Dimensional contraction via Markov transportation distance

François Bolley, Ivan Gentil, Arnaud Guillin

▶ To cite this version:

Dimensional contraction via Markov transportation distance

François Bolley[*], Ivan Gentil[†] and Arnaud Guillin[‡]

April 23, 2014

Abstract

It is now well known that curvature conditions *à la* Bakry-Émery are equivalent to contraction properties of the heat semigroup with respect to the classical quadratic Wasserstein distance. However, this curvature condition may include a dimensional correction which up to now had not induced any strenghtening of this contraction. We first consider the simplest example of the Euclidean heat semigroup, and prove that indeed it is so. To consider the case of a general Markov semigroup, we introduce a new distance between probability measures, based on the semigroup, and adapted to it. We prove, in the setting of a compact Riemannian manifold, that this Markov transportation distance satisfies the same properties as the Wasserstein distance does in the specific case of the Euclidean heat semigroup, namely dimensional contraction properties and Evolution Variational Inequalities.

Key words: Diffusion equations, Wasserstein distance, Markov semigroups, Curvature-dimension bounds.

1 Introduction

Contraction properties of (Markov) semigroups are an important probabilistic and analytic tool: for instance they enable to study the existence of invariant probability measures, or the stability and long time behaviour of solutions to various linear (Fokker-Planck, kinetic Fokker-Planck, ...) or non linear (McKean-Vlasov, porous medium, Boltzmann, ...) partial differential equations. An important aspect is of course the distance in which we measure this contraction. Recent progress has shown that the Wasserstein distance is a particularly relevant and natural choice, in particular, but not only, for dynamics which have been interpreted as gradient flows for this distance (see for example [Ott01, CT05, CMV06, CGM08, BGM10, NPS11, BGG12] and the

[*]Ceremade, Umr Cnrs 7534, Université Paris-Dauphine, Place de Lattre de Tassigny, F-75775 Paris cedex 16. bolley@ceremade.dauphine.fr

[†]Institut Camille Jordan, Umr Cnrs 5208, Université Claude Bernard Lyon 1, 43 boulevard du 11 novembre 1918, F-69622 Villeurbanne cedex. gentil@math.univ-lyon1.fr

[‡]Institut Universitaire de France et Laboratoire de Mathématiques, Umr Cnrs 6620, Université Blaise Pascal, Avenue des Landais, F-63177 Aubière cedex. guillin@math.univ-bpclermont.fr

reference books [AGS08, Vil09]). Here and below the Wasserstein distance between two Borel probability measures ν and μ on a Polish metric space (E, d) is defined by

$$W_2(\mu, \nu) = \inf \left(\int d^2(x, y) d\pi(x, y) \right)^{1/2},$$

where the infimum runs over all probability measures π on $E \times E$ with marginals μ and ν. We refer again to [AGS08, Vil09] for a reference presentation of this distance and its interplay with the optimal transportation problem.

On the other hand, geometric properties of metric spaces are an important and vast topic with many diverse issues, and the Wasserstein distance has provided new insight on them, see [OW05, vRS05, Stu06, LV09, AGS12a, AGS12b]. A particularly relevant notion is the one of curvature which has recently attracted much attention. It turns out that it can be handled in terms of a contraction property in Wasserstein distance, as follows.

Let $(H_t)_{t \geqslant 0}$ denote the heat semigroup on a smooth and complete (and connex) Riemannian manifold (M, g): it solves the heat equation $\partial_t u = \Delta_g u$ where Δ_g is the Laplace-Beltrami operator on M. Let also μ be the Riemannian measure on (M, g) and d the associated Riemannian distance. Then a fundamental result, due to M. von Renesse and K.-T. Sturm in [vRS05], says that the Ricci curvature of the manifold is bounded from below by a constant $R \in \mathbb{R}$ if and only if

$$W_2(H_t f\mu, H_t g\mu) \leq e^{-Rt} W_2(f\mu, g\mu) \tag{1}$$

for any $t \geqslant 0$ and any probability densities f, g with respect to μ. Diverse proofs and generalizations of this contraction result are given in [OW05, Wan04, BGL13].

A crucial challenging problem now consists in understanding the role of the dimension in the contraction property in Wasserstein distance. Indeed curvature and dimension are jointly considered in the synthetic definition by Lott-Sturm-Villani [Stu06, LV09], contraction properties, gradient commutation type properties or the Bakry-Émery curvature-dimension condition. It is for instance well known that, given $R \in \mathbb{R}$ and $n \geqslant 1$, the $CD(R, n)$ curvature-dimension condition proposed by D. Bakry and M. Émery in [BÉ85], see section 4, is satisfied for the Laplace-Beltrami operator on an n-Riemannian manifold if the Ricci curvature of the manifold is uniformly bounded from below by R.

This has been very recently performed in the following two remarkable results, by deriving an upper bound on the distance $W_2(H_t f\mu, H_s g\mu)$ with two *different* times $s, t > 0$:

• The first result is due to K. Kuwada in [Kuw13] : the Ricci curvature of the n-dimensional manifold M is bounded from below by a constant $R \in \mathbb{R}$ if and only if

$$W_2^2(H_t f\mu, H_s g\mu) \leq A(s, t, R) W_2^2(f\mu, g\mu) + B(s, t, n, R), \tag{2}$$

for any $s, t > 0$ and any probability densities f, g with respect to μ, for appropriate functions $A, B \geqslant 0$. In the case $R = 0$ the bound simplifies into

$$W_2^2(H_t f\mu, H_s g\mu) \leq W_2^2(f\mu, g\mu) + 2n(\sqrt{t} - \sqrt{s})^2, \tag{3}$$

stated independently in [BGL13].

• The second result is due to M. Erbar, K. Kuwada and K.-T. Sturm [EKS13] : the Ricci curvature of the n-dimensional manifold M is bounded from below by a constant $R \in \mathbb{R}$ if and only if

$$s_{\frac{R}{n}} \left(\frac{1}{2} W_2(H_t f\mu, H_s g\mu) \right)^2 \leq e^{-R(t+s)} s_{\frac{R}{n}} \left(\frac{1}{2} W_2(f\mu, g\mu) \right)^2 + \frac{n}{R}(1 - e^{-R(s+t)}) \frac{(\sqrt{t} - \sqrt{s})^2}{2(t+s)}$$

for any $s, t > 0$ and any probability densities f, g with respect to μ. Here $s_r(x) = \sin(\sqrt{r}x)/\sqrt{r}$ if $r > 0$, $s_r(x) = \sinh(\sqrt{-r}x)/\sqrt{-r}$ if $r < 0$ and $s_0(x) = x$, hence recovering (3) for $R = 0$.

Observe that in these two results the dimension dependent additional term in the right-hand side is positive, and appears only when the two solutions are considered at different times s and t.

A first aim of this paper is to take the dimension into account and to improve inequality (1) for solutions considered at the *same* time. For instance in section 2 we prove that

$$W_2^2(H_t f dx, H_t g dx) \leq W_2^2(f dx, g dx) - \frac{2}{n} \int_0^t \left(\text{Ent}_{dx}(H_u f) - \text{Ent}_{dx}(H_u g) \right)^2 du \qquad (4)$$

for the heat semigroup on $\mathbb{R}^n$, any $t \geqslant 0$ and any probability densities f, g with respect to the Lebesgue measure dx; here $\text{Ent}_\mu(g) = \int g \log g d\mu$ is the entropy. This inequality improves on (1) since the Euclidean space $\mathbb{R}^n$ has null Ricci curvature and then satisfies (1) with $R = 0$. Let us observe that a dimensional contraction property in a Wasserstein distance with a modified cost was derived by Wang in [Wan11].

A second aim it to obtain dimensional contraction inequalities for more general Markov semigroups. For that purpose we will work with a new distance called *Markov transportation distance*, based on the generator of the semigroup, and adapted to it and to the Bakry-Émery curvature-dimension condition formulation. It is defined by a modification of the following dynamical formulation of the Wasserstein distance proposed by J.-D. Benamou and Y. Brenier in [BB00] : for any probability densities f and g with respect to the Lebesgue measure in $\mathbb{R}^n$,

$$W_2(f dx, g dx) = \inf \Big(\int_0^1 \int \frac{|w_s|^2}{\rho_s} dx ds \Big)^{1/2}$$

where the infimum runs over all paths $(\rho_s)_{s \in [0,1]}$ and vector fields $(w_s)_{s \in [0,1]}$ such that $\partial_s \rho_s + \nabla \cdot w_s = 0$, $\rho_0 = f$ and $\rho_1 = g$; here $\nabla \cdot$ stands for the divergence operator on $\mathbb{R}^n$. This dynamical approach is the starting point of the definition in [DNS09, DNS12] of generalized distances. Instead, one can consider the quantity

$$\inf \Big(\int_0^1 \int \frac{|\nabla h_s|^2}{\rho_s} dx ds \Big)^{1/2}$$

where the infimum runs over all paths $(\rho_s, h_s)_{s \in [0,1]}$ such that $\partial_s \rho_s + \nabla \cdot (\nabla h_s) = 0$, $\rho_0 = f$ and $\rho_1 = g$. This quantity is more adapted to the context of a general Markov semigroup, and gives us the way to define the Markov transportation distance. Given a Markov generator L on a space E, with carré du champ Γ and invariant measure μ, we let

$$T_2(f\mu, g\mu) = \inf \Big(\int_0^1 \int \frac{\Gamma(h_s)}{\rho_s} d\mu ds \Big)^{1/2}$$

for two probability densities f and g with respect to μ, under the constraints $\partial_s \rho_s + L h_s = 0, \rho_0 = f, \rho_1 = g$. In this abstract formulation, discrete and non-local operators can be studied in a similar way. A fundamental instance is that of $L = \Delta_g - \nabla V \cdot \nabla$ on a Riemannian manifold (E, g), with carré du champ $\Gamma(f) = |\nabla f|^2$ and invariant measure μ with density e^{-V}. This will be the main example in this article, and in this case $W_2(f\mu, g\mu) \leq T_2(f\mu, g\mu)$ since the infima defining the distances run over a smaller set for T_2 than for W_2.

The paper is organized as follows. In section 2 we show in a simple way how to reach the dimension dependent contraction property (4) in Wasserstein distance for the specific heat semigroup on $\mathbb{R}^n$. It will give a flavor of the results proved and the methods used below in the Markov transportation distance T_2 in our context of Markov semigroups $(P_t)_{t \geqslant 0}$ on a connected and compact Riemannian manifold.

This distance is properly defined in section 3, together with fundamental properties and examples. In particular we derive an Otto-Villani theorem for T_2 : a logarithmic Sobolev inequality implies a transportation Talagrand inequality.

Section 4 is devoted to our main application : the contraction property under the curvature-dimension condition $CD(R, n)$ on our semigroup. Under this condition we prove that

$$T_2^2(P_T f\mu, P_T g\mu) \leq e^{-2RT} T_2^2(f\mu, g\mu) - \frac{2}{n} \int_0^T e^{-2R(T-t)} (\mathrm{Ent}_\mu(P_t g) - \mathrm{Ent}_\mu(P_t f))^2 dt,$$

for any $T > 0$ and any probability densities f, g with respect to the invariant measure μ.

In section 5 we briefly consider the so-called evolution variational inequalities (EVI in short). These inequalities say that if the Ricci curvature of a manifold is bounded from below by a constant $R \in \mathbb{R}$, then

$$W_2^2(f\mu, H_t g\mu) - W_2^2(f\mu, g\mu) \leq -\frac{e^{-2Rt} - 1 + 2Rt}{2Rt} W_2^2(f\mu, g\mu) + 2t(\mathrm{Ent}_\mu(f) - \mathrm{Ent}_\mu(H_t g)). \quad (5)$$

for the heat semigroup $(H_t)_{t \geqslant 0}$, any $t \geqslant 0$ and any probability densities f, g with respect to μ. This inequality characterizes $(H_t)_{t \geqslant 0}$ as the gradient flow of the entropy with respect to the Wasserstein distance. This interpretation has been made by R. Jordan, D. Kinderlehrer and F. Otto in [JKO98], and has led to numerous developments, see in particular the seminal paper [OV00] and the huge contribution of [AGS08]. In section 5 we explain how to obtain a dimensional EVI for the Wasserstein distance and the Euclidean heat semigroup, and then for the Markov transportation distance and our Markov semigroup under a curvature-dimension condition.

In Section 6 we briefly investigate natural generalizations of the Markov transportation distance.

Many questions are left aside in this work, such as the general existence of geodesics, dual formulations and further equivalence between the obtained contraction and curvature conditions. The purpose of this work is rather to show the interest of the T_2 distance, and these questions will be further investigated elsewhere. We haven chosen to present the Markov transportation distance in the classical setting of a compact Riemannian manifold to properly prove the dimensional contraction inequality. We are convinced that the new distance can be defined in a general setting.

Since this work was completed, the second author [Gen13] has extended the bound (4) to the heat semigroup on an n-dimensional compact Riemannian manifold. Moreover, related results are studied in the work [AMS14] in preparation.

2 The heat equation on $\mathbb{R}^n$

This section is devoted to the simple derivation of a dimension dependent contraction property for the heat semigroup $(H_t)_{t \geqslant 0}$ on $\mathbb{R}^n$. It is defined by

$$H_t f(x) = \int_{\mathbb{R}^n} f(y) \frac{e^{-\frac{|x-y|^2}{4t}}}{(4\pi t)^{n/2}} \, dy$$

and is obtained as the solution of the heat equation $\partial_t u = \Delta u$; here Δ is the usual Laplace operator in $\mathbb{R}^n$.

For this semigroup, the bound (1) is classical with $R = 0$ and μ the Lebesgue measure dx on $\mathbb{R}^n$, and is optimal in the sense that equality holds for all t if g is obtained from f by a translation in $\mathbb{R}^n$. Let us see how to simply turn this classical bound into a more precise dimension dependent bound.

Following [DNS12], let $(R_t)_{t \geqslant 0}$ be the heat semigroup acting on $\mathbb{R}^n$-valued maps, coordinate by coordinate. It satisfies

$$H_t(\nabla \cdot w) = \nabla \cdot (R_t w) \tag{6}$$

for all $\mathbb{R}^n$-valued functions w. This semigroup acting on vectors will be the main tool in our derivation. We omit regularity issues which are carefully considered in [DNS12].

As recalled in the introduction, the Benamou-Brenier Theorem ensures that

$$W_2^2(f dx, g dx) = \inf \int_0^1 \int \frac{|w_s|^2}{\rho_s} ds dx \tag{7}$$

for any probability measures $f dx$ and $g dx$ in $\mathbb{R}^n$; here the infimum runs over all couples $(\rho_s, w_s)_{s \in [0,1]}$ such that

$$\partial_s \rho_s + \nabla \cdot w_s = 0 \tag{8}$$

where, for all $s \in [0,1]$, ρ_s is a probability density with respect to Lebesgue measure, $\rho_0 = f$ and $\rho_1 = g$.

Let now $(\rho_s, w_s)_{s \in [0,1]}$ interpolate the densities f and g with the constraint (8). Then $(H_t(\rho_s))_{s \in [0,1]}$ interpolates the densities $H_t f$ and $H_t g$ and, by (6), the couple $(H_t(\rho_s), R_t(w_s))_{s \in [0,1]}$ satisfies (8). Then, by (7),

$$W_2^2(H_T f dx, H_T g dx) \leq \int_0^1 \int \frac{|R_T(w_s)|^2}{H_T(\rho_s)} ds dx \tag{9}$$

for any $T \geqslant 0$. Moreover:

Lemma 2.1 *Let $F : \mathbb{R}^n \to \mathbb{R}^n$ and g a positive probability density with respect to the Lebesgue measure, with F and g smooth. Then, for all $T \geqslant 0$*

$$\int \frac{|R_T F|^2}{H_T g} dx \leq \int \frac{|F|^2}{g} dx - \frac{2}{n} \int_0^T \left(\int \frac{R_t F \cdot \nabla H_t g}{H_t g} dx \right)^2 dt.$$

Proof
◁ We let

$$\Lambda(t) = \int \frac{|R_t F|^2}{H_t g} dx$$

for $t \geqslant 0$ and prove that

$$\Lambda'(t) \leq -\frac{2}{n}\left(\int \frac{R_t F \cdot \nabla H_t g}{H_t g}dx\right)^2,$$

which will prove the lemma by time integration. Indeed

$$\Lambda'(t) = \int \left(2\frac{R_t F \cdot \Delta R_t F}{H_t g} - \frac{\Delta H_t g|R_t F|^2}{(H_t g)^2}\right)dx.$$

For notational simplicity, we let $\bar{F} = R_t F$, $\bar{g} = H_t g$ and then $\bar{G} = \log \bar{g}$. Since

$$0 = \int \Delta\left(\frac{|\bar{F}|^2}{\bar{g}}\right)dx = \int 2\nabla(|\bar{F}|^2)\cdot\nabla\left(\frac{1}{\bar{g}}\right) + \frac{1}{\bar{g}}\Delta(|\bar{F}|^2) + |\bar{F}|^2\Delta\left(\frac{1}{\bar{g}}\right)dx,$$

we obtain

$$
\begin{aligned}
\Lambda'(t) &= -\int \frac{2}{\bar{g}}\left(\frac{1}{2}\Delta|\bar{F}|^2 - \bar{F}\cdot\Delta\bar{F} + \nabla(|\bar{F}|^2)\nabla\bar{G} + |\bar{F}|^2|\nabla\bar{G}|^2\right)dx \\
&= -\int \frac{2}{\bar{g}}\sum_{1\leq i,j\leq n}\left(\partial_i\bar{F}_i + \bar{F}_i\partial_j\bar{G}\right)^2 dx \leq -\int \frac{2}{\bar{g}}\sum_{1\leq i\leq n}\left(\partial_i\bar{F}_i + \bar{F}_i\partial_i\bar{G}\right)^2 dx \\
&= -\frac{2}{n}\int \bar{g}\left(\sum_{1\leq i\leq n}\frac{\partial_i\bar{F}_i}{\bar{g}} + \frac{\bar{F}_i\partial_i\bar{G}}{\bar{g}}\right)^2 dx \\
&\leq -\frac{2}{n}\left(\sum_{1\leq i\leq n}\int\left(\partial_i\bar{F}_i + \frac{\bar{F}_i\partial_i\bar{g}}{\bar{g}}\right)dx\right)^2 = -\frac{2}{n}\left(\int \frac{\bar{F}\cdot\nabla\bar{g}}{\bar{g}}dx\right)^2
\end{aligned}
$$

by the Cauchy-Schwarz inequality, the Jensen inequality for the probability measure $\bar{g}\,dx$ and the relation $\int \sum_i \partial_i\bar{F}_i\,dx = 0$. $\triangleright$

Then, by Lemma 2.1 and the Cauchy-Schwarz inequality (with respect to the measure ds),

$$\int_0^1 \int \frac{|R_T(w_s)|^2}{H_T(\rho_s)}dsdx \leq \int_0^1 \int \frac{|w_s|^2}{\rho_s}dsdx - \frac{2}{n}\int_0^T\left(\int\int_0^1 \frac{R_t(w_s)\cdot\nabla H_t(\rho_s)}{H_t(\rho_s)}dxds\right)^2 dt.$$

Moreover the couple $(H_t(\rho_s), R_t(w_s))$ satisfies (8), so

$$\int \frac{R_t(w_s)\cdot\nabla H_t(\rho_s)}{H_t(\rho_s)}dx = \partial_s\int H_t(\rho_s)\log H_t(\rho_s)dx,$$

and then

$$\int\int_0^1 \frac{R_t(w_s)\cdot\nabla H_t(\rho_s)}{H_t(\rho_s)}dxds = \mathrm{Ent}_{dx}(H_t f) - \mathrm{Ent}_{dx}(H_t g).$$

Then inequality (9) leads to the following refined contraction inequality for the heat semigroup in $\mathbb{R}^n$:

Proposition 2.2 *Let $(H_t)_{t\geqslant 0}$ be the heat semigroup on $\mathbb{R}^n$. Then for any probability densities f and g in $\mathbb{R}^n$ such that $W_2(fdx, gdx) < \infty$, for any $T > 0$,*

$$W_2^2(H_T fdx, H_T gdx) \leq W_2^2(fdx, gdx) - \frac{2}{n}\int_0^T \left(\mathrm{Ent}_{dx}(H_t f) - \mathrm{Ent}_{dx}(H_t g)\right)^2 dt. \qquad (10)$$

6

Remark 2.3 *By comparison with* (3), *the dimension brings a negative correction term in the contraction property. The bound* (10) *is again an equality if g is obtained from f by translation. Finally, a Taylor expansion of* (10), *for T close to 0 and g close to f, for any given f, implies back the curvature dimension $CD(0, n)$ for the Laplace operator (see section 4 below for the precise definition of the curvature-dimension condition).*

Remark 2.4 *Note that this result not only gives a correction for equal times, but also for different times s, t in the spirit of [BGL13], [EKS13] or [Kuw13]. Let indeed $s \leq t$: then applying the contraction estimate* (10) *to $P_{t-s}f$ and g and then using* (3) *lead to*

$$W_2^2(H_t f dx, H_s g dx) \leq W_2^2(f dx, g dx) + n(t - s) - \frac{2}{n} \int_0^s \left(\mathrm{Ent}_{dx}(H_{t-s+u}f) - \mathrm{Ent}_{dx}(H_u g) \right)^2 du.$$

Our main goal is then the extension of this contraction result to general Markov semigroups satisfying a $CD(R, n)$ condition, which will be given in Theorem 4.5 : there the Markov transportation distance will prove to be an adapted and efficient tool.

3 The Markov transportation distance

3.1 Definition

Let (E, g) be a C^∞ compact connected Riemannian manifold with volume measure dx, and V be a smooth function on E with $\int e^{-V} dx = 1$. Let also μ be the Borel probability measure on E with density e^{-V}. Let $(P_t)_{t \geqslant 0}$ be the Markov semigroup on E with infinitesimal generator

$$L = \Delta_g - \nabla V \cdot \nabla$$

where Δ_g is the Laplace-Beltrami operator on E. The generator is defined on a dense subspace $\mathcal{D}(L)$ of $L^2(\mu)$, which includes the algebra $\mathcal{A}$ of (bounded) smooth functions (in our context, smooth means C^∞). This algebra is itself stable by L and P_t.

The semigroup is *reversible* with respect to μ, in the sense that

$$\int f \, P_t g \, d\mu = \int g \, P_t f \, d\mu \qquad \text{or equivalently} \qquad \int f \, Lg \, d\mu = \int g \, Lf \, d\mu \tag{11}$$

for all $f, g \in \mathcal{A}$ and $t \geqslant 0$,

Such a Markov semigroup admits a Markov probability kernel, that is for any function $f \in L^2(\mu)$, $t \geqslant 0$ and $x \in E$,

$$P_t f(x) = \int_{\mathbb{R}^n} f(y) \, p_t(x, dy).$$

The carré du champ operator, defined on functions $f, g \in \mathcal{A}$ by the general expression

$$\Gamma(f, g) = \frac{1}{2} \Big(L(fg) - f \, Lg - g \, Lf \Big) \in \mathcal{A},$$

satisfies $\Gamma(f, f) = |\nabla f|^2$ where $|\nabla f|$ stands for the length of the vector ∇f; for simplicity we shall let $\Gamma(f) = \Gamma(f, f)$. The Dirichlet form $\mathcal{E}_\mu(f) = \int \Gamma(f) d\mu$ is defined on its domain $\mathcal{D}(\mathcal{E}_\mu) \subset L^2(\mu)$, which also includes $\mathcal{A}$.

The generator satisfies the following so-called *diffusion* property :

$$L\Phi(g) = \Phi'(g)Lg + \Phi''(g)\Gamma(g) \tag{12}$$

for any smooth function Φ and any function $g \in \mathcal{A}$. In particular $\Gamma(\Phi'(g)) = \Phi''^2(g)\Gamma(g)$.

We refer to [BGL14] for further details on these notions, and for their definition in a more general setting (there, the triple (E, Γ, μ) is called a compact Markov triple).

Before giving the definition of the Markov transportation distance, we need to define the paths between probability densities. We let $\mathcal{F}$ be the set of positive probability densities (with respect to μ) in $\mathcal{A}$.

Definition 3.1 *For a couple $(\rho_s, h_s)_{s \in [0,1]}$ of smooth functions on $[0,1] \times E$ with ρ_s in $\mathcal{F}$ we define*

$$\varphi(\rho_s, h_s) = \int \frac{\Gamma(h_s)}{\rho_s} d\mu \in [0, \infty)$$

for $s \in [0,1]$, and the action

$$\Phi(\rho, h) = \int_0^1 \varphi(\rho_s, h_s) ds = \int_0^1 \int \frac{\Gamma(h_s)}{\rho_s} d\mu\, ds \in [0, \infty].$$

For f, g in $\mathcal{F}$ we call admissible path between f and g such a couple $(\rho_s, h_s)_{s \in [0,1]}$ for which moreover

$$\partial_s \rho_s + L h_s = 0, \tag{13}$$

$\rho_0 = f$ and $\rho_1 = g$. We let $\mathcal{A}(f, g)$ be the set of admissible paths between f and g.

Definition 3.2 *The Markov transportation distance is defined for $f, g \in \mathcal{F}$ by*

$$T_2(f\mu, g\mu) = \inf \left(\int_0^1 \int \frac{\Gamma(h_s)}{\rho_s} d\mu\, ds \right)^{1/2},$$

where the infimum runs over all admissible paths $(\rho_s, h_s)_{s \in [0,1]} \in \mathcal{A}(f, g)$.

For given $f, g \in \mathcal{F}$, there exists a smooth function $h \in \mathcal{A}$ such that $Lh = f - g$. In particular $\mathcal{A}(f, g)$ is nonempty since $\rho_s = sf + (1-s)g$, associated with the function $h_s = h$ (independent of s), is an admissible path.

Moreover, $T_2(f\mu, g\mu)$ is well defined and finite since it is bounded from above by

$$\int_0^1 \varphi(sf + (1-s)g, h)\, ds = \int_0^1 \int \frac{\Gamma(h)}{sf + (1-s)g}\, ds\, d\mu = \int \Gamma(h) \frac{\log(f) - \log(g)}{f - g}\, d\mu$$

which is finite since $h \in \mathcal{A}$. For instance, if $f, g > \eta$ for some $\eta > 0$, then

$$T_2^2(f\mu, g\mu) \leq \frac{1}{\eta} \int \Gamma(h) d\mu < \infty.$$

3.2 Remarks and examples

The Markov transportation distance heavily depends on both the reference measure μ and the generator L.

As it has been presented in the introduction, the Markov transportation distance is a generalization of the Benamou-Brenier dynamical formulation of the Wasserstein distance W_2 (see [BB00]). In our setting of a compact Riemannian manifold E equipped with the probability

measure $d\mu(x) = e^{-V(x)}dx$, it was indeed proven by F. Otto and M. Westdickenberg [OW05, Prop. 4.3 and (4.15)] that, given two positive densities (with respect to μ) f and g in $\mathcal{F}$,

$$W_2^2(f\mu, g\mu) = \inf \int_0^1 \int \frac{|w_s|^2}{\rho_s} d\mu ds$$

where the infimum runs over all smooth vector fields $(\rho_s, w_s)_{s \in [0,1]}$ on $[0,1] \times E$ with $\rho_s > 0$ in $\mathcal{F}$ satisfying

$$\partial_s \rho_s + \nabla \cdot w_s - \nabla V \cdot w_s = 0, \qquad \rho_0 = f, \rho_1 = g. \tag{14}$$

By comparison, in this setting,

$$T_2^2(f\mu, g\mu) = \inf \int_0^1 \int \frac{|\nabla h_s|^2}{\rho_s} d\mu ds$$

where the infimum runs over all smooth functions $(\rho_s, h_s)_{s \in [0,1]}$ satisfying

$$\partial_s \rho_s + \Delta_g h_s - \nabla V \cdot \nabla h_s = 0, \qquad \rho_0 = f, \rho_1 = g. \tag{15}$$

But $(\rho_s, w_s = \nabla h_s)$ satisfies (14) for any $(\rho_s, h_s)_{s \in [0,1]} \in \mathcal{A}(f, g)$, so $T_2^2(f\mu, g\mu) \geqslant W_2^2(f\mu, g\mu)$.

The Markov transportation distance as defined in Definitions 3.1 and 3.2 can also be considered in the more general setting of a Polish measure space, a generator L, its carré du champ Γ and a reference measure μ, but below we prefer to stick to our framework to be able to properly justify our computation. Here is however an example in the discrete case.

In the case of a countable state space E, a Markov semigroup $(P_t)_{t \geqslant 0}$ is described by an infinite matrix of positive kernels $(p_t(x,y))_{(x,y) \in E \times E}$, $t \geqslant 0$, such that for all $t \geqslant 0$ and $x \in E$, and any positive function f on E,

$$P_t f(x) = \sum_{y \in E} f(y) \, p_t(x, y).$$

For any $x \in E$, $p_t(x, .)$ is a probability measure on E. The generator L is given by an infinite matrix $(L(x,y))_{(x,y) \in E \times E}$, where for any finitely supported function f on E,

$$Lf(x) = \sum_{y \in E} L(x, y) f(y).$$

For the matrix L to be a generator, it is required that $L(x, y) \geqslant 0$ whenever $x \neq y$, and $\sum_y L(x, y) = 0$ for every $x \in E$. The carré du champ operator is defined on finitely supported functions f by

$$\Gamma(f)(x) = \frac{1}{2} \sum_{y \in E} L(x, y) \big[f(x) - f(y) \big]^2, \quad x \in E.$$

The measure is μ reversible if

$$\mu(x) L(x, y) = \mu(y) L(y, x).$$

Let us illustrate the discrete setting with the two point space $\{a, b\}$. The generator is unique up to a multiplicative factor, and is given by $Lf(a) = \kappa(f(b) - f(a))$ and $Lf(b) = \kappa(f(a) - f(b))$ for a nonnegative constant κ; moreover the carré du champ is constant, equal to

$$\Gamma(f) = \frac{\kappa}{2}(f(b) - f(a))^2,$$

and the reversible measure is $\mu = \frac{1}{2}(\delta_a + \delta_b)$. There one can simply and explicitly compute a geodesic curve for the T_2 distance between two generic measures $(1-r)\delta_a + r\delta_b$ and $(1-t)\delta_a + t\delta_b$ with $0 < r, t < 1$.

Let indeed $(\rho_s, h_s)_{s \in [0,1]}$ be an admissible path between $2(1-r)\mathbb{I}_a + 2r\mathbb{I}_b$ and $2(1-t)\mathbb{I}_a + 2t\mathbb{I}_b$. Then there exists a map $\varphi : [0,1] \to [0,1]$ such that $\varphi(0) = r$ and $\varphi(1) = t$, and

$$\rho_s = 2\varphi(s)\mathbb{I}_b + 2(1 - \varphi(s))\mathbb{I}_a.$$

The map h_s has to satisfy $2\varphi'(s)(\mathbb{I}_b - \mathbb{I}_a) = -Lh_s$ for $s \in [0,1]$, that is $\varphi'(s) = (h_s(b) - h_s(a))\kappa/2$. It remains to minimize

$$\int_0^1 \int \frac{\Gamma(h_s)}{\rho_s} d\mu ds = \frac{\kappa}{2} \int_0^1 \int \frac{(h_s(b) - h_s(a))^2}{\rho_s} d\mu ds = \frac{1}{\kappa} \int_0^1 \varphi'(s)^2 \Big(\frac{1}{\rho_s(a)} + \frac{1}{\rho_s(b)}\Big) ds$$

Since $\rho_s(a) = 2(1 - \varphi(s))$ and $\rho_s(b) = 2\varphi(s)$ we need to minimize

$$\frac{1}{2\kappa} \int_0^1 \varphi'(s)^2 \Big(\frac{1}{\varphi(s)} + \frac{1}{1 - \varphi(s)}\Big) ds$$

over all functions φ such that $\varphi(0) = r$ and $\varphi(1) = t$. The Euler-Lagrange equation is

$$2\varphi'' \Big(\frac{1}{\varphi} + \frac{1}{1 - \varphi}\Big) = \varphi'^2 \Big(\frac{1}{\varphi^2} - \frac{1}{(1 - \varphi)^2}\Big).$$

It implies that $\varphi'^2 = a\varphi(1 - \varphi)$ for some $a > 0$. We let $r = \sin^2 \theta$ and $t = \sin^2 \omega$ with $0 < \theta, \omega < \pi/2$. This solves into $\varphi(s) = \sin^2(\frac{\sqrt{a}}{2}s + \varepsilon\theta)$ for $s \in [0,1]$, and some $\varepsilon = \pm 1$. For such a φ,

$$\varphi'(s)^2 \Big(\frac{1}{\varphi(s)} + \frac{1}{1 - \varphi(s)}\Big) = \frac{\varphi'^2}{\varphi(1 - \varphi)} = a, \quad s \in [0,1].$$

But the smallest a for which $\varphi(1) = t$ is $a = 4(\omega - \theta)^2$ (and $\varepsilon = 1$) : hence φ is given by $\varphi(s) = \sin^2(s\omega + (1 - s)\theta)$. This implies that

$$T_2^2\big((2(1 - r)\mathbb{I}_a + 2r\mathbb{I}_b)\mu, (2(1 - t)\mathbb{I}_a + 2t\mathbb{I}_b)\mu\big) = \frac{2(\omega - \theta)^2}{\kappa}$$

for the T_2 distance defined as in Definitions 3.1 and 3.2. Moreover

$$\int \frac{\Gamma(h_s)}{\rho_s} d\mu = \frac{2(\omega - \theta)^2}{\kappa}$$

for all $s \in [0,1]$, so for such a φ the path (ρ_s, h_s) is a geodesic between $2(1 - r)\mathbb{I}_a + 2r\mathbb{I}_b$ and $2(1 - t)\mathbb{I}_a + 2t\mathbb{I}_b$.

In this example, in a discrete setting, one has obtained the existence of a geodesic path between two probability measures associated to the Markov transportation distance. Let us recall that there is no such geodesic for the Wasserstein distance.

3.3 General properties of T_2

We first exhibit ε-geodesics for the T_2 distance. Actually, we shall see below how properties on the distance and curvature-dimension bounds can be obtained without geodesics.

Proposition 3.3 (ε-geodesics) *Let $f, g \in \mathcal{F}$ and let $\varepsilon > 0$. Then there exists an ε-geodesic map, that is an admissible path $(\rho_s, h_s)_{s \in [0,1]} \in \mathcal{A}(f, g)$ such that for all $s \in [0, 1]$,*

$$\varphi(\rho_s, h_s) = \int \frac{\Gamma(h_s)}{\rho_s} d\mu \leq T_2^2(f\mu, g\mu) + \varepsilon.$$

Proof

$\triangleleft$ Let $\varepsilon > 0$ and an admissible path $(\rho_s, h_s) \in \mathcal{A}(f, g)$ such that

$$\Phi(\rho, h) = \int_0^1 \int \frac{\Gamma(h_s)}{\rho_s} d\mu\, ds \leq T_2^2(f\mu, g\mu) + \varepsilon.$$

It is easy to see that there exists $a > 0$ such that

$$\int_0^1 \sqrt{\varphi(\rho_u, h_u) + a}\, du = \sqrt{\Phi(\rho, h) + \varepsilon}.$$

Then let $\beta : [0, 1] \to [0, 1]$ be defined by

$$s = \frac{\displaystyle\int_0^{\beta(s)} \sqrt{\varphi(\rho_u, h_u) + a}\, du}{\sqrt{\Phi(\rho, h) + \varepsilon}}$$

for $s \in [0, 1]$. The function β is increasing and differentiable in $[0, 1]$ and satisfies $\beta(0) = 0$ and $\beta(1) = 1$, so $(\rho_{\beta(s)}, \beta'(s) h_{\beta(s)}) \in \mathcal{A}(f, g)$. Moreover

$$
\begin{aligned}
\varphi(\rho_{\beta(s)}, \beta'(s) h_{\beta(s)}) = \beta'(s)^2 \varphi(\rho_{\beta(s)}, h_{\beta(s)}) &= (\Phi(\rho, h) + \varepsilon) \frac{\varphi(\rho_{\beta(s)}, h_{\beta(s)})}{\varphi(\rho_{\beta(s)}, h_{\beta(s)}) + a} \\
&\leq \Phi(\rho, h) + \varepsilon \leq T_2^2(f\mu, g\mu) + 2\varepsilon
\end{aligned}
$$

for any $s \in [0, 1]$. This means that the couple $(\rho_\beta, \beta' h_\beta)$ is a 2ε-geodesic. $\triangleright$

Proposition 3.4 *The space $(\mathcal{F}, T_2)$ is a metric space.*

Proof

$\triangleleft$ For any $f \in \mathcal{F}$, then $T_2(f\mu, f\mu) = 0$ by choosing ρ_s constant equal to f. Conversely, if f and g in $\mathcal{F}$ are such that $T_2(f\mu, g\mu) = 0$, then $W_2(f\mu, g\mu) = 0$ since it is smaller than $T_2(f\mu, g\mu)$, as seen in section 3.2; hence $f = g$. Moreover T_2 is a symmetric function with respect to the two densities.

Let now f, g and h in $\mathcal{F}$. Let (ρ_s^1, h_s^1) (resp. (ρ_s^2, h_s^2)) be an ε-geodesic map between f and g (resp. g and h). Let $\alpha \in (0, 1)$ and define $(\rho_s, h_s)_{s \in [0,1]}$ by

$$
\rho_s = \begin{cases} \rho_{s/\alpha}^1, & \text{if } s \in [0, \alpha], \\ \rho_{(s-\alpha)/(1-\alpha)}^2, & \text{if } s \in [\alpha, 1]. \end{cases}
\qquad
h_s = \begin{cases} \dfrac{1}{\alpha} h_{s/\alpha}^1, & \text{if } s \in [0, \alpha), \\ \dfrac{1}{1-\alpha} h_{(s-\alpha)/(1-\alpha)}^2, & \text{if } s \in [\alpha, 1]. \end{cases}
$$

Then the couple (ρ_s, h_s) is an admissible path between f and h. Moreover

$$T_2^2(f\mu, h\mu) \leq \Phi(\rho, h) = \int_0^\alpha \left(\frac{1}{\alpha}\right)^2 \varphi(\rho_{s/\alpha}^1, h_{s/\alpha}^1)ds + \int_\alpha^1 \left(\frac{1}{1-\alpha}\right)^2 \varphi(\rho_{(s-\alpha)/(1-\alpha)}^2, h_{(s-\alpha)/(1-\alpha)}^2)ds.$$

Now ρ_s^1 and ρ_s^2 are ε-geodesics, so

$$\varphi(\rho_{s/\alpha}^1, h_{s/\alpha}^1) \leq T_2^2(f\mu, g\mu) + \varepsilon$$

and

$$\varphi(\rho_{(s-\alpha)/(1-\alpha)}^2, h_{(s-\alpha)/(1-\alpha)}^2) \leq T_2^2(g\mu, h\mu) + \varepsilon.$$

Hence

$$T_2^2(f\mu, h\mu) \leq \frac{1}{\alpha}(T_2^2(f\mu, g\mu) + \varepsilon) + \frac{1}{1-\alpha}(T_2^2(g\mu, h\mu) + \varepsilon).$$

Now, choose

$$\alpha = \frac{T_2(f\mu, g\mu)}{T_2(f\mu, g\mu) + T_2(g\mu, h\mu)}$$

and let ε go to 0 to obtain the triangular inequality

$$T_2(f\mu, h\mu) \leq T_2(f\mu, g\mu) + T_2(g\mu, h\mu).$$

$\triangleright$

Proposition 3.5 (Tensorization) *Let* $(P_t^i)_{t\geq 0}$, $i \in \{1, \cdots, N\}$ *be N Markov semigroups on compact connected Riemannian manifolds E_i with probability measure μ_i, with generators L_i and carrés du champ Γ_i as in Section 3.1. Then one can define a product semigroup $P_t = \otimes_{i=1}^N P_t^i$ on the product space $E = \times_{i=1}^N E_i$ with $\mu = \otimes_{i=1}^N \mu_i$ with generator $L = \oplus_{i=1}^N L_i$ and carré du champ $\Gamma = \oplus_{i=1}^N \Gamma_i$.*

Then, for any densities $f(x) = \prod_{i=1}^N f_i(x_i)$ and $g(x) = \prod_{i=1}^N g_i(x_i)$ $(x = (x_1, \cdots, x_N))$ in $\mathcal{F}$,

$$T_2^2(f\mu, g\mu) \geqslant \sum_{i=1}^N T_{2,i}^2(f_i\mu_i, g_i\mu_i). \tag{16}$$

Proof
$\triangleleft$ For simplicity we prove the result for $N = 2$. Let (ρ_s, h_s) be an admissible path between the densities $f_1(x)f_2(y)$ and $g_1(x)g_2(y)$. Let $\rho_s^1(x) = \int \rho_s(x, y)d\mu_2(y)$ and $h_s^1(x) = \int h_s(x, y)d\mu_2(y)$, and let ρ_s^2 and h_s^2 similarly defined. Then

$$\varphi(\rho_s, h_s) = \int \frac{\Gamma(h_s)}{\rho_s} d\mu_1 d\mu_2 = \int \frac{\Gamma_1(h_s)}{\rho_s} d\mu_1 d\mu_2 + \int \frac{\Gamma_2(h_s)}{\rho_s} d\mu_1 d\mu_2. \tag{17}$$

Let us first prove that

$$\int \frac{\Gamma_1(h_s)}{\rho_s} d\mu_2 \geqslant \frac{\Gamma_1(h_s^1)}{\rho_s^1}, \tag{18}$$

and similarly for the second coordinate. Since

$$\Gamma_1(f)(x) = \lim_{t\to 0} \frac{1}{2t} \int \int (f(y_1) - f(y_2))^2 p_t^1(x, dy_1) p_t^1(x, dy_2) \tag{19}$$

for every function f, and for the Markov kernel p_t^1 of the semigroup $(P_t^1)_{t\geqslant 0}$ (see for instance [BGL14]), then for all x

$$\int \frac{\Gamma_1(h_s)(x,y)}{\rho_s(x,y)}d\mu_2(y) =$$

$$\rho_s^1(x)\lim_{t\to 0}\frac{1}{2t}\int\int\int\left(\frac{h_s(z_1,y)-h_s(z_2,y)}{\rho_s(x,y)}\right)^2\frac{\rho_s(x,y)}{\rho_s^1(x)}\,d\mu_2(y)\,p_t^1(x,dz_1)\,p_t^1(x,dz_2)$$

$$\geqslant \frac{1}{\rho_s^1(x)}\lim_{t\to 0}\frac{1}{2t}\int\int(\int(h_s(z_1,y)-h_s(z_2,y))d\mu_2(y))^2\,p_t^1(x,dz_1)\,p_t^1(x,dz_2)=\frac{\Gamma_1(h_s^1)(x)}{\rho_s^1(x)}$$

by the Cauchy-Schwarz inequality for the probability measure $\frac{\rho_s(x,y)}{\rho_s^1(x)}d\mu_2(y)$.

By (17) and (18) written for both variables we obtain

$$\varphi(\rho_s,h_s)\geqslant \varphi_1(\rho_s^1,h_s^1)+\varphi_2(\rho_s^2,h_s^2).$$

After integration over $s\in[0,1]$, we get, for any admissible path $(\rho_s,h_s)\in\mathcal{A}(f_1f_2,g_1g_2)$:

$$\Phi(\rho,h)=\int_0^1\varphi(\rho_s,h_s)ds\geqslant\Phi_1(\rho^1,h^1)+\Phi_2(\rho^2,h^2)\geqslant T_2^2(f_1\mu_1,g_1\mu_1)+T_2^2(f_2\mu_2,g_2\mu_2)\quad(20)$$

since (ρ_s^i,h_s^i) is an admissible path between f_i and g_i, for $i=1,2$. The result follows by optimizing over (ρ_s,h_s). $\quad\triangleright$

3.4 First application : the Talagrand inequality

As explained in the introduction, we will recover classical bounds as the contraction properties in the Markov transportation distance. We first make some observations on related functional inequalities.

The so-called Otto-Villani Theorem says that a logarithmic Sobolev inequality with constant C (see (24) below) implies the Talagrand inequality

$$W_2^2(f\mu,\mu)\leq 4C\,\mathrm{Ent}_\mu(f)$$

for all probability densities f. This inequality has first been derived by M. Talagrand in [Tal96] and linked with the logarithmic Sobolev inequality in [OV00] (see also [BGL01]). The proofs in [OV00, GL13] are based on the general inequality

$$W_2^2(P_tf\mu,f\mu)\leq t(\mathrm{Ent}_\mu(f)-\mathrm{Ent}_\mu(P_tf))\tag{21}$$

(see also [GKO13]). We prove the same inequality for the larger T_2 distance :

Proposition 3.6 *For our diffusion semigroup $(P_t)_{t\geqslant 0}$ and any f in $\mathcal{F}$, there holds*

$$T_2^2(P_tf\mu,f\mu)\leq t\,(\mathrm{Ent}_\mu(f)-\mathrm{Ent}_\mu(P_tf))\tag{22}$$

for every $t\geqslant 0$, and in particular

$$\limsup_{t\to 0^+}\frac{T_2(P_tf\mu,f\mu)}{t}\leq\sqrt{\int\frac{\Gamma(f)}{f}d\mu}.\tag{23}$$

Proof

◁ Let f in $\mathcal{F}$ be given. Then $(P_{st}f, -tP_{st}f)_{s\in[0,1]}$ is an admissible couple between f and $P_t f$. By definition of the distance T_2, it implies that

$$T_2^2(P_t f\mu, f\mu) \leq t \int_0^t \int \frac{\Gamma(P_r f)}{P_r f} d\mu dr$$

by change of time variable. Moreover

$$\frac{d}{dr}\mathrm{Ent}_\mu(P_r f) = \frac{d}{dr}\int P_r f \log P_r f \, d\mu = -\int \frac{\Gamma(P_r f)}{P_r f} \, d\mu$$

by diffusion property of the semigroup. This leads to (22) by integrating in r. ▷

Corollary 3.7 *In our notation, assume that the probability measure μ satisfies a logarithmic Sobolev inequality with constant C, that is,*

$$\mathrm{Ent}_\mu(f) \leq C \int \frac{\Gamma(f)}{f} d\mu, \tag{24}$$

for any f in $\mathcal{F}$. Then

$$T_2^2(f\mu, P_T f\mu) \leq 4C\mathrm{Ent}_\mu(f)$$

for any f in $\mathcal{F}$. In particular, if for instance T_2 is lower semicontinuous with respect to narrow convergence, then μ satisfies a Talagrand type inequality for the distance T_2, namely

$$T_2^2(f\mu, \mu) \leq 4C\mathrm{Ent}_\mu(f) \tag{25}$$

for any f in $\mathcal{F}$.

Proof

◁ Let f in $\mathcal{F}$ be given, and let $\varphi(t) = \mathrm{Ent}_\mu(P_t f)$. Then (23) and semigroup properties imply

$$\frac{d^+}{dt}T_2(P_t f\mu, f\mu) \leq \sqrt{-\varphi'(t)}.$$

Moreover the logarithmic Sobolev inequality for μ ensures that $\varphi'(t) \leq -\varphi(t)/C$, and thus

$$\sqrt{-\varphi'(t)} \leq -\sqrt{4C}\left(\sqrt{\varphi(t)}\right)'.$$

Let now $T > 0$. Then

$$T_2(P_T f\mu, f\mu) \leq \int_0^T \frac{d^+}{dt}T_2(P_t f\mu, f\mu)dt \leq -\sqrt{4C}\int_0^T \left(\sqrt{\varphi(t)}\right)'dt$$

$$= \sqrt{4C}\left(\sqrt{\mathrm{Ent}_\mu(f)} - \sqrt{\mathrm{Ent}_\mu(P_T f)}\right) \leq \sqrt{4C}\sqrt{\mathrm{Ent}_\mu(f)}.$$

Moreover $P_T f\mu$ narrowly converges to $f\mu$. Hence

$$T_2(\mu, f\mu) \leq \liminf_{T\to\infty} T_2(P_T f\mu, f\mu) \leq \sqrt{4C}\sqrt{\mathrm{Ent}_\mu(f)}$$

if T_2 is lower semicontinuous with respect to narrow convergence. ▷

Inequality (25) will be useful in the following section in the derivation of refined convergence rates.

4 Contraction property under the curvature-dimension condition $CD(R, n)$

In this section we prove a dimension dependent contraction property in the Markov transportation distance. We will see that the Γ_2-calculus is a well adapted and efficient tool.

4.1 Curvature condition, examples and useful commutation properties

The Γ_2-operator, or iterated carré du champ operator, is defined on functions $f \in \mathcal{A}$ by the general expression

$$\Gamma_2(f) = \frac{1}{2}\Big(L\Gamma(f) - 2\Gamma(f, Lf)\Big).$$

Definition 4.1 *The Markov semigroup $(P_t)_{t \geqslant 0}$ (associated to the generator L) is said to satisfy a curvature-dimension condition $CD(R, n)$ for $R \in \mathbb{R}$ and $n \geqslant 1$ if*

$$\Gamma_2(f) \geqslant \rho\,\Gamma(f) + \frac{1}{n}(Lf)^2,$$

for all functions $f \in \mathcal{A}$.

This criterion has been introduced in the seminal paper [BÉ85] by D. Bakry and M. Émery. For instance the Laplace-Beltrami operator on the sphere $S^n \subset \mathbb{R}^{n+1}$ satisfies a $CD(n-1, n)$ condition. More generally, for a complete Riemannian manifold M with dimension d, equipped with the Laplace-Beltrami operator Δ_g and the Riemannian measure dx, the curvature-dimension condition $CD(R, d)$ holds for Δ_g if the Ricci curvature of M is uniformly bounded from below by R. An explicit condition on the potential V which is equivalent to a $CD(R, n)$ condition for $L = \Delta_g - \nabla V \cdot \nabla$ is given in [BGL14, Appendix C.6] for instance. Observe that n need not be the dimension on the manifold.

One of the main results concerning the curvature-dimension condition $CD(R, \infty)$ is a regularity property of the Markov semigroup. The $CD(R, \infty)$ conditions holds for a diffusive Markov semigroup if and only if for any function $f \in \mathcal{A}$

$$\Gamma(P_t(f)) \leq e^{-2Rt}\Big(P_t\sqrt{\Gamma(f)}\Big)^2. \tag{26}$$

This result, proved in [Bak94], is the key point for many applications such as logarithmic Sobolev inequalities, Harnack parabolic inequalities, etc. (see [BGL14]). Gradient bounds, in a weaker form, also hold under the $CD(R, n)$ condition with finite n (see [BL06] and [Wan11]). Here is a new such bound which will be the key point for the Markov transportation distance.

Lemma 4.2 *Let $(P_t)_{t \geqslant 0}$ be a diffusion Markov semigroup, and let $R \in \mathbb{R}$ and $n \geqslant 1$. The following assertions are equivalent :*

(i) *The Markov semigroup satisfies a $CD(R, n)$ condition.*

(ii) *For all functions $f, g \in \mathcal{A}$ with $g > 0$ and all $t \geqslant 0$*

$$\frac{\Gamma(P_t f)}{P_t g} \leq e^{-2Rt} P_t\left(\frac{\Gamma(f)}{g}\right) - \frac{2}{n}\int_0^t e^{-2Ru}\frac{\big[LP_t f - P_u(\Gamma(P_{t-u}f, \log P_{t-u}g))\big]^2}{P_t g}du. \tag{27}$$

In particular, under the $CD(R,n)$ condition and for any $t \geqslant 0$, $f \in \mathcal{A}$ and $g \in \mathcal{F}$,

$$\int \frac{\Gamma(P_t f)}{P_t g} d\mu \leq e^{-2Rt} \int \frac{\Gamma(f)}{g} d\mu - \frac{2}{n} \int_0^t e^{-2R(t-u)} \left(\int \frac{\Gamma(P_u f, P_u g)}{P_u g} d\mu \right)^2 du. \tag{28}$$

Proof

◁ Let us first prove that (i) implies (ii). We let $t > 0$ and $f, g \in \mathcal{A}$ be fixed, with $g > 0$. Then we define

$$\Lambda(s) = P_s \left(\frac{\Gamma(P_{t-s}f)}{P_{t-s}g} \right)$$

for $s \in [0, t]$, and then $F = P_{t-s}f$ and $G = P_{t-s}g$. Then

$$\Lambda'(s) = P_s \left(-2\frac{\Gamma(F, LF)}{G} + \Gamma(F)\frac{LG}{G^2} + L\left(\frac{\Gamma(F)}{G} \right) \right).$$

But

$$L(hk) = 2\Gamma(h, k) + hLk + kLh$$

for any function $h, k \in \mathcal{A}$, so the diffusion property (12) and the definition of Γ_2 lead to

$$\Lambda'(s) = 2P_s \left(\frac{1}{G} \left[\Gamma_2(F) - \Gamma(\Gamma(F), \log G) + \Gamma(F)\Gamma(\log G) \right] \right).$$

Now Lemma 4.3 below, applied with $f = F$ and $g = -\log G$, ensures that

$$\Gamma_2(F) - \Gamma(\Gamma(F), \log G) + \Gamma(F)\Gamma(\log G) \geqslant R\,\Gamma(F) + \frac{1}{n}(LF - \Gamma(F, \log G))^2.$$

Since $G \geqslant 0$, this gives

$$\Lambda'(s) \geqslant 2R\Lambda(s) + \frac{2}{n}P_s \left(\frac{[LF - \Gamma(F, \log G)]^2}{G} \right) \geqslant 2R\Lambda(s) + \frac{2}{n}\frac{[LP_t f - P_s(\Gamma(P_{t-s}f, \log P_{t-s}g))]^2}{P_t g}$$

by the Cauchy-Schwarz inequality for the Markov kernel of P_s and semigroup properties. Inequality (27) follows by integration over $s \in [0, t]$.

Let us now assume (ii) and let $g = 1$. Then inequality (27) writes

$$\Gamma(P_t f) \leq e^{-2Rt} P_t \Gamma(f) - \frac{2}{n} \int_0^t e^{-2Ru}(LP_t f)^2 du.$$

Taking the time derivative at $t = 0$ implies back the $CD(R,n)$ condition.

Let us finally prove (28): integrating (27) with respect to μ gives

$$\int \frac{\Gamma(P_t f)}{P_t g} d\mu \leq e^{-2Rt} \int \frac{\Gamma(f)}{g} d\mu - \frac{2}{n} \int_0^t e^{-2Ru} \left[\int \frac{[LP_t f - P_u(\Gamma(P_{t-u}f, \log P_{t-u}g))]^2}{P_u g} d\mu \right] du$$

by invariance property of μ. Then the Cauchy-Schwarz inequality for the measure μ implies (28) by recalling that $\int g d\mu = 1$, invariance property of μ and change of time variable. ▷

16

Lemma 4.3 *For a diffusion Markov semigroup, under the curvature-dimension condition $CD(R, n)$ (with $R \in \mathbb{R}$), for all functions $f, g \in \mathcal{A}$*

$$\Gamma_2(f) + \Gamma(\Gamma(f), g) + \Gamma(f)\Gamma(g) \geqslant R\,\Gamma(f) + \frac{1}{n}(Lf + \Gamma(f, g))^2. \tag{29}$$

Proof

◁ The proof is inspired from Lemma 5.4.4, p. 83 of [ABC$^+$00]. Let $f, g \in \mathcal{A}$ and $x_0 \in E$. Let Φ be a smooth map on $\mathbb{R}^2$ such that

$$\partial_2\Phi = \partial_{11}^2\Phi = \partial_{22}^2\Phi = 0, \quad \partial_1\Phi = 1 \quad \text{and} \quad \partial_{12}^2\Phi = \frac{1}{2}$$

at the point $(f(x_0), g(x_0))$. Then the $CD(R, n)$ condition applied to the function $\Phi(f, g)$ at the point x_0 yields

$$\Gamma_2(\Phi(f, g)) \geqslant R\,\Gamma\Phi(f, g) + \frac{1}{n}(L\Phi(f, g))^2.$$

The usual change of variable rules for the Γ and Γ_2 operators (see for instance [ABC$^+$00, p. 83]) imply

$$\Gamma_2(f) + \Gamma(\Gamma(f), g) + \frac{1}{2}\big[\Gamma(f, g)^2 + \Gamma(f)\Gamma(g)\big] \geqslant R\Gamma(f) + \frac{1}{n}(Lf + \Gamma(f, g))^2.$$

The result follows since $\Gamma(f, g)^2 \leq \Gamma(f)\Gamma(g)$. ▷

Remark 4.4 *Without the dimension, namely under the curvature-dimension condition $CD(R, \infty)$, inequality (27) is a direct consequence of inequality (26). Indeed (26) implies*

$$\frac{\Gamma(P_t f)}{P_t g} \leq e^{-2Rt}\frac{P_t(\sqrt{\Gamma(f)})^2}{P_t g} \leq e^{-2Rt}P_t\left(\frac{\Gamma(f)}{g}\right),$$

by the Cauchy-Schwarz inequality for the Markov kernel of P_t.

4.2 Contraction property under $CD(R, n)$

The following result is a extension of Proposition 2.2 for the heat semigroup on $\mathbb{R}^n$ to our diffusion Markov semigroup on a manifold under the curvature-dimension condition $CD(R, n)$. For $R = 0$ we precisely recover the bound obtained in Wasserstein distance for the heat semigroup on $\mathbb{R}^n$.

Theorem 4.5 *Let $(P_t)_{t \geqslant 0}$ be our diffusion Markov semigroup on E satisfying a $CD(R, n)$ condition with $R \in \mathbb{R}$ and $n \geqslant 1$. Then, for all $f, g \in \mathcal{F}$ and $T \geqslant 0$,*

$$T_2^2(P_T f\mu, P_T g\mu) \leq e^{-2RT}T_2^2(f\mu, g\mu) - \frac{2}{n}\int_0^T e^{-2R(T-t)}(\mathrm{Ent}_\mu(P_t g) - \mathrm{Ent}_\mu(P_t f))^2 dt. \tag{30}$$

Proof

◁ Let (ρ_s, h_s) be an admissible path between f and g. Then $(P_t(\rho_s), P_t(h_s))_{s \in [0,1]}$ is also an admissible path between $P_t f$ and $P_t g$.

Then inequality (28) of Lemma 4.2, applied at time T to the functions h_s and ρ_s, implies

$$\int \frac{\Gamma(P_T h_s)}{P_T \rho_s} d\mu \leq e^{-2RT} \int \frac{\Gamma(h_s)}{\rho_s} d\mu - \frac{2}{n} \int_0^T e^{-2R(T-t)} \left(\int \frac{\Gamma(P_t(\rho_s), P_t(h_s))}{P_t(\rho_s)} d\mu \right)^2 dt.$$

Integrating over $s \in [0,1]$ and the Cauchy-Schwarz inequality imply

$$\int_0^1 \int \frac{\Gamma(P_T h_s)}{P_T \rho_s} d\mu \, ds \leq e^{-2RT} \int_0^1 \int \frac{\Gamma(h_s)}{\rho_s} d\mu \, ds$$
$$- \frac{2}{n} \int_0^T e^{-2R(T-t)} \left(\int_0^1 \int \frac{\Gamma(P_t(\rho_s), P_t(h_s))}{P_t(\rho_s)} d\mu \, ds \right)^2 dt.$$

We finally obtain (30) since, letting $\varphi(s) = \mathrm{Ent}_\mu(P_t \rho_s)$,

$$\int_0^1 \int \frac{\Gamma(P_t(\rho_s), P_t(h_s))}{P_t(\rho_s)} d\mu \, ds = \int_0^1 \varphi'(s) \, ds = \varphi(1) - \varphi(0) = \mathrm{Ent}_\mu(P_t g) - \mathrm{Ent}_\mu(P_t f)$$

by the reversibility property (11) of the semigroup. $\triangleright$

Remark 4.6 *As noted in the introduction, this result has a particular new flavor. Indeed the recent results (2)-(3) in [BGL13], [EKS13] and [Kuw13] present a dimensional correction term for the contraction property, but for solutions at different times only. If the approaches for these inequalities are slightly different, it would be of interest to obtain a dimensional correction term for our contraction also in different times. A possible approach could be through Evolution variational inequalities, as studied in the next section, as the contraction result in [EKS13] is deduced from these inequalities.*

Assuming that μ is a probability measure, and taking $g = 1$, under the $CD(R, \infty)$ condition, the following bound

$$T_2^2(P_T f\mu, \mu) \leq e^{-2Rt} T_2^2(f\mu, \mu)$$

holds for the T_2 distance as it does for the W_2 distance. The following corollary gives a more precise bound under the $CD(R, n)$ condition:

Corollary 4.7 *Let $(P_t)_{t \geq 0}$ be our diffusion Markov semigroup on E satisfying a $CD(R, n)$ condition with $R > 0$ and $n \geq 1$. Then, in the framework of Corollary 3.7, for all $f \in \mathcal{F}$ and $T \geq 0$,*

$$T_2^2(P_T f\mu, \mu) \leq e^{-2RT} T_2^2(f\mu, \mu) \frac{1}{1 + n\, R\, T_2^2(f\mu, \mu) \frac{1 - e^{-2RT}}{4(n-1)^2}}.$$

Proof
$\triangleleft$ Taking $g = 1$ in (30), the map $\Lambda(t) = e^{2Rt} T_2^2(P_T f\mu, \mu)$ satisfies

$$\Lambda'(t) \leq -\frac{nR^2}{2(n-1)^2} e^{-2Rt} \Lambda(t)^2.$$

Observe indeed that, under the $CD(R, n)$ condition, the measure μ satisfies the logarithmic Sobolev inequality (24) with constant $C = \frac{n-1}{2Rn}$ (see [BGL14]), whence a Talagrand inequality (25) with constant $4C$ by Corollary 3.7. The conclusion follows by time integration. $\triangleright$

5 Evolution variational inequalities

Evolution variational inequalities (EVI in short) have recently been developed as a connection between curvature conditions $CD(R, \infty)$ (usually in the sense of the commutation of the semi-group and the carré du champ), heat semigroups and the notion of curvature bound introduced by J. Lott, K.-T. Sturm and C. Villani (see [Stu06] and [LV09]). We refer to the recent work [AGS12a] for a nearly complete picture in this dimensionless setting. However, no dimensional EVI, namely related to a $CD(R, n)$ curvature-dimension condition, were known until M. Erbar, K. Kuwada and K.-T. Sturm very recently proved in [EKS13] that $CD(R, n)$ is (roughly) equivalent to

$$\frac{d}{dt} s_{\frac{R}{n}} \left(\frac{1}{2} W_2(f\mu, H_t g\mu) \right)^2 \leq -R\, s_{\frac{R}{n}} \left(\frac{1}{2} W_2(f\mu, H_t g\mu) \right)^2 + \frac{n}{2} \left(1 - e^{-\frac{1}{n}(\text{Ent}_\mu(f) - \text{Ent}_\mu(H_t g))} \right) \quad (31)$$

on the heat semigroup. Here $s_r(x) = \sin(\sqrt{r}x)/\sqrt{r}$ if $r > 0$ and $s_r(x) = \sinh(\sqrt{-r}x)/\sqrt{-r}$ if $r < 0$. Forgetting for a time the map s_r, which is equivalent to x for small x, and using that $1 - e^{-x} \leq x$, this inequality clearly appears to improve the classical EVI obtained under a $CD(R, \infty)$ type condition.

The main goal of this section is twofold. First, we will give a (time integrated) EVI for the Markov transportation distance T_2 (rather than the usual Wasserstein distance), under the $CD(R, \infty)$ condition. Then we will see how the possible existence of geodesics can lead to a dimensional EVI, here with a negative corrective term in the spirit of the contraction result in Theorem 4.5. As in section 2 we will start with the Euclidean heat equation, i.e. under a $CD(0, n)$ condition, obtaining a dimensional EVI in Wasserstein distance. Assuming the existence of smooth geodesics for T_2, we will see that one can obtain a statement under the $CD(R, n)$.

Evolution variational inequality for T_2 under $CD(R, \infty)$

Theorem 5.1 *Let $(P_t)_{t \geq 0}$ be our diffusion Markov semigroup satisfying a $CD(R, \infty)$ condition with $R \in \mathbb{R}$. Then, for all $f, g \in \mathcal{F}$,*

$$T_2^2(f\mu, P_t g\mu) - T_2^2(f\mu, g\mu) \leq -\frac{e^{-2Rt} - 1 + 2Rt}{2Rt} T_2^2(f\mu, g\mu) + 2t(\text{Ent}_\mu(f) - \text{Ent}_\mu(P_t g)).$$

Proof

◁ Let (ρ_s, h_s) be an admissible path between f and g. Then $(P_{ts}(\rho_s), P_{ts}(h_s - t\rho_s))$ is an admissible path between f and $P_t g$, so

$$T_2^2(f\mu, P_t g\mu) \leq \int_0^1 \int \frac{\Gamma(P_{ts}(h_s - t\rho_s))}{P_{ts}(\rho_s)} d\mu\, ds$$

$$= \int_0^1 \int \frac{\Gamma(P_{ts}(h_s))}{P_{ts}(\rho_s)} d\mu\, ds - 2t \int_0^1 \int \frac{\Gamma(P_{ts}(h_s - t\rho_s), P_{ts}(\rho_s))}{P_{ts}(\rho_s)} d\mu\, ds - t^2 \int_0^1 \int \frac{\Gamma(P_{ts}(\rho_s))}{P_{ts}(\rho_s)} d\mu\, ds. \quad (32)$$

By Lemma 4.2 under the $CD(R, \infty)$ curvature condition, then for all s

$$\int \frac{\Gamma(P_{ts}(h_s))}{P_{ts}(\rho_s)} d\mu \leq e^{-2Rts} \int \frac{\Gamma(h_s)}{\rho_s} d\mu \leq e^{-2Rts} \left(T_2^2(f\mu, g\mu) + \varepsilon \right)$$

19

if the admissible path (ρ_s, h_s) is an ε-geodesic, for given $\varepsilon > 0$. Moreover

$$\int \frac{\Gamma(P_{ts}(h_s - t\rho_s), P_{ts}(\rho_s))}{P_{ts}(\rho_s)} d\mu = \frac{d}{ds} \int P_{ts}(\rho_s) \log P_{ts}(\rho_s) d\mu$$

by reversibility. Hence, forgetting the last term in (32) and integrating in $s \in [0, 1]$ conclude the argument by letting ε go to 0. $\triangleright$

Remark 5.2 *Combining the contraction result of Theorem 4.5 under a $CD(R, n)$ condition and this EVI (of course valid under $CD(R, n)$) we may get a contraction type result in the T_2 distance for $s < t$, in the spirit of [EKS13] (see Remark 2.4).*

To obtain dimension dependent bounds under the $CD(R, n)$ condition we will need geodesics. As above for the contraction property, let us see first which additional term coming from the dimension appears for the Euclidean heat equation.

5.1 A dimensional EVI in Wasserstein distance for the heat equation in $\mathbb{R}^n$

For simplicity the EVI is here described in its time derivative form, but it may be easily justified by first considering an integrated form of the EVI.

As in section 2, the Benamou-Brenier formulation (7) is the starting point. Let (ρ_s, w_s) be an admissible path between f and g, satisfying the constraint (8). Then $(H_{ts}(\rho_s), R_{ts}(w_s) - t\nabla H_{ts}(\rho_s))$ is an admissible path between f and $H_t g$, and satisfying (8), so by (7),

$$W_2^2(f dx, H_t g dx) \leq \int_0^1 \int \frac{|R_{ts}(w_s) - t\nabla H_{ts}(\rho_s)|^2}{H_{ts}(\rho_s)} ds dx. \tag{33}$$

Assume further that (ρ_s, w_s) is a minimizer in the Benamou-Brenier formulation, that is,

$$W_2^2(f dx, g dx) = \int_0^1 \int \frac{|w_s|^2}{\rho_s} ds dx.$$

The path (ρ_s) is then a geodesic path between $f dx$ and $g dx$ with respect to the Wasserstein distance (see [DNS09] for more details). In particular inequality (33) is an equality at time $t = 0$ and, formally, the time derivative of (33) at $t = 0$ implies

$$\frac{d}{dt} W_2^2(f dx, H_t g dx) \Big|_{t=0} \leq \frac{d}{dt} \int_0^1 \int \frac{|R_{ts}(w_s) - t\nabla H_{ts}(\rho_s)|^2}{H_{ts}(\rho_s)} ds dx \Big|_{t=0}. \tag{34}$$

The term on the right-hand side is controlled by the following lemma.

Lemma 5.3 *Let $(\rho_s, w_s)_{s \in [0,1]}$ be a couple satisfying the constraint (8), where ρ_s is a probability density with respect to Lebesgue measure. Letting*

$$\Lambda(t) = \int_0^1 \int \frac{|R_{ts}(w_s) - t\nabla H_{ts}(\rho_s)|^2}{H_{ts}(\rho_s)} ds dx$$

for $t \geqslant 0$, then

$$\Lambda'(0) \leq -\frac{2}{n} \int_0^1 s \left(\int \frac{w_s \cdot \nabla \rho_s}{\rho_s} dx \right)^2 ds - 2 \int_0^1 \int \frac{w_s \cdot \nabla \rho_s}{\rho_s} ds dx.$$

20

We skip the proof since it is almost the same as for Lemma 2.1.

Now Lemma 5.3 and (34) imply

$$\frac{d}{dt}W_2^2(fdx, H_tgdx)\Big|_{t=0} \leq -\frac{2}{n}\int_0^1 s\left(\int \frac{w_s \cdot \nabla \rho_s}{\rho_s}dx\right)^2 ds - 2\int_0^1 \int \frac{w_s \cdot \nabla \rho_s}{\rho_s}dsdx.$$

Letting $\varphi(s) = \int \rho_s \log \rho_s dx$, the relation (8) between w_s and ρ_s implies

$$\varphi'(s) = \int \frac{w_s \cdot \nabla \rho_s}{\rho_s}dx,$$

so that

$$\frac{d}{dt}W_2^2(fdx, H_tgdx)|_{t=0} \leq -\frac{2}{n}\int_0^1 s(\varphi'(s))^2 ds - 2\int_0^1 \varphi'(s)ds.$$

Then the Jensen inequality for the measure $2\,s\,ds$ and an integration by parts give

$$\frac{d}{dt}W_2^2(fdx, H_tgdx)|_{t=0} \leq -\frac{4}{n}\left(\varphi(1) - \int_0^1 \varphi(s)ds\right)^2 - 2(\varphi(1) - \varphi(0)).$$

We have obtained the following result :

Proposition 5.4 *Let $(H_t)_{t\geqslant 0}$ be the heat semigroup on $\mathbb{R}^n$. Then, for any probability densities f and g in $\mathbb{R}^n$ such that $W_2(fdx, gdx) < \infty$,*

$$\frac{1}{2}\frac{d}{dt}W_2^2(fdx, H_tgdx)\Big|_{t=0} \leq -\frac{2}{n}\left(\text{Ent}_{dx}(g) - \int_0^1 \text{Ent}_{dx}(\rho_s)\,ds\right)^2 + \text{Ent}_{dx}(f) - \text{Ent}_{dx}(g) \quad (35)$$

where $(\rho_s)_{s\in[0,1]}$ is a geodesic path between f and g for the Wasserstein distance.

At the time being, we have not been able to get a "geodesic" free version of this dimensional EVI. Note however, once again, that the correction term is quite different in nature from the one obtained for example in [EKS13]. We will see in the next subsection that we obtain results in the same flavor with the T_2 distance.

5.2 A dimension dependent EVI for T_2 in the geodesic case

In this subsection we will assume the existence of smooth geodesics for the T_2 distance, for any $f, g \in \mathcal{F}$, there exists an admissible path $(\rho_s, h_s)_{s\in[0,1]} \in \mathcal{A}(f, g)$ such that for all $s \in [0, 1]$,

$$\int \frac{\Gamma(h_s)}{\rho_s}d\mu = T_2^2(f\mu, g\mu).$$

Therefore, this section is formal, as also we should first consider the integrated form of EVI (which can be obtained, but is however quite difficult to read). We will closely follow the approach used above for the Euclidean heat equation and the Wasserstein distance.

We then consider our diffusion Markov semigroup $(P_t)_{t\geqslant 0}$ and, for $f, g \in \mathcal{F}$, a geodesic path (ρ_s, h_s) between f and g for the associated T_2 distance. Then $(P_{ts}(\rho_s)), P_{ts}(h_s - t\rho_s))$ is an admissible path between for f and P_tg. In particular

$$T_2^2(f\mu, P_tg\mu) \leq \int_0^1 \int \frac{\Gamma(P_{ts}(h_s - t\rho_s))}{P_{ts}(\rho_s)}d\mu ds. \quad (36)$$

We will use the following adaptation of Lemma 5.3 to our setting, which does not need (ρ_s, h_s) to be a geodesic path:

Lemma 5.5 *Let $(\rho_s)_{s\in[0,1]}$ be a smooth path in $\mathcal{F}$, $(h_s)_{s\in[0,1]}$ be a smooth path in $\mathcal{A}$, and*

$$\Lambda(t) = \int_0^1 \int \frac{\Gamma(P_{ts}(h_s - t\rho_s))}{P_{ts}(\rho_s)} d\mu ds$$

for $t \geqslant 0$. Then, under the $CD(R,n)$ condition,

$$\Lambda'(0) \leq -2R \int_0^1 s \int \frac{\Gamma(h_s)}{\rho_s} d\mu ds - \frac{2}{n} \int_0^1 s \left(\int \frac{\Gamma(h_s, \rho_s)}{\rho_s} d\mu \right)^2 ds - 2 \int_0^1 \int \frac{\Gamma(h_s, \rho_s)}{\rho_s} d\mu ds.$$

Since (ρ_s, h_s) is a geodesic path, then (36) is an equality at time $t = 0$, so, taking the time derivative at $t = 0$ and using Lemma 5.5,

$$\begin{aligned}
\frac{1}{2}\frac{d}{dt}T_2^2(f\mu, P_t g\mu)\Big|_{t=0} &\leq -R \int_0^1 s \int \frac{\Gamma(h_s)}{\rho_s} d\mu ds \\
&\quad -\frac{1}{n} \int_0^1 s \left(\int \frac{\Gamma(h_s, \rho_s)}{\rho_s} d\mu \right)^2 ds - \int_0^1 \int \frac{\Gamma(h_s, \rho_s)}{\rho_s} d\mu ds.
\end{aligned}$$

Again (ρ_s, h_s) is a geodesic path, so

$$\int \frac{\Gamma(h_s)}{\rho_s} d\mu = T_2^2(f\mu, g\mu)$$

for all s. Letting again $\varphi(s) = \int \rho_s \log \rho_s d\mu$, the inequality may then be rewritten as

$$\frac{1}{2}\frac{d}{dt}T_2^2(f\mu, P_t g\mu)\Big|_{t=0} \leq -\frac{R}{2} T_2^2(f\mu, g\mu) - \frac{1}{n} \int_0^1 s\, \varphi'(s)^2\, ds - \int_0^1 \varphi'(s) ds.$$

Hence, by the Jensen inequality for the measure $2sds$ and an integration by parts, the $CD(R,n)$ condition and the existence of geodesics ensure the following dimensional EVI:

$$\frac{1}{2}\frac{d}{dt}T_2^2(f\mu, P_t g\mu)\Big|_{t=0} \leq -\frac{R}{2}T_2^2(f\mu, g\mu) - \frac{2}{n}\left(\mathrm{Ent}_\mu(g) - \int_0^1 \mathrm{Ent}_\mu(\rho_s)\, ds \right)^2 + \mathrm{Ent}_\mu(f) - \mathrm{Ent}_\mu(g)\,,\tag{37}$$

where (ρ_s) is a geodesic path between $f\mu$ and $g\mu$ for the T_2 distance.

6 Φ-entropies versus usual entropy

There are many ways of extending the Markov transportation distance. Here we present the one associated with Φ-entropies, well adapted to the Γ_2-calculus. For the Wasserstein distance this generalization has been formulated in [DNS09, DNS12].

Let again $(P_t)_{t\geqslant 0}$ be a diffusion Markov semigroup with invariant measure μ. Let ξ be a $\mathcal{C}^2$ positive function on $(0, +\infty)$ with $1/\xi$ concave. Let also

$$\mathrm{Ent}_\mu^\Phi(f) = \int \Phi(f) d\mu - \Phi\left(\int f d\mu \right)$$

be the Φ-entropy of a positive map f, with $\Phi'' = \xi$. The Φ-entropies have been studied for instance in [Bak94, Cha04, AD05, BG10]. By analogy with Definition 3.1, for $\Phi(x) = x \log x$ and $\xi(x) = 1/x$:

Definition 6.1 *For $f, g \in \mathcal{F}$ we let*

$$T_\xi(f\mu, g\mu) = \inf \left(\int_0^1 \int \Gamma(h_s)\xi(\rho_s) d\mu ds \right)^{1/2}$$

where the infimum runs over all admissible paths $(\rho_s, h_s)_{s \in [0,1]} \in \mathcal{A}(f, g)$.

For a general map ξ this distance shares the same properties of existence of ε-geodesics and tensorization as the distance T_2, which can be proved as in section 3. For instance:

Proposition 6.2 (Tensorization) *Let ξ be a $\mathcal{C}^2$ positive function on $(0, +\infty)$ with $1/\xi$ concave. Let $(P_t^i)_{t \geqslant 0}$, $i \in \{1, \cdots, N\}$ be N Markov semigroups on compact connected Riemannian manifolds E_i with probability measure μ_i, with generators L_i and carrés du champ Γ_i as in Section 3.1. Let $P_t = \otimes_{i=1}^N P_t^i$ on the product space $E = \times_{i=1}^N E_i$, as in Proposition 3.5.*
  Then, for any densities $f(x) = \prod_{i=1}^N f_i(x_i)$ and $g(x) = \prod_{i=1}^N g_i(x_i)$ $(x = (x_1, \cdots, x_N))$ in $\mathcal{F}$,

$$T_\xi^2(f\mu, g\mu) \geqslant \sum_{i=1}^N T_\xi^2(f_i\mu_i, g_i\mu_i). \tag{38}$$

Proof
◁ The argument follows the proof of Proposition 3.5. The key bound (18) is replaced by

$$\int \Gamma_1(h_s)\xi(\rho_s) d\mu_2 \geqslant \Gamma_1(h_s^1)\xi(\rho_s^1).$$

This is a consequence of the definition (19) of the carré du champ and of the fact that the map $(x, y) \mapsto x^2 \xi(y)$ is convex under our assumption on ξ, see [Cha04]. ▷

Contraction and evolution variational inequalities

Theorem 6.3 *Let $(P_t)_{t \geqslant 0}$ be our diffusion Markov semigroup satisfying a $CD(R, \infty)$ condition with $R \in \mathbb{R}$. Then for any $f, g \in \mathcal{F}$ and $t \geqslant 0$, the contraction property*

$$T_\xi^2(P_t f\mu, P_t g\mu) \leq e^{-2Rt} T_\xi^2(f\mu, g\mu) \tag{39}$$

holds, as well as the Evolution Variational Inequality

$$T_\xi^2(f\mu, P_t g\mu) - T_\xi^2(f\mu, g\mu) \leq -\frac{e^{-2Rt} - 1 + 2Rt}{2Rt} T_\xi^2(f\mu, g\mu) + 2t \left(\mathrm{Ent}_\mu^\Phi(f) - \mathrm{Ent}_\mu^\Phi(P_t g) \right), \tag{40}$$

where $\Phi'' = \xi$.

Proof
◁ The proof follows the idea of the classical case of Theorems 4.5 and 5.1. It uses the relation

$$\int_0^1 \int \Gamma(P_{ts}(h_s - t\rho_s), P_{ts}(\rho_s))\xi(P_{ts}(\rho_s)) \, d\mu \, ds = \mathrm{Ent}_\mu^\Phi(P_t g) - \mathrm{Ent}_\mu^\Phi(f).$$

for any admissible path (ρ_s, h_s) between f and g, and in particular

$$\int_0^1 \int \Gamma(\rho_s, h_s)\xi(\rho_s) \, d\mu \, ds = \mathrm{Ent}_\mu^\Phi(g) - \mathrm{Ent}_\mu^\Phi(f),$$

and the following lemma. ▷

Lemma 6.4 *Let $f, g \in \mathcal{A}$ with $g > 0$ and let*

$$\Lambda(t) = \int \Gamma(P_t f)\xi(P_t g)d\mu$$

for $t \geqslant 0$. Then, under the curvature condition $CD(R, \infty)$,

$$\Lambda'(t) \leq -2R\Lambda(t) - \int \xi^2(P_t g)\left(-\frac{1}{\xi}\right)''(P_t g)\Gamma(P_t f)\Gamma(P_t g)d\mu \tag{41}$$

Proof

◁ We only briefly check the proof since it follows the one of Lemma 4.2. For any $t \geqslant 0$,

$$\Lambda'(t) = \int \left[2\Gamma(P_t f, L P_t f)\xi(P_t g) + \Gamma(P_t f)\xi'(P_t g)L P_t g\right]d\mu.$$

In the notation $G = P_t g$ and $F = P_t f$, the invariance property $\int L\big[\Gamma(F)\xi(G)\big]d\mu = 0$ and the diffusion property of L give that

$$\int \Gamma(F)\xi'(G)LG d\mu = -\int \left[2\xi'(G)\Gamma(\Gamma(F), G) + \xi''(G)\Gamma(G)\Gamma(F) + \xi(G)L\Gamma(F)\right]d\mu.$$

Hence, using the definition of the Γ_2 operator,

$$\Lambda'(t) = -2\int \xi(G)\left[\Gamma_2(F) + \Gamma(\Gamma(F), \log \xi(G)) + \frac{\xi''(G)\xi(G)}{2\xi'(G)^2}\Gamma(F)\Gamma(\log \xi(G))\right]d\mu$$

$$= -2\int \xi(G)\left[\Gamma_2(F) + \Gamma(\Gamma(F), \log \xi(G)) + \Gamma(F)\Gamma(\log \xi(G))\right]d\mu - \int \xi^2(G)\left(-\frac{1}{\xi}\right)''(G)\Gamma(F)\Gamma(G)d\mu.$$

Then Lemma 4.3 for $n = \infty$, applied to $f = F$ and $g = \log \xi(G)$, implies inequality (41). ▷

The particular case of power functions

Poincaré and logarithmic Sobolev inequalities belong to the family of Φ-entropy inequalities, namely for $\Phi(x) = x^2/2$ and $\Phi(x) = x \log x$ respectively (see [Cha04]). An interpolation family of inequalities between them consist in the Beckner inequalities, for $\Phi_p(x) = \frac{x^p}{p(p-1)}$. It has been proved in [AD05, BG10] how to refine these Beckner inequalities under the curvature-dimension condition $CD(R, \infty)$. In the same way, the contraction inequalities proved in [DNS12] and in (39) for a general Φ can be made more precise for these power functions, as follows.

For $p \in (1, 2)$ we let $\xi_p(x) = x^{p-2}$ and $\Phi_p(x) = \frac{x^p}{p(p-1)}$ for $x > 0$, so that $\Phi_p'' = \xi_p$.

Theorem 6.5 (Refined contraction inequality) *Let $(P_t)_{t\geqslant 0}$ be our diffusion Markov semigroup satisfying a $CD(R, \infty)$ condition with $R \in \mathbb{R}$. Then, for any $f, g \in \mathcal{F}$ and $t \geqslant 0$,*

$$T_{\xi_p}^2(P_t f\mu, P_t g\mu) \leq e^{-2Rt}T_{\xi_p}^2(f\mu, g\mu)$$

$$- 4\frac{2-p}{p^2(p-1)}\int_0^t e^{-2R(t-u)}\left(\sqrt{\int (P_u f)^p d\mu} - \sqrt{\int (P_u g)^p d\mu}\right)^2 du. \tag{42}$$

Proof

◁ Let (ρ_s, h_s) be an admissible path between f and g, and $\Lambda(t,s) = \int \Gamma(P_t(h_s))\xi_p(P_t(\rho_s))d\mu$. Then inequality (41) for h_s and ρ_s writes

$$\partial_t\Lambda(t,s) \leq -2R\Lambda(t,s) - (2-p)(p-1)\int (P_t(\rho_s))^{p-4}\Gamma(P_t(\rho_s))\Gamma(P_t(h_s))d\mu.$$

But

$$\int (P_t(\rho_s))^{p-4}\Gamma(P_t(\rho_s))\Gamma(P_t(h_s))d\mu \geqslant \frac{\left(\int (P_t(\rho_s))^{p-2}\Gamma(P_t(\rho_s), P_t(h_s))d\mu\right)^2}{\int P_t(\rho_s)^p d\mu}$$

by the Cauchy-Schwarz inequality, so

$$\partial_t\Lambda(t,s) \leq -2R\Lambda(t,s) - \frac{2-p}{p}\frac{[\partial_s\varphi(t,s)]^2}{\varphi(t,s)}$$

where

$$\varphi(t,s) = \frac{1}{p(p-1)}\int (P_t(\rho_s))^p d\mu.$$

Integrating over $s \in [0,1]$ and applying the Gronwall inequality in t, we obtain

$$\int_0^1 \Lambda(t,s)ds \leq e^{-2Rt}\int_0^1 \Lambda(0,s)ds - \frac{2-p}{p}\int_0^t\int_0^1 e^{-2R(t-u)}\frac{[\partial_s\varphi(u,s)]^2}{\varphi(u,s)}dsdu.$$

But

$$\int_0^1 \frac{[\partial_s\varphi(u,s)]^2}{\varphi(u,s)}ds \geqslant \left(\int_0^1 \frac{\partial_s\varphi(u,s)}{\sqrt{\varphi(u,s)}}ds\right)^2 = 4(\sqrt{\varphi(u,1)} - \sqrt{\varphi(u,0)})^2$$

again by the Cauchy-Schwarz inequality. The result follows by optimizing over (ρ_s, h_s). ▷

In the limit case where $p = 2$, the improvement in the contraction inequality disappears in (42), as observed in the refined Beckner inequalities of [AD05, BG10]. Morever this improvement goes to 0 when p goes to 1, hence recovering the classical contraction inequality (30)-(39) under the curvature condition $CD(R,\infty)$.

References

[ABC+00] C. Ané, S. Blachère, D. Chafaï, P. Fougères, I. Gentil, F. Malrieu, C. Roberto, and G. Scheffer. *Sur les inégalités de Sobolev logarithmiques*, volume 10 of *Panoramas et Synthèses*. Société Math. de France, Paris, 2000.

[AD05] A. Arnold and J. Dolbeault. Refined convex Sobolev inequalities. *J. Funct. Anal.*, 225(2):337–351, 2005.

[AGS08] L. Ambrosio, N. Gigli, and G. Savaré. *Gradient flows in metric spaces and in the space of probability measures.* Lectures in Math. ETH Zürich. Birkhäuser, Basel, 2008.

[AGS12a] L. Ambrosio, N. Gigli, and G. Savaré. Bakry-Emery-curvature-dimension condition and Riemannian Ricci curvature bounds. *Preprint*, 2012.

[AGS12b] L. Ambrosio, N. Gigli, and G. Savaré. Metric measure spaces with Riemannian Ricci curvature bounded from below. To appear in *Duke Math. J.*

[AMS14]   L. Ambrosio, A. Mondino, and G. Savaré. Nonlinear diffusion equations and curvature conditions in metric measure spaces. *Work in preparation.*

[Bak94]   D. Bakry. L'hypercontractivité et son utilisation en théorie des semigroupes. In *Lectures on probability theory (Saint-Flour, 1992)*, volume 1581 of *Lecture Notes in Math.*, pages 1–114. Springer, Berlin, 1994.

[BB00]   J.-D. Benamou and Y. Brenier. A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem. *Numer. Math.*, 84(3):375–393, 2000.

[BÉ85]   D. Bakry and M. Émery. Diffusions hypercontractives. In *Séminaire de probabilités, XIX, 1983/84*, Lecture Notes in Math. 1123, pages 177–206. Springer, Berlin, 1985.

[BG10]   F. Bolley and I. Gentil. Phi-entropy inequalities for diffusion semigroups. *J. Math. Pures Appl. (9)*, 93(5):449–473, 2010.

[BGG12]   F. Bolley, I. Gentil, and A. Guillin. Convergence to equilibrium in Wasserstein distance for Fokker-Planck equations. *J. Funct. Anal.*, 263(8):2430–2457, 2012.

[BGL01]   S. Bobkov, I. Gentil, and M. Ledoux. Hypercontractivity of Hamilton-Jacobi equations. *J. Math. Pures Appl. (9)*, 80(7):669–696, 2001.

[BGL13]   D. Bakry, I. Gentil, and M. Ledoux. On Harnack inequality and optimal transportation. To appear in *Ann. Sc. Norm. Super. Pisa Cl. Sci.*, 2013.

[BGL14]   D. Bakry, I. Gentil, and M. Ledoux. *Analysis and geometry of Markov diffusion operators*, volume 338 of *Grund. Math. Wiss.* Springer, Berlin, 2014.

[BGM10]   F. Bolley, A. Guillin, and F. Malrieu. Trend to equilibrium and particle approximation for a weakly selfconsistent Vlasov-Fokker-Planck equation. *Math. Mod. Num. Anal.*, 44 (5):867–884, 2010.

[BL06]   D. Bakry and M. Ledoux. A logarithmic Sobolev form of the Li-Yau parabolic inequality. *Rev. Mat. Iberoam.*, 22(2):683–702, 2006.

[CGM08]   P. Cattiaux, A. Guillin, and F. Malrieu. Probabilistic approach for granular media equations in the non uniformly convex case. *Prob. Theor. Rel. Fields*, 140(1-2):19–40, 2008.

[Cha04]   D. Chafaï. Entropies, convexity, and functional inequalities: on Φ-entropies and Φ-Sobolev inequalities. *J. Math. Kyoto Univ.*, 44(2):325–363, 2004.

[CMV06]   J. A. Carrillo, R. J. McCann, and C. Villani. Contractions in the 2-Wasserstein length space and thermalization of granular media. *Arch. Rat. Mech. Anal.*, 179:217–263, 2006.

[CT05]   J. A. Carrillo and G. Toscani. Wasserstein metric and large-time asymptotics of nonlinear diffusion equations. In *New trends in math. physics*. World Sci., Singapore, 2005.

[DNS09]   J. Dolbeault, B. Nazaret, and G. Savaré. A new class of transport distances between measures. *Calc. Var. Part. Diff. Eq.*, 34(2):193–231, 2009.

[DNS12]   J. Dolbeault, B. Nazaret, and G. Savaré. From Poincaré to Logarithmic Sobolev inequalities: a gradient flow approach. *SIAM J. Math. Anal.*, 44(5):3186–3216, 2012.

[EKS13]   M. Erbar, K. Kuwada, and K.-T. Sturm. On the equivalence of the entropic curvature-dimension condition and Bochner's inequality on metric measure spaces. *Preprint*, 2013.

[Gen13]   I. Gentil. Dimensional contraction in Wasserstein distance for diffusion semigroups on a Riemannian manifold. *Preprint*, 2013.

[GKO13]   N. Gigli, K. Kuwada, and S. Ohta. Heat flow on Alexandrov spaces. *Comm. Pure Appl. Math.*, 66:307–331, 2013.

[GL13]   N. Gigli and M. Ledoux. From log Sobolev to Talagrand: a quick proof. *Discrete Contin. Dyn. Syst.*, 33:1927–1935, 2013.

[JKO98]   R. Jordan, D. Kinderlehrer, and F. Otto. The variational formulation of the Fokker-Planck equation. *SIAM J. Math. Anal.*, 29(1):1–17, 1998.

[Kuw13]   K. Kuwada. Space-time Wasserstein controls and Bakry-Emery-Wang type gradient estimates. *Preprint*, 2013.

[LV09]   J. Lott and C. Villani. Ricci curvature for metric-measure spaces via optimal transport. *Ann. of Math. (2)*, 169(3):903–991, 2009.

[NPS11]   L. Natile, M. A. Peletier, and G. Savaré. Contraction of general transportation costs along solutions to Fokker-Planck equations with monotone drifts. *J. Math. Pures Appl.*, 95:18–35, 2011.

[Ott01]   F. Otto. The geometry of dissipative evolution equations: the porous medium equation. *Comm. Part. Diff. Eq.*, 26(1-2):101–174, 2001.

[OV00]   F. Otto and C. Villani. Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality. *J. Funct. Anal.*, 173(2):361–400, 2000.

[OW05]   F. Otto and M. Westdickenberg. Eulerian calculus for the contraction in the Wasserstein distance. *SIAM J. Math. Anal.*, 37(4):1227–1255, 2005.

[Stu06]   K.-T. Sturm. On the geometry of metric measure spaces. I and II. *Acta Math.*, 196(1):65–177, 2006.

[Tal96]   M. Talagrand. Transportation cost for Gaussian and other product measures. *Geom. Funct. Anal.*, 6(3):587–600, 1996.

[Vil09]   C. Villani. *Optimal transport, Old and new*, volume 338 of *Grund. Math. Wiss.* Springer, Berlin, 2009.

[vRS05]   M.-K. von Renesse and K.-T. Sturm. Transport inequalities, gradient estimates, entropy and Ricci curvature. *Comm. Pure Appl. Math.*, 68:923–940, 2005.

[Wan04]   F.-Y. Wang. *Functional Inequalities, Markov Processes, and Spectral Theory.* Science Press, Beijing, 2004.

[Wan11]   F.-Y. Wang. Equivalent semigroup properties for the curvature-dimension condition. *Bull. Sci. Math.*, 135(6-7):803–815, 2011.