holds similarly, with $C = \frac{1}{\kappa}$, if $d\mu = e^{-V} dx$ with $V(x) - \frac{\kappa}{2}|x|^2$ convex for some $\kappa > 0$.

The preceding result is actually implied by the Otto-Villani theorem [17] ensuring that if μ satisfies a logarithmic Sobolev inequality in the sense that

$$\mathrm{T}_1\big(\nu \,|\, \mu\big) \,\leq\, \frac{C_{\mathrm{LS}}}{2}\, \mathrm{I}\big(\nu \,|\, \mu\big) \tag{2}$$

for every $\nu << \mu$, with $\mathrm{I}(\nu \,|\, \mu) = \int_M \frac{|\nabla f|^2}{f} d\mu$ the Fisher information, then (1) holds with constant $C = C_{\mathrm{LS}}$.

As another result in this framework, it has been shown recently by Y. Ding [8] that if μ satisfies a Poincaré inequality in the sense that

$$\mathrm{Var}_\mu(f) \,\leq\, C_{\mathrm{P}} \int_M |\nabla f|^2 d\mu \tag{3}$$

for some $C_{\mathrm{P}} > 0$ and every smooth f on M, then for every $\alpha > 1$ and every $\nu << \mu$,

$$\mathrm{W}_2^2(\nu, \mu) \,\leq\, C(\alpha, C_{\mathrm{P}})\, \mathrm{T}_\alpha\big(\nu \,|\, \mu\big) \tag{4}$$

where $C = C(\alpha, C_{\rm P}) > 0$. (For a prior result on $\mathbb{R}$ with $\alpha = 2$, see [13].) Since the logarithmic Sobolev inequality (2) implies the Poincaré inequality (with $C_{\rm P} = C_{\rm LS}$), this result might be thought of as an analogue of the Otto-Villani theorem at the level of the Poincaré inequality (3). It is not possible to reach $\alpha = 1$ in (4) under only (3) in general.

Conversely, if μ satisfies the preceding transportation cost inequality (4) for some $\alpha \geq 1$, $C > 0$ and all $\nu << \mu$, it satisfies a Poincaré inequality (3) (with constant $C_{\rm P} = \frac{\alpha C}{2}$). The claim may be deduced in various ways (cf. e.g. [17, 3, 20, 2]). In particular, it may be seen to follow from a standard result in optimal transport (cf. [17], [20, p. 588]) indicating that if $d\nu_\varepsilon = (1 + \varepsilon g) d\mu$ as $\varepsilon \to 0$, then

$$\lim_{\varepsilon \to 0} \frac{1}{\varepsilon^2} \, {\rm W}_2^2(\nu_\varepsilon, \mu) \, = \, \|g\|^2_{{\rm H}^{-1,2}(\mu)} \tag{5}$$

where

$$\|g\|_{{\rm H}^{-1,2}(\mu)} \, = \, \bigg(\int_M \big| \nabla((-{\rm L})^{-1} g) \big|^2 d\mu \bigg)^{1/2}$$

is the dual Sobolev norm (for $g : M \to \mathbb{R}$ with $\int_M g d\mu = 0$ in the suitable domain, and ${\rm L} = \Delta - \nabla V \cdot \nabla$ the diffusion operator with invariant measure μ – see Section 2). By a Taylor expansion on (4), it then follows that

$$\int_M \big| \nabla((-{\rm L})^{-1} g) \big|^2 d\mu \, \leq \, \frac{\alpha C}{2} \int_M g^2 d\mu$$

from which the assertion follows (by the formal change $g = (-{\rm L})^{1/2} h$).

The purpose of this work is to discuss some variations and improvements on the preceding statements, together with some simplified arguments. To this task, we rely on two distinct interpolation schemes, the linear interpolation and the classical semigroup interpolation, on the Kantorovich dual description of the metrics ${\rm W}_p$. The linear interpolation gives rise to bounds on the Monge-Kantorovich metrics by Sobolev-type norms which quantify the limit (5), already emphasized in [15] following [1]. This is presented in Section 2, while Section 3 illustrates the conclusion in the context of Coulomb transport inequalities, improving in a specific case a result of [6] and [10]. Using the semigroup interpolation, we provide, in Section 4, a stronger version of the Otto-Villani theorem under a polynomial, rather than exponential, decay of entropy along the heat flow. We also emphasize a direct simple proof of (4). The final Section 5 outlines yet another approach to the Otto-Villani theorem in the context of the Schrödinger variational principle following the recent development [7].

2 Linear interpolation

This section presents a simple linear interpolation argument in the study of transportation cost inequalities. It is mainly based on the Kantorovich dual description of the metric ${\rm W}_p(\nu, \mu)$ as

$${\rm W}_1(\nu, \mu) \, = \, \sup \bigg(\int_M \varphi \, d\nu - \int_M \varphi \, d\mu \bigg) \tag{6}$$

where the supremum runs over all 1-Lipschitz maps $\varphi : M \to \mathbb{R}$, and for $p > 1$,

$$\frac{1}{p}\,\mathrm{W}_p^p(\nu,\mu) \,=\, \sup\bigg(\int_M Q_1\varphi\, d\nu - \int_M \varphi\, d\mu\bigg) \tag{7}$$

where the supremum is taken over all bounded continuous functions $\varphi : M \to \mathbb{R}$ and where

$$Q_s\varphi(x) \,=\, \inf_{y\in M}\Big[\varphi(y) + \frac{d(x,y)^p}{ps^{p-1}}\Big], \quad s > 0,\ x \in M,$$

is the infimum-convolution Hopf-Lax semigroup. It is classical (cf. e.g. [9]) that $Q_s\varphi(x)$, $s > 0$, $x \in M$, solves the Halmiton-Jacobi equation

$$\frac{d}{ds}\,Q_s\varphi \,=\, -\frac{1}{q}\,|\nabla Q_s\varphi|^q. \tag{8}$$

in $(0,\infty)\times M$ with initial condition φ, where $\frac{1}{p}+\frac{1}{q}=1$.

The first observation below is a control of the Monge-Kantorovich metric $\mathrm{W}_p(\nu,\mu)$ by the $\mathrm{H}^{-1,p}$-Sobolev norm of the Radon-Nykodim derivative of ν with respect to μ. In the weighted Riemannian framework, recall the second order differential operator $\mathrm{L} = \Delta - \nabla V\cdot\nabla$ for which the integration by parts formula

$$\int_M \varphi(-\mathrm{L}\psi)d\mu \,=\, \int_M \nabla\varphi\cdot\nabla\psi\, d\mu \tag{9}$$

holds true for all smooth $\varphi,\psi : M \to \mathbb{R}$. Denote by $(P_t)_{t\ge 0}$ the Markov semigroup with infinitesimal generator L [2]. Formally the inverse $(-\mathrm{L})^{-1}$ of the non-negative operator $-\mathrm{L}$ may be described by

$$(-\mathrm{L})^{-1} \,=\, \int_0^\infty P_t\, dt$$

acting on mean zero functions in the suitable domain, a core of which being the set C_c^∞ of C^∞ compactly supported functions on M. Whenever the spectrum $\sigma(-\mathrm{L})$ of $-\mathrm{L}$ is discrete, $(-\mathrm{L})^{-1}$ can be spectrally represented on a suitable function f as

$$(-\mathrm{L})^{-1}f \,=\, \sum_{\lambda\in\sigma(-\mathrm{L})\setminus\{0\}} \frac{1}{\lambda}\, f_\lambda u_\lambda \tag{10}$$

where $(u_\lambda)_{\lambda\in\sigma(-\mathrm{L})}$ is an $\mathrm{L}^2(\mu)$ orthonormal basis of eigenvectors and $f_\lambda = \langle f, u_\lambda\rangle$. Such a picture occurs on a compact manifold for example. On $\mathbb{R}^n$ equipped with the standard Gaussian measure μ, the family of Hermite polynomials provides an orthonormal basis of eigenvectors of $\mathrm{L}^2(\mu)$ with eigenvalues $\lambda_k = k$, $k\in\mathbb{N}$ (counted with mutiplicity).

Define then, for every $p \ge 1$, the dual Sobolev norm $\mathrm{H}^{-1,p}(\mu)$ by

$$\|g\|_{\mathrm{H}^{-1,p}(\mu)} \,=\, \bigg(\int_M \big|\nabla((-\mathrm{L})^{-1}g)\big|^p\, d\mu\bigg)^{1/p}$$

for functions $g : M \to \mathbb{R}$ with $\int_M g d\mu = 0$ for which $\nabla((-\mathrm{L})^{-1}g)$ exists and belongs to $\mathrm{L}^p(\mu)$. In the particular case $p = 2$, the integration by parts formula (9) and the symmetry of $(P_t)_{t\geq 0}$ yield

$$\begin{aligned}\int_M \big|\nabla((-\mathrm{L})^{-1}g)\big|^2 d\mu \,&=\, \int_M g(-\mathrm{L})^{-1}g\, d\mu \\ &=\, \int_0^\infty \int_M gP_t g\, d\mu\, dt \\ &=\, 2\int_0^\infty \int_M (P_t g)^2 d\mu\, dt,\end{aligned} \tag{11}$$

and in particular a simpler description of the admissible functions g.

The following statement is an energy estimate on the Monge-Kantorovich distance $\mathrm{W}_p(\nu, \mu)$ between two probability measures ν and μ with ν absolutely continuous with respect to μ by the dual Sobolev norm $\mathrm{H}^{-1,p}(\mu)$ of the Radon-Nikodym density $f = \frac{d\nu}{d\mu}$. It has been emphasized earlier in [15], following [1]; the proof is reproduced here for completeness. An alternate independent proof, based on the Benamou-Brenier formula, is established in [18] (and presented in [19]).

Proposition 1. *For any $1 \leq p < \infty$, and for all $d\nu = f d\mu$ with $f - 1$ in the domain of the dual Sobolev norm $\mathrm{H}^{-1,p}(\mu)$,*

$$\mathrm{W}_p(\nu, \mu) \,\leq\, p\, \|f - 1\|_{\mathrm{H}^{-1,p}(\mu)}.$$

As explained in the introduction, when $p = 2$, Proposition 1 is closely related to Poincaré-type inequalities and their connection with transportation cost inequalities via the limit (5). In view of this asymptotics, the inequality of Proposition 1 is of the correct order for $p = 2$ up to a factor 4.

Proof. Let first $p > 1$. By a standard regularization procedure, it may be assumed that f is smooth and that $f > 0$. Set then $g = f - 1$ so that $g > -1$ and $\int_M g d\mu = 0$. Let $\theta : [0, 1] \to [0, 1]$ be increasing, smooth, with $\theta(0) = 0$ and $\theta(1) = 1$. For every bounded continuous $\varphi : M \to \mathbb{R}$, by the Hamilton-Jacobi equation (8),

$$\begin{aligned}\int_M Q_1\varphi\, d\nu - \int_M \varphi d\mu \,&=\, \int_0^1 \frac{d}{ds} \int_M \big(1 + \theta(s)g\big) Q_s\varphi\, d\mu\, ds \\ &=\, \int_0^1 \int_M \Big[\theta'(s)g\, Q_s\varphi - \big(1 + \theta(s)g\big)\frac{1}{q}|\nabla Q_s\varphi|^q\Big] d\mu\, ds \\ &=\, \int_0^1 \int_M \Big[-\theta'(s)\nabla\big((-\mathrm{L})^{-1}g\big)\cdot\nabla Q_s\varphi - \big(1 + \theta(s)g\big)\frac{1}{q}|\nabla Q_s\varphi|^q\Big] d\mu\, ds\end{aligned}$$

where we used integration by parts (9) in the last step.

By Young's inequality $a \cdot b \leq \frac{|a|^p}{p} + \frac{|b|^q}{q}$,

$$\int_M Q_1\varphi\, d\nu - \int_M \varphi\, d\mu \,\leq\, \frac{1}{p}\int_0^1 \theta'(s)^p \int_M \frac{|\nabla((-\mathrm{L})^{-1}g)|^p}{[1 + \theta(s)g]^{p-1}}\, d\mu\, ds$$

and since $g > -1$,

$$\int_M Q_1\varphi\, d\nu - \int_M \varphi\, d\mu \,\le\, \frac{1}{p}\int_0^1 \frac{\theta'(s)^p}{[1-\theta(s)]^{p-1}}\, ds \int_M \big|\nabla((-\mathrm{L})^{-1}g)\big|^p\, d\mu.$$

Therefore, by the Kantorovich duality formula (7),

$$\mathrm{W}_p^p(\nu,\mu) \,\le\, \int_0^1 \frac{\theta'(s)^p}{[1-\theta(s)]^{p-1}}\, ds \int_M \big|\nabla((-\mathrm{L})^{-1}g)\big|^p\, d\mu.$$

The optimal choice of θ is provided by $\theta(s) = 1-(1-s)^p$ for which $\int_0^1 \frac{\theta'(s)^p}{[1-\theta(s)]^{p-1}}\, ds = p^p$. The proof is thereby completed by the Kantorovich duality formula (7). The conclusion extends (or by a direct argument) to $p=1$. Proposition 1 is established. □

Note that under a Poincaré inequality (3) for μ, the semigroup $(P_t)_{t\ge 0}$ decays exponentially in $\mathrm{L}^2(\mu)$, that is, whenever $g : M \to \mathbb{R}$ is in $\mathrm{L}^2(\mu)$ with mean zero,

$$\int_M (P_t g)^2\, d\mu \,\le\, e^{-2t/C_\mathrm{P}} \int_M g^2 d\mu.$$

Hence, from Proposition 1 and (11), for every ν with $f = \frac{d\nu}{d\mu}$,

$$\mathrm{W}_2^2(\nu,\mu) \,\le\, 8\int_0^\infty \int_M [P_t f - 1]^2 d\mu \,\le\, 4\, C_\mathrm{P}\, \mathrm{Var}_\mu(f) \,=\, 4\, C_\mathrm{P}\, \mathrm{T}_2\big(\nu\,|\,\mu\big).$$

This is of course much improved by (4).

3 Coulomb transport inequality

In this section, we illustrate Proposition 1 of the preceding section in the context of Coulomb gas transport inequalities as recently emphasized in [6]. The context and notation are taken from this article.

The n-dimensional ($n \ge 2$) Coulomb kernel k is defined as usual by $x \in \mathbb{R}^n \mapsto k(x) = \frac{1}{|x|^{n-2}}$ if $n \ge 3$ and $x \in \mathbb{R}^n \mapsto k(x) = \log\frac{1}{|x|}$ if $n = 2$. Given a probability measure μ on the Borel sets of $\mathbb{R}^n$, its Coulomb energy, with values in $\mathbb{R} \cup \{+\infty\}$, is defined by

$$\mathcal{E}(\mu) \,=\, \int_{\mathbb{R}^n\times\mathbb{R}^n} k(x-y) d\mu(x) d\mu(y).$$

For any ν, μ, probability measures on $\mathbb{R}^n$, with compact support and finite Coulomb energy, the quantity $\mathcal{E}(\nu-\mu)$ is well defined, finite, non-negative, and vanishes if and only if $\nu = \mu$, and its square root defines a metric (on the subspace of such measures).

It has been shown in [6] that, for every compact subset $D \subset \mathbb{R}^n$, there exists a constant $C_D > 0$ such that for every probability measures ν, μ supported in D with finite Coulomb energy,

$$\mathrm{W}_1(\nu,\mu)^2 \,\le\, C_D\, \mathcal{E}(\nu-\mu). \tag{12}$$

As a consequence of Proposition 1, when μ is the uniform measure on a compact subset D in $\mathbb{R}^n$, the result may be improved to the W_2 metric.

Proposition 2. *If μ is the uniform normalized measure on a compact non-empty subset $D \subset \mathbb{R}^n$, for any probability measure ν supported in D,*

$$\mathrm{W}_2^2(\nu,\mu) \,\leq\, \frac{4|D|^2}{c_n}\,\mathcal{E}(\nu-\mu)$$

where $|D|$ is the volume of D and $c_n = 2\pi$ if $n=2$ and $c_n = n(n-2)|B|$, with B the unit ball in $\mathbb{R}^n$.

Proof. Assume that ν has a smooth density f with respect to μ, and set

$$U(x) \,=\, \int_{\mathbb{R}^n} k(x-y)g(y)d\mu(y), \quad x \in \mathbb{R}^n,$$

where $g = f-1$. The Poisson equation expresses that, in the sense of distributions, $\Delta k = -c_n\delta_0$. Hence $\Delta U = -\frac{c_n}{|D|}\,g$. Now Proposition 1 (with $p=q=2$) expresses that

$$\mathrm{W}_2^2(\nu,\mu) \,\leq\, 4\int_{\mathbb{R}^n} \big|\nabla\big((-\Delta)^{-1}g\big)\big|^2 d\mu.$$

As in [6], in order to freely use integration by parts in the proof of this inequality, it is worthwhile mentioning that for $\varphi : M \to \mathbb{R}$, non-negative, bounded and continuous, there exists $0 \leq \widetilde{\varphi} \leq \varphi$ with compact support such that $\widetilde{\varphi} = \varphi$ on D and

$$\int_M Q_1\varphi\, d\nu - \int_M \varphi\, d\mu \,=\, \int_M Q_1\widetilde{\varphi}\, d\nu - \int_M \widetilde{\varphi}\, d\mu.$$

Next, since $(-\Delta)^{-1}g = \frac{|D|}{c_n}\,U$, it follows that

$$\mathrm{W}_2^2(\nu,\mu) \,\leq\, \frac{4|D|^2}{c_n^2}\int_{\mathbb{R}^n} |\nabla U|^2 d\mu.$$

Finally, by integration by parts,

$$\int_{\mathbb{R}^n} |\nabla U|^2 dx \,=\, -\int_{\mathbb{R}^n} U\Delta U dx \,=\, \frac{c_n}{|D|}\int_{\mathbb{R}^n} U(x)g(x)dx \,=\, c_n\,\mathcal{E}(\eta).$$

This is the conclusion of the proposition when the density f is smooth. The technical smoothing properties to reach arbitrary distributions are developed in [6]. □

Conclusions similar to the ones developed in [6] have been recently considered on compact Riemannian manifods in [10]. Proposition 2 admits an analogue in this setting. Namely, whenever M is a compact Riemannian manifold, without boundary, we let as usual μ denote the normalized volume element. If $G : M \times M \to (-\infty, +\infty]$ is the Green's function for the Laplace-Beltrami operator Δ such that $\int_M G_x d\mu = 0$ for every $x \in M$, then $\Delta G_x = -\delta_x + 1$ in the distributional sense (here $G_x(y) = G(x,y)$, $y \in M$). If ρ is signed measure on M such that $\int_{M\times M} G(x,y) d|\rho|(x) d|\rho|(y) < \infty$, define

$$\mathcal{E}(\rho) \,=\, \int_{M\times M} G(x,y) d\rho(x) d\rho(y).$$

Arging exactly as for Proposition 2, it may then be established that for any $\nu << \mu$ such that $\mathcal{E}(\nu) < \infty$,

$$\mathrm{W}_2^2(\nu,\mu) \,\leq\, 4\,\mathcal{E}(\nu - \mu).$$

4 Semigroup interpolation

In this section, we replace the linear interpolation by a semigroup interpolation. This approach was developed by K. Kuwada [14] (see also [2]), and we simply refine here the argument by an additional parameter.

We start again from the Kantotovich dual description (7), with $p > 1$ and q the dual exponent. Given $d\nu = f d\mu$ with a smooth positive density f, for any (bounded continuous) $\varphi : M \to \mathbb{R}$,

$$\int_M Q_1\varphi \, d\nu - \int_M \varphi \, d\mu \,=\, -\int_0^\infty \frac{d}{dt} \int_M Q_{\lambda(t)}\varphi \, P_t f d\mu \, dt$$

where $\lambda(t) > 0$, λ smooth, decreasing, $\lambda(0) = 1$ and $\lambda(\infty) = 0$. Therefore

$$\begin{aligned}\int_M Q_1\varphi \, d\nu - \int_M \varphi \, d\mu \,&=\, \int_0^\infty \int_M \Big[\frac{\lambda'(t)}{q} |\nabla Q_{\lambda(t)}\varphi|^q \, P_t f - Q_{\lambda(t)}\varphi \, \mathrm{L} P_t f \Big] d\mu \, dt \\ &=\, \int_0^\infty \int_M \Big[\frac{\lambda'(t)}{q} |\nabla Q_{\lambda(t)}\varphi|^q \, P_t f + \nabla Q_{\lambda(t)}\varphi \cdot \nabla P_t f \Big] d\mu \, dt.\end{aligned}$$

By Young's inequality,

$$\int_M Q_1\varphi \, d\nu - \int_M \varphi \, d\mu \,\leq\, \int_0^\infty \frac{1}{-p\lambda'(t)} \int_M (P_t f)^{-1} |\nabla P_t f|^p \, d\mu \, dt.$$

As a consequence, with the notation $\mathrm{I}_p(t) = \int_{\mathbb{R}^d} (P_t f)^{-1} |\nabla P_t f|^p \, d\mu$, $t \geq 0$, for the p-Fisher information along the semigroup,

$$\mathrm{W}_p^p(\nu,\mu) \,\leq\, \int_0^\infty \frac{\mathrm{I}_p(t)}{-\lambda'(t)} \, dt. \tag{13}$$

When $p = q = 2$, a first choice of interest is

$$\lambda(t) = \frac{1}{c} \int_t^\infty \sqrt{\mathrm{I}_2(s)} \, ds, \quad t \geq 0,$$

where $c = \int_0^\infty \sqrt{\mathrm{I}_2(s)} \, ds$ (assumed to be finite). Then $-\lambda'(t) = -\frac{1}{c} \sqrt{\mathrm{I}_2(t)}$ so that

$$\mathrm{W}_2^2(\nu,\mu) \,\leq\, c \int_0^\infty \sqrt{\mathrm{I}_2(t)} \, dt,$$

that is

$$\mathrm{W}_2(\nu,\mu) \;\le\; \int_0^\infty \sqrt{\mathrm{I}_2(t)}\, dt.$$

We recover in this way an inequality emphasized in [17, 20].

To illustrate other possibilities, still with $p = q = 2$, let

$$\mathrm{T}_1(t) \;=\; \int_M P_t f \log P_t f \, d\mu, \quad t \ge 0,$$

be the relative entropy along the semigroup $(P_t)_{t\ge 0}$. Since $\mathrm{T}_1'(t) = -\mathrm{I}_2(t)$, by integration by parts,

$$\mathrm{W}_2^2(\nu,\mu) \;\le\; \frac{\mathrm{T}_1(0)}{-\lambda'(0)} + \int_0^\infty \frac{\lambda''(t)}{\lambda'(t)^2}\, \mathrm{T}_1(t) dt. \tag{14}$$

It remains to appropriately choose the function λ. A first instance of interest is the Otto-Villani theorem. A logarithmic Sobolev inequality (2) for μ, with constant $C_{\mathrm{LS}} > 0$, is equivalent to the exponential decay of entropy

$$\mathrm{T}_1(t) \le e^{-2t/C_{\mathrm{LS}}}\, \mathrm{T}_1(0), \quad t \ge 0, \tag{15}$$

along the flow ([2]). Choose then $\lambda(t) = e^{-t/C_{\mathrm{LS}}}$ to get that

$$\mathrm{W}_2^2(\nu,\mu) \;\le\; 2C_{\mathrm{LS}}\, \mathrm{T}_1(0) \;=\; 2C_{\mathrm{LS}}\, \mathrm{T}_1\big(\nu \,|\, \mu\big)$$

which amounts to the quadratic transportation cost inequality (1). We therefore recover here the Otto-Villani theorem from a logarithmic Sobolev inequality to a quadratic transportation cost inequality.

But actually, rather than the exponential decay (15), only a relatively soft, polynomial, decay of entropy along the flow allows for such a quadratic transportation cost inequality. We have for example the following result.

Proposition 3. *In the preceding notation, given $d\nu = f d\mu$, assume that*

$$\mathrm{T}_1(t) \;\le\; \frac{C}{(1+t)^\beta}\, \mathrm{T}_1(0), \quad t \ge 0,$$

for some $\beta > 1$ and $C > 0$. Then

$$\mathrm{W}_2^2(\nu,\mu) \;\le C'\, \mathrm{T}_1\big(\nu \,|\, \mu\big)$$

for some $C' > 0$ (only depending on C and β).

The proposition follows from the choice of $\lambda(t) = (1+t)^{-\beta}$, $t \ge 0$, with $0 < \beta < \alpha - 1$. It should be mentioned that, under a curvature lower bound (in the sense of weighted manifolds [2]), P. Cattiaux and A. Guillin showed in [5, Lemma 4.10] that any decay of entropy yields the transportation cost inequality, actually the logarithmic Sobolev inequality itself. One may wonder for the minimal decay of entropy ensuring the quadratic transportation cost inequality.

Let us now use (14) in a somewhat different direction. Since $\mathrm{T}_1 \leq \mathrm{T}_\alpha$, $\alpha \geq 1$,

$$\mathrm{W}_2^2(\nu, \mu) \,\leq\, \frac{\mathrm{T}_\alpha(0)}{-\lambda'(0)} + \int_0^\infty \frac{\lambda''(t)}{\lambda'(t)^2}\, \mathrm{T}_\alpha(t) dt$$

with the corresponding notation $\mathrm{T}_\alpha(t) = \mathrm{T}_\alpha(P_t f)$, $t \geq 0$. It holds that

$$\frac{d}{dt}\, \mathrm{T}_\alpha(t) \,=\, -\alpha \int_M (P_t f)^{\alpha-2} |\nabla P_t f|^2 d\mu.$$

Consider now, for $\alpha \in [1, 2]$, the Beckner inequality for μ,

$$\frac{1}{1-\frac{1}{\alpha}} \bigg[\int_M h^2 d\mu - \Big(\int_M |h|^{2/\alpha} d\mu \Big)^\alpha \bigg] \,\leq\, C \int_M |\nabla h|^2 d\mu \tag{16}$$

for some $C > 0$ and all smooth $h : M \to \mathbb{R}$. Note that if μ satisfies a logarithmic Sobolev inequality, then it satisfies all the Beckner inequalities, and conversely if it satisfies a Poincaré inequality, it satisfies the Beckner inequalities for $\alpha \in (1, 2]$ (cf. [2]).

Assume then that μ satisfies the Beckner inequality (16) for $\alpha \in (1, 2]$. As explained in [5] and is easy to check, by Jensen's inequality, the application of (16) to $h = (P_t f)^{\alpha/2}$ implies that

$$\mathrm{T}_\alpha(t) \,\leq\, -\frac{C}{4}\, \frac{d}{dt}\, \mathrm{T}_\alpha(t).$$

Therefore $t \mapsto e^{4t/C} \mathrm{T}_\alpha(t)$ is non-increasing, that is

$$\mathrm{T}_\alpha(t) \,\leq\, e^{-4t/C}\, \mathrm{T}_\alpha(0)$$

for every $t \geq 0$. Choose then $\lambda(t) = e^{-8t/C}$, $t \geq 0$, to reach the following statement.

Proposition 4. *Assume that μ satisfies the Beckner inequality* (16) *for $\alpha \in (1, 2]$. Then, for every $\nu << \mu$,*

$$\mathrm{W}_2^2(\nu, \mu) \,\leq\, \Big(1 + \frac{C}{8}\Big)\, \mathrm{T}_\alpha\big(\nu \,|\, \mu\big).$$

Since a Poincaré inequality implies a Beckner inequality for $\alpha \in (1, 2]$, the result amounts to the main conclusion of [8]. For the specific value $\alpha = 2$, alternate arguments have been developed recently in [16], producing in particular a sharper constant in the latter together with further characterizations of the Poincaré inequality.

5 Entropic transportation cost inequality

This short section is yet another approach to the Otto-Villani theorem via recent developments [11, 12, 7] on the Schrödinger problem. It is actually another way to interpret the reverse hypercontractivity argument already emphasized in [3]. The setting is the one of the preceding sections.

Let $\varepsilon > 0$ and set, for ν probability measure on M,

$$\frac{1}{2}\,\mathrm{W}_{2,\varepsilon}^2(\nu,\mu) \,=\, \sup\bigg(\int_M v_1^\varepsilon(\varphi)d\nu - \int_M \varphi\, d\mu\bigg)$$

where the supremum is over all bounded continuous $\varphi : M \to \mathbb{R}$ and where

$$v^\varepsilon \,=\, v_t^\varepsilon(\varphi) \,=\, -2\varepsilon \log P_{\varepsilon t}(e^{-\varphi/2\varepsilon}), \quad t > 0,$$

is the viscous Hamilton-Jacobi solution of $\partial_t v^\varepsilon + \frac{1}{2}\,|\nabla v^\varepsilon|^2 - \varepsilon \mathrm{L} v^\varepsilon \,=\, 0$ with initial condition φ. As $\varepsilon \to 0$, $v_t^\varepsilon \to Q_t\varphi$ (in a sense to be made precise).

We consider here an inequality of the form

$$\mathrm{W}_{2,\varepsilon}^2(\nu,\mu) \,\leq\, 2C'\,\mathrm{T}_1\big(\nu\,|\,\mu\big) \tag{17}$$

holding for any $\nu << \mu$. The Kantorovich duality transforms equivalently (cf. [4, 3, 2]) this inequality into

$$\int_M e^{\frac{1}{C'}v_1^\varepsilon(\varphi)}d\mu \,\leq\, e^{\frac{1}{C'}\int_M \varphi d\mu}$$

holding for any $\varphi : M \to \mathbb{R}$ (bounded continuous). This is achieved in the standard way with $\psi = \frac{1}{C'}[v_1^\varepsilon(\varphi) - \int_M \varphi d\mu]$ and the choice of $f = \frac{d\nu}{d\mu} = \frac{e^\psi}{\int_M e^\psi d\mu}$. (When $\varepsilon = 0$, this is the usual dual description of the quadratic transportation cost inequality.)

Next, with $\phi = e^{-\varphi/2\varepsilon}$, the latter amounts to

$$\|P_\varepsilon \phi\|_{-2\varepsilon/C'} \,\geq\, \|\phi\|_0 \,=\, e^{\int_M \log\phi d\mu} \tag{18}$$

holding for every non-negative ϕ.

Recall that the logarithmic Sobolev inequality for μ, with constant $C > 0$,

$$\mathrm{H}\big(\nu\,|\,\mu\big) \,\leq\, \frac{1}{2C}\int_M \frac{|\nabla f|^2}{f}\,d\mu$$

for every $d\nu = f d\mu$, is equivalent to hypercontractivity, and also to reverse hypercontractivity in the form of

$$\|P_t f\|_q \,\geq\, \|f\|_p$$

for every non-negative f and some (any) $-\infty < q < p < 1$ such that $e^{2t/C} = \frac{q-1}{p-1}$. In particular therefore, the inequality (17) holds with

$$C' \,=\, \frac{2\varepsilon}{e^{2\varepsilon/C} - 1} \,<\, C.$$

This is the constant put forward in [7]. It is optimal as checked on the example of the Ornstein-Uhlenbeck semigroup (with $C = 1$) and for the test function $\phi = e^{\lambda x - \lambda^2/2}$. It yields that

Now, as $\varepsilon \to 0$, the inequality amounts the quadratic transportation cost inequality (1), providing another approach to the Otto-Villani theorem.

References

[1] L. Ambrosio, F. Stra, D. Trevisan. A PDE approach to a 2-dimensional matching problem. *Probab. Theory Related Fields* 173, 433-478 (2019).

[2] D. Bakry, I. Gentil, M. Ledoux. *Analysis and geometry of Markov diffusion operators.* Grundlehren der mathematischen Wissenschaften 348. Springer (2014).

[3] S. Bobkov, I. Gentil, M. Ledoux. Hypercontractivity of Hamilton-Jacobi equations. *J. Math. Pures et Appliquées* 80, 669–696 (2001).

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

[5] P. Cattiaux, A. Guillin. On quadratic transportation cost inequalities *J. Math. Pures et Appliquées* 86, 342–361 (2006).

[6] D. Chafaï, A. Hardy, M. Maïda. Concentration for Coulomb gases and Coulomb transport inequalities. *J. Funct. Anal.* 275, 1447–1483 (2018).

[7] G. Conforti, L. Ripani. Around the entropic Talagrand inequality (2018).

[8] Y. Ding. A note on quadratic transportation and divergence inequality. *Statist. Probab. Letters* 100, 115–123 (2015).

[9] L. Evans. *Partial differential equations.* Graduate Studies in Mathematics 19. American Mathematical Society (1998).

[10] D. García-Zelada. Concentration for Coulomb gases on compact manifolds (2018).

[11] I. Gentil, C. Léonard, L. Ripani. About the analogy between optimal transport and minimal entropy. *Ann. Fac. Sci. Toulouse Math.* 26, 569–601 (2017).

[12] I. Gentil, C. Léonard, L. Ripani. Dynamical aspects of generalized Schrödinger problem via Otto calculus – A heuristic point of view (2018).

[13] B. Jourdain. Equivalence of the Poincaré inequality with a transport-chi-square inequality in dimension one. *Electron. Commun. Probab.* 17, 1–12 (2012).

[14] K. Kuwada. Duality on gradient estimates and Wasserstein controls. *J. Funct. Anal.* 258, 3758–3774 (2010).

[15] M. Ledoux. On optimal matching of Gaussian samples. *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI) 457, Veroyatnost' i Statistika.* 25, 226–264 (2017).

[16] Y. Liu. The Poincaré inequality and quadratic transportation-variance inequalities (2019).

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

[18] R. Peyre. Comparison between W_2 distance and $\dot{\mathrm{H}}^{-1}$ norm, and localization of Wasserstein distance (2016). To appear *ESAIM: Control, Optimisation and Calculus of Variations.*

[19] F. Santambrogio. *Optimal Transport for Applied Mathematicians.* Progress in Nonlinear Differential Equations and Their Applications. Birkhäuser (2015).

[20] C. Villani. *Optimal transport. Old and new.* Grundlehren der mathematischen Wissenschaften 338. Springer (2009).

Institut de Mathématiques de Toulouse
Université de Toulouse – Paul-Sabatier, F-31062 Toulouse, France
& Institut Universitaire de France
ledoux@math.univ-toulouse.fr